Daniel(Chaofan) Tao

423 Towerview Dr., Durham, North Carolina

(+1) 919-884-1590 | chaofan.tao@duke.edu | www.danieltao.me | danieltao | chaofan daniel tao

Education

Duke University
Durham, NC

B.S of Computer Science (CS) *August, 2017-December, 2020*

- GPA: 3.9/4.0
- Named to the Dean's List (top third in the school) every semester
- **Courses:** Machine Learning (graduate level), Database, Operating Systems, Computer Architecture, Internet, Data Structure, Algorithms

Experience

Intersystems
Cambridge, MA

Data Platform Development Intern *May, 2019-August, 2019*

- Worked in the Core Development team on a project about **PMML** using ObjectScript
- Compared the IRIS PMML with standard JPMML with Java and Python, fixed over a dozen existing bugs in IRIS PMML
- Added functionality to support NN, SVM, Kmeans, Naive Bayes, and Random Forest in IRIS PMML with ObjectScripts
- Improved the runtime performance of IRIS PMML by over 30 percent by optimizing data input methods
- Worked on a industrial codebase that would influence real customers, wrote unit tests to ensure quality
- Used Swarm and Perforce for source control

Duke Prediction Analysis Lab
Durham, NC

Undergraduate Research Assistant *October, 2017-Present*

- Researched the application of interpretable NN in medical area with the help of prof. **Cynthia Rudin**
- Published ***This Looks Like That: Deep Learning for Interpretable Image Recognition*** on arXiv
- Applied PPNet to give a self-explained diagnosis of breast cancer mammograms, achieved **state-of-the-art** AUC
- Used existing attention NN and CNN models such as VGG, ResNet, RA-CNN, R-CNN, U-Net, etc.
- Used Tensorflow, Keras, Scikit-learn and Pytorch for the dataset of Breast Cancer
- Used GPU computing and data parallelism to speed up training

Duke Information Initiative
Durham, NC

Software Developer Intern *May, 2018-August, 2018*

- Built an **interactive web app** with Plotly to demonstrate and analyze datasets of single cell sequencing
- Built a pipeline to reduce dimension, cluster, and visualize single cell sequencing data
- Used TSNE, PCA, autoencoder, KMeans, and other deep learning clustering methods
- Achieved cluster accuracy of **76%** on testing dataset

Skills

Programming **Fluent**: Python, Java **Proficient**: JavaScript, C, Cpp, Assembly, Latex

Tools Git, Perforce, Maven, SQL

Web Dev. HTML, CSS, Bootstrap, React, Django

Languages English, Chinese

Projects

<Router Simulator>
Durham, NC

C *April, 2019*

- Built a virtual router using Routing Information Protocol (RIP).
- The router could perform tasks like ping, traceroute, etc. dynamically

<Duke AI for Art Competition>
Durham, NC

Python, Pytorch *February, 2019*

- Used Neural Style method to generate art works
- Implemented the Neural Style algorithm using tensorflow
- See https://github.com/danieltao/DukeAIforArt for some art pieces

<Stack Simulator>
Durham, NC

C *March, 2019*

- Programmed a stack manager simulator that simulate the behavior of how to manage stack in memory
- Used first fit, and performs coalescing and splitting correctly