XIAOCHI LI

george_li@gwu.edu | (571)225-1934 Arlington, VA 22202
LinkedIn: https://www.linkedin.com/in/xiaochi-li/ GitHub: https://xc-li.github.io/

EDUCATION

The George Washington University Washington, D.C., USA
Master of Science in Data Science, GPA:4.0 / 4.0 05/2019
Course: Machine Learning, Deep Learning, Design and Analysis of Algorithms, Natural Language Processing, Bayesian Methods, High Performance Computing and Parallel Computing

East China University of Science & Technology (ECUST) Shanghai, China
Bachelor of Science in Economics, GPA: 3.54 / 4.0 07/2017

TECHNICAL SKILLS

Proficient in Python (Pandas, Scikit Learn, TensorFlow, NLTK), R, Tableau, Linux Operation, SQL

DATA SCIENCE PROJECTS (Read detail on: https://xc-li.github.io/)

Machine Learning: Loan Default Prediction Spring 2018

- Led a 3-student group to build an end-to-end machine learning pipeline on 887K Lending Club data, and achieved 70% recall score on loan default prediction
- Applied feature engineering, over-sampling and fine-tuned various supervised models such as Random Forest, Logistic Regression with Scikit Learn to optimize performance

Deep Learning: Facial Expression Recognition Fall 2018

- Led a 3-student group to classify 12K facial expression images into 7 categories by developing a convolutional neural network (CNN) with TensorFlow and Keras on AWS, achieved 74% accuracy
- Built a real time facial expression recognition program with OpenCV3 and demonstrated in the class

Natural Language Processing: Yelp Review Sentiment Modeling Fall 2018

- Developed a sentiment classification model on 144K Yelp reviews with NLTK and Scikit Learn, and achieved 76% accuracy

Performed several tokenization and vectorization methods to improve model performance

WORK EXPERIENCE

TA for Bayesian Methods Washington, D.C., USA
GWU Data Science Program 08/2018 – 12/2018

- Tutored the students with programming assignments, and helped professor designing homework and quizzes

Business Intelligence Intern Shanghai, China
Michelin (China) Investment Co., Ltd. 07/2016 – 02/2017

- Self-designed a program using Python's Regulate Expression to auto-extract data information of tires, drastically reducing the work period from 10 weeks to 2-3 days
- Conducted data collection and analytical research to establish a company wide database for predicting sales figures and market shares of different types of tires