Yanzhao Qian (Archibald)

A modern machine learning algorithm enthusiast who is problem-driven and data-oriented with three years of coding and data analytics expertise.

https://github.com/ArchibaldChain
+1(825) 984 6233
archibaldchain@gmail.com
www.linkedin.com/in/yanzhao-qian-997409185/

EDUCATION

University of Calgary — Calgary, AB, Canada
M.Sc. in Statistics (GPA: 3.925/4.00) — Sept. 2021 – Present

Southern University of Science and Technology — Shenzhen, China
B.Sc. in Mathematics and Applied Mathematics (GPA: 3.69/4.00) — Sept. 2017- June 2021

SKILLS

Programming: Python (NumPy, Pandas, Scikit-Learn, PyTorch, etc.), R, SQL, MATLAB, Java, C++
Tools: Bash, Microsoft Office, Git, Cloud Computing
Analytical Tools: Data Mining, GLMs, Bayesian Inference, Machine Learning/Deep Learning, Data Visualization

PROJECT EXPERIENCE

Statistical Analysis of Ookla Internet Speeds for Rural Canadian Communities — Calgary
Data Analyst — May 2022

- Analyzed Ookla internet speed dataset provided by Statistical Society of Canada.
- Uncovered the unbalanced development of Canadian internet between rural area and urban area using **Permutation Test**.
- Predicted the progress of internet development using **Logistic Regression** with 91% accuracy in the test data.

Data Analysis of Washington DC Crimes — Shenzhen
Team Leader — Apr. 2021 - May 2021

- Led a team of 3 members processing and analyzing the Data of Washington DC Crimes that happened from 2019 to 2021.
- Classified districts in Washington DC into low, medium, and high severity areas using **Kernel Density Estimation** and **K-means**.
- Visualized the crime severity in Python matplotlib.

Invasive Hornets Detection and Classification in Mathematical Contest in Modeling (MCM) — Feb 2022

- Used **YOLO-v5** to detect the insect in pictures.
- Augmented the unbalanced dataset and trained a **transfer-learning** convolutional neural network (CNN) using **ResNet-50** to classify the harmful hornet among bees' pictures with an accuracy of 96% in test dataset.
- Constructed a hornet detection system to help the local agriculture department trace and detect the spreading of the invasive pest.

RESEARCH EXPERIENCE

Cross-Validation Under-Estimated True Error of Machine Learning Algorithms in Genomic Data — Sept. 2021 - Present

- Developed a statistical structure using covariance between Training set and Test set to adjust the Cross-Validation Error.
- Tested corrected CV error for the models of **Linear Mixed-Effects Model (LMM), Best Linear Unbiased Predictor (BLUP) and Ridge** with simulated dataset from 1000 Genomes Project, which have closer estimation of generalization error (expectation of test error) than CV error.
- Trying to generalize the CV correction structure to the Bayesian model.

Particle Merging Method Development in Computational Mathematics — Nov. 2020 - May 2021

- Constructed a method of Moment Invariant Space to calculate the theoretical solution of particle merging problems.
- Developed a software that implemented the method above using MATLAB Symbolic toolbox.

TEACHING EXPERIENCE

Teaching Assistant in University of Calgary — Nov. 2020 – Present

- TA in **Statistical Modeling** for graduated Data Science students.
- TA in **Time Series Analysis**, **Probability & Statistics**, **Linear Algebra** and **Calculus** for undergraduate students.

Student Tutor — Feb. 2019 – June 2020

- Taught undergraduate students **Java** as a supplementary learning outside their classes.
- Helped to solve their problems in understanding coding in **Java** and the idea of object-oriented programming (**OOP)** and they all got A's in that course.