

SHAOJIE WANG

729 Richardson Rd. ♦ Rochester, N.Y. 14623

(+1) 585 210 5550 ♦ shaojie.wang@rochester.edu ♦ <https://github.com/Jossome>

EDUCATION

University of Rochester

September 2017 - Present

M.S. Computer Science

Advisor: Chenliang Xu

Renmin University of China

September 2013 - June 2017

B.S. Mathematics

Jingdong Scholarship first prize.

PUBLICATIONS

S. Wang, W. Zhao, Z. Kou, C. Xu. “*How to Make a BLT Sandwich?* Learning to Reason towards Instructional Video Understanding”. ArXiv 2018.

W. Zhao, **S. Wang**, Z. Xie, J. Shi, C. Xu. “GAN-EM: GAN based EM Learning Framework”. ArXiv 2018.

J. Zhang, **S. Wang**, S. Xu, M. Yu. Stock Price Manipulation Detection based on Machine Learning Technology: Evidence in China. In GRMSE 2016: Geo-Spatial Knowledge and Intelligence pp 150-158.

RESEARCH & PROJECT

How to Make a BLT Sandwich?

Learning to Reason towards Instructional Video Understanding

June - November 2018

First Author

In submission

- YouQuek (YouCook Question Answering) dataset for instructional video reasoning is proposed.
- A novel Recurrent GCN (RGCN) is proposed for learning both order and relation information.
- Multiple modalities are explored to model instructional features.

GAN-EM: GAN Based EM Learning Framework

April - November 2018

Second author

In submission

- We are the first to design general EM algorithm based on GAN, which highly enhances the learning ability of EM by relaxing the prior distribution assumption.
- A new loss for discriminator is designed to enable the framework to perform clustering and classification.
- GAN-EM gets state-of-the-art clustering and semi-supervised classification results on multiple datasets.

Attention Signals for Stock Market Forecasting Using Hilbert-Huang Transform and Incremental Learning

September 2016 - May 2017

B.S. Thesis

- Click ratio is proposed as a new indicator to model the attention of investors.
- Attention as signal is processed using Hilbert-Huang transformation for better performance.
- Incremental learning is applied to alleviate the problem of exploding data in stock market.

Even a Bumbling Bee can Attack Video Understanding Systems

Fall 2018

Leader

Work in progress

- A flying patch is proposed to perturb the video understanding tasks, e.g., action recognition.

- The patch size, moving region and speed are discussed, showing that temporal dimension helps reduce the patch size compared to adversarial patch for images.

Disguise Attack: An Extreme Test on Adversarial Examples

Spring 2018

Leader

- A new untargeted attack method “disguise attack” is proposed to minimize the classification probability for the real class label along with the difference between the modified pixel and its surroundings.
- YUV color space is applied on calculating the pixel-wise distance to fool human eyes.

Cache-Conscious Concurrent HAMT in Rust

Spring 2018

Contributor

- Both blocking and lock-free HAMT to achieve concurrency are implemented.
- The cache-conscious implementation without concurrency performs much better than the Rust official `HashMap` when the size is large.
- The lock-free version of cache-conscious HAMT implementation has good locality and very low miss rate, and is also scalable.

Voyager for Exoplanets: Mining on Imbalanced Data

Fall 2017

Leader

- A novel system classifying exoplanets and habitable exoplanets classification is constructed.
- An ensemble model of SVMs with under sampling is proposed to deal with extremely imbalanced data.
- For habitable exoplanet classification, we have recall of 92.3%, precision of 85.7%, and AUC of 88.5%.

EXPERIENCE

Lab of Information System Application, RUC, Beijing

September 2015 - June 2017

Volunteer Researcher

- Participated in the project “Prediction of exhibition influence” cooperating with Jiguang, mainly working with semantic analysis on Weibo.
- Participated in “Book sales prediction”, using SVM for regression task.

Tianyu Online Mart, Beijing

August 2014 - January 2017

Co-founder; Data Analyst

- Founded the company cooperating with a classmate, committed to building an online supermarket with the fastest delivery in campus.
- Programmed a customer data analysis system using Python for effective management.
- Developed location based logistic strategy among multiple dormitory buildings on campus.

RELEVANT COURSES

CS Courses

Advanced Topics in Computer Vision
Machine Learning
Parallel and Distributed System
Algorithm

Math Courses

Probability Theory
Stochastic Process
Functional Analysis
Optimization

LANGUAGES & SKILLS

Proficient

Python, C

Familiar

OCaml, C++, Rust, Java, Ruby, L^AT_EX, Matlab, SQL, JavaScript

Preferences

PyTorch, Vim