JONATHAN TRAN

https://github.com/JonathanDLTran
15359 NW Blakely Lane ⋄ Portland, OR ⋄ 97229
971-777-9511 ⋄ jdt98@cornell.edu

EDUCATION

Cornell University, College of Arts and Sciences, Ithaca, NY *Expected May 2021*
Bachelor of Arts - Double Major in Statistical Science and Computer Science
Cumulative GPA: 4.091 • Dean's List Three Semesters

RELEVANT COURSES

Machine Learning • Data Science • Linear Models • Probability Theory • Theory of Statistics
Python Programming • Linear Algebra • Functional Programming and Data Structures • Discrete Mathematics

WORK EXPERIENCE

Cornell University Lab of Ornithology - Data Analyst Jan 2019 - Present
· Evaluated accuracy of machine learning algorithm that tags elephant rumbles
· Employed specialized software such as Raven to classify specific sounds for the Elephant Language Project

CS 1110 / CS 2110 - Teaching Assistant August 2018 - Present
· Hosted weekly consulting office hours to allow for one on one coding and conceptual help for around 15 students
· Guided students towards assignment solutions and helped debug Python and Java coding mistakes
· Jointly graded labs, assignments and exams as part of a course staff serving 550+ students

National History Bee and Bowl - Contract Writer May 2017 - Present
· Crafted high quality and factually accurate questions that test a standard high school curriculum in history
· Ensured integrity and security of question distribution
· Collaborate with writing team and delivered one third of all questions for all Nationals events

PROJECTS

Index and Search Querying System October 2019
Implemented index and search crawler which traverses filesystem for .txt files and indexes words in all files
· Created Red Black Tree from scratch in order to allow for insertion, search and deletion in logarithmic time
· Allows user to search for any keywords, with the ability to filter using "and" and "or" keywords
· Tested OCaml code extensively to expose any errors, and also guarantee quick running time (2s) on over 1 million words

Data Visualization Contest December 2018
· Produced unique visualization comparing world versus American energy consumption
· Won Second Place during in-class competition

Modeling Crime Rate March 2019
Constructed multiple linear model in R using AIC and BIC criteria in order to evaluate connections between crime rate and predictors such as population and average resident age
Enhanced model by adding significant polynomial and interaction terms after consulting diagnostic plots to increase SSR

Classification on MNIST Data Set November 2018
· Implemented classifiers using logistic regression, support vector machine methods and unsupervised clustering methods
· Experimented with different parameters and kernels to raise classification rate on test images
· Trained a classifier with 98.06% success rate for hand written numerical digits from the test data set
· Extensively Utilized Python packages like csv, Pandas, Numpy and SciKit Learn to produce model

SKILLS AND ACTIVITIES

Languages	Python (experienced), R (proficient), OCaml (proficient), Java, Matlab
Software Packages	Git, Github, SciKit Learn, Pandas, Scipy, Numpy, Matplotlib, R Markdown, JMP
Activities	Cornell Statistics Applied Machine Learning Team, Ivy League Data Analytics Conference