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

Facebook *Menlo Park, CA*

Software Engineering Intern *May, 2020-August, 2020*

- Incoming SWE intern

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 **up to 800%** 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**
- 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

TOURTech *Raleigh, NC*

Contractor *December, 2019-April, 2020*

- Built a log anomaly detection model powered by machine learning with LSTM and invariance clustering
- Led and coordinated a team of five people

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

Publications

This Looks Like That: Deep Learning for Interpretable Image Recognition *Durham, NC*

NeurIPS 2019 — spotlight *2019*

- Proposed a new neural network architecture that uses prototypes to produce interpretable classification results
- Got accepted as a **spotlight (2 %)** paper

Skills

Programming	**Fluent**: Python, Java	**Proficient**: JavaScript, C, Cpp, Assembly, Latex
Tools & Web Dev.	**Tools**: Git, Perforce, Maven	**Web Dev.**: HTML, CSS, Bootstrap, React
Database	MySQL, PostgreSQL, MongoDB, XQuery	

Projects & Clubs

<DAML> *Durham, NC*

Co-Founder & Leader Of Data Science *September, 2019-Present*

- Started Duke Applied Machine Learning(dukeaml.com) group that applies machine learning in real life problems
- Customers we have: US Department of Defense, Duke VR center, Duke Woo Center, TOURTech, Big Spoon Roaster, REIN AI, etc.
- Lead data scientist teams consist of P.hDs, undergrads, and masters