Yao Ming

CYT 3007, HKUST
Clear Water Bay, Hong Kong
(+852) 5499 4004
(+86) 18666218906
yaoming.thu@gmail.com
www.myaooo.com

Education

Aug 2016 - Present **Ph.D. in Computer Science**
Hong Kong University of Science and Technology, Hong Kong

- Supervisor: Prof. Huamin Qu
- Research Interests: Interpretable/Explainable Machine Learning, Visualization.
- Cumulative Grade Average (CGA): 4.188/4.3
- Selected Courses: Advanced Algorithm Techniques, Machine Learning, Parallel Programming, Computer Vision, Advanced Statistics, Theory of Computation.

Aug 2012 - Jul 2016 **B.S. in Civil Engineering, B.S. in Economics**
Tsinghua University, Beijing, China

- Ranking: 1st of 93 students; Overall GPA: 93/100.
- Selected Courses (CS): Data Structure & Algorithms, Software Engineering, Database, Operating System, Advanced Computer Graphics.

Research / Work Experience

Jun 2019 - Present **Alibaba Damo Academy**, Hangzhou, China
Research Intern (advised by Dr. Hongxia Yang)

- Developing **interpretable user-commodity embedding algorithms** for large-scale recommendation systems. Planning to deploy and test the algorithms at Alibaba's online platforms.
- Programming: Python (Tensorflow & PyTorch), SQL (Hive).

Aug 2018 - Dec 2018 **Robert Bosch LLC**, Sunnyvale, CA
Research Intern (advised by Dr. Panpan Xu), Human Machine Interaction Group

- Developed **interpretable sequence learning** algorithm, ProSeNet (with PyTorch and Tensorflow), which provides interpretability while retaining the state-of-the-art performance.
- Applied the algorithm to *Vehicle Fault Prediction*, *Text Classification*, *Protein Classification*, and *Electrocardiogram (ECG) Classification*. Cleaned, processed, and prepared the dataset for benchmarking.
- Designed, developed, and deployed a web-based visual analytics system using React and Flask for understanding and steering ProSeNet. The system supports online fine-tuning and rich user interactions.
- ProSeNet was accepted by **KDD' 19**, the top conference in data mining for **oral presentation**.

Jan 2018 - Jun 2018 **New York University**, NY
Research Intern (advised by Prof. Enrico Bertini), VIDA Lab

- Developed a **model-agnostic** algorithm (in C and Python) that learns **surrogate rule lists** for explaining any classification models for *non-experts*. The algorithm achieves a fidelity of ~90% on different datasets.
- Designed and built rule-based explanatory visual interfaces, **RuleMatrix**, for understanding and analyzing machine learning models using Python (Keras and Scikit-Learn) and Javascript (React and D3).
- The system was deployed in an internal machine learning platform at NYU, presented and published on **IEEE VIS' 18** and **IEEE TVCG** (the top conference & journal in visualization).

Aug 2016 - Present **HKUST-WeChat Joint Lab on AI Technology**, Hong Kong
Lab Member

Project RNNVis (published on **IEEE VIS' 17'** | **Yelp Data Challenge Grand Prize Winner**)

- Proposed a technique for explaining the learned features of hidden units of RNNs using Tensorflow.
- Designed and implemented a web-based interface for understanding, exploring, and comparing the hidden memories of RNNs using Flask, MongoDB, and Javascript (Vue.js, D3.js).

Project ATMSeer (published on **ACM CHI' 18**)

- Participated as a major developer in ATMSeer, a user-centered auto-machine learning system, a joint project collaborated with the Data to AI Lab from MIT.
- Designed and implemented the APIs of the Auto-ML web service using Flask and SQL.

Skills

Programming	Proficient in Python, C++, and JavaScript/Typescript Familiar with MATLAB and Java
ML Tools	PyTorch, TensorFlow, Keras, Scikit-Learn
Web Dev	Flask, MongoDB, SQL, React, Node.js, Vue.js, D3.js

Selected Publications

Interpretable and Steerable Sequence Learning via Prototypes

Yao Ming, Panpan Xu, Huamin Qu, Liu Ren.

Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, 2019.

[***Top Conference in Data Mining*** | Accepted for **Oral Presentation**, 9% Acceptance Rate]

ProtoSteer: Steering Deep Sequence Model with Prototypes

Yao Ming, Panpan Xu, Furui Cheng, Huamin Qu, Liu Ren.

IEEE Transactions on Visualization and Computer Graphics, 2019 (accepted).

[***Top Journal in Visualization***]

ATMSeer: Increasing Transparency and Controllability in Automated Machine Learning

Qianwen Wang, **Yao Ming**, Zhihua Jin, Qiaomu Shen, Dongyu Liu, Micah J. Smith, Kalyan Veeramachaneni, Huamin Qu.

Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 2019.

[***Top Conference in Human-Computer-Interaction*** | 24% Acceptance Rate]

RuleMatrix: Visualizing and Understanding Classifiers using Rules

Yao Ming, Huamin Qu, Enrico Bertini.

IEEE Transactions on Visualization and Computer Graphics, 2018.

[***Top Journal in Visualization*** | 23% Acceptance Rate | Explainable ML]

Understanding Hidden Memories of Recurrent Neural Networks

Yao Ming, Shaozu Cao, Ruixiang Zhang, Zhen Li, Yuanzhe Chen, Yangqiu Song, Huamin Qu.

IEEE Visualization Conference (VAST), 2017.

[***Top Conference in Visualization*** | 24% Acceptance Rate]

Software

RuleMatrix A general rule-based machine learning model explanation tool featured with an interactive visualization interface in Jupyter Notebook. https://github.com/rulematrix/rule-matrix-py

PySBRL A C implementation of the Scalable Bayesian Rule List algorithm (supporting multi-class classification) with a python wrapper. https://github.com/myaooo/pysbrl

RNNVis A web-based app for explaining the hidden states of RNN models in Tensorflow. https://github.com/myaooo/rnnvis

Selected Honors

2019 SENG Top Research Postgraduate Award

2019 KDD Student Travel Award

2018 **Yelp Dataset Challenge Round 10 Grand Prize Award**

2016 - 2020 **Hong Kong PhD Fellowship (HKPF)**

2016 **Outstanding Graduate of Beijing**

2015 Nomination of **Tsinghua Top Talent Scholarship**

2014 **National Scholarship**

Invited Talks

Aug 2019 Interpretable and Steerable Sequence Learning via Prototypes
ACM KDD Conference, Anchorage, AK, U.S..

Oct 2018 Explainable Machine Learning via Surrogate Rules
Bay Area Visual Analytics Symposium, Sunnyvale, CA, U.S..

Oct 2018 RuleMatrix: Visualizing and Understanding Classifiers with Rules
IEEE VIS Conference, Berlin, Germany.

Oct 2017 Understanding Hidden Memories of Recurrent Neural Network
IEEE VIS Conference, Phoenix, AZ, U.S..

Services

Reviewer of
ACM Conference on Human Factors in Computing Systems (CHI), 2019
ACM Conference on Human Factors in Computing Systems (CHI) Late Breaking Work, 2019
IEEE Transactions on Visualization and Computer Graphics (TVCG), 2017, 2018
IEEE VIS (VAST, InfoVis, and SciVis) Conference, 2018, 2019
IEEE Eurographics/VGTC Conference on Visualization (EuroVis), 2019
IEEE Computer Graphics and Applications (CG&A) Magazine, 2018
China Vis, 2019

Last updated on: September 14, 2019