Xian Sun

Email: sunxian1107@gmail.com   Tel: +1 9843294985

EDUCATION

Duke University, Durham, North Carolina   **Aug 2019 - May 2021**
Master of Science in Computer Engineering   GPA: 3.97 (Max: 4.0)
Relevant Coursework: Computer Vision, Machine Learning, Deep Learning, Random Signal and Noise,
          Systems Programming & Engineering, Software Engineering

Jilin University, Changchun, China   **Aug 2015 - Jun 2019**
Bachelor of Science in Electrical Engineering   GPA: 3.90 (Max: 4.0)
Relevant Coursework: Advanced Mathematics, Probability and Statistics, Digital Signal Processing
          Embedded Systems, Signal and Systems, Linear Algebra

RESEARCH & PROJECT EXPERIENCE

Research Assistant, Computer Vision Lab, Duke University   **June 2020 - Present**
Advisor: Dr. Carlo Tomasi
- Built deep learning models to segment damages (semantic segmentation) on the leaf dataset whose imbalance ratio=400.
- Developed Leaf-Image-Editor to clean the Dataset and designed a copy/paste data augmentation algorithm to further improve the accuracy. Working on a two-stage model to improve the overall performance.

Research Intern, Deep Learning (DL) & ReRAM Research Group, Duke University   **Oct 2020 - Apr 2021**
Advisors: Dr. Krishnendu (Krish) Chakrabarty, Dr. Biresh Joardar
- Saved more than 80 percent of crossbars (128x128) for VGG and ResNet, via extremely sparse networks by iterative magnitude pruning.
- Designed one-shot pruning with the threshold, which enabled sparse networks (the sparsity < 2%) for AlexNet and VGG.

Research Assistant, Almost Matching Exactly Lab for Statistical Machine Learning, Duke University   **July 2020 - Apr 2021**
Advisors: Dr. Cynthia Rudin, Dr. Sudeepa Roy, Dr. Alexander Volfovsky
- Developed the Python package for Fast Large-scale Almost Matching Exactly on Database (FLAME-DB) Algorithm and R package for Adaptive Hyper Box Matching (AHB) Algorithm. Both packages support missing data handling, average treatment effect (ATE), average treatment effect on treated (ATT).
- Improved the speed of the FLAME-DB, with the combination of fixed and adaptive weight matching. The time Complexity is reduced to $O(N)$ from $O(N^2)$.

Improved Regularization of Convolutional Neural Networks, Duke University   **Dec 2020**
- Implemented Cutout and Mix-up to address overfitting problems for ResNet18, 34 and 50.
- Found the best combination of Cutout and Mix-up by grid search on ResNet18-CIFAR10 and transferred it to other datasets like SVHN and Fashion MNIST.

Exploring Probabilistic Classifiers on Binary Classification, Duke University   **Apr 2020**
- Implemented Bayes classifier with different assumptions regarding covariance structure of dataset and linear discriminant classifier and logistic discriminant to four different 2-dimensional datasets.
- Applied cross-validation to estimate classifier performance using the available training data and compared three classifiers and analyzed the advantages and disadvantages of them.

Software Development: RISK Game in Java, Duke University   **Apr 2021**
- Designed and implemented the game server and chat server with NIO and state machine.
- Handled concurrency requests with multi-thread programming from multiple users and server-client communication with TCP socket programming.

Software Development: Mini Linux Command Shell, Duke University   **Nov 2019**
- Designed a miniature version of Linux Command Shell functionally with C++
- Implemented command processing functions and functions to modify environment variables.

INDUSTRIAL EXPERIENCE

Machine learning Intern, Neocova, United States   **June 2020- Aug 2020**
- Built a linear regression model that calculates the percentage change in valuation of the model and a binomial logistic model that calculates the probability of positive/negative growth.
- Selected 5 most significant statistically variables with LASSO and 5 demographic variables with Decision Tree from more than 500 variables.

Electrical & Software Intern, Firmenich Aromatics (China) Co. Ltd   **Apr 2019 - Jun 2019**
- Designed and implemented a wireless scoring board that allowed blind participants in food assessment lab to rate various features of food products, that then sent data to a computer automatically.
- Improved the lab working efficiency by streamlining the data collecting process.

RELEVANT SKILLS

- Languages: Python, Java, C++, R, C, SQL, Markdown, basic HTML/CSS
- Frameworks: PyTorch, Scikit-learn, NumPy, Pandas, TensorFlow
- Software: Conda, Jupyter Notebook, PyCharm, IntelliJ, RStudio, PostgreSQL

TEACHING EXPERIENCE

- Instructor (Inspirit AI), AI Scholars and Deep Learning System Design (middle and high school)   2021
- Teaching Assistant to Dr. Carlo Tomasi for the graduate course CS527: Computer Vision   2021
- Teaching Assistant to Dr. Ivan Mura for the graduate course ECE650K: System Programming & Engineering   2021
- Teaching Assistant to Dr. Loren Nolte for the graduate course ECE581: Random Signal and Noise   2020
- Teaching Assistant to Dr. Elchanan Solomon for the undergraduate course Math216: Linear Algebra and Differential Equations   2020

HONORS & AWARDS

National
- China National Scholarship, China Ministry of Education   2017&2018

University
- ECE Merit Scholarship, Duke University   2020
- Valedictorian for 2019 Commencement, Jilin University   2019
- Top Ten Outstanding Student, Jinlin University (10 recipients, academics, service and leadership)   2019
- Outstanding Volunteer, Jilin University (>600 volunteer hours)   2018
- CASC Scholarship, China Aerospace Science and Technology Corporation (10 recipients, academics)   2018