Zifeng Wang 王子丰

wangzf18@mails.tsinghua.edu.cn

EDUCATION BACKGROUND

Tsinghua University — **Shenzhen, China**
M.S., Data Science, Tsinghua-Berkeley Shenzhen Institute (TBSI) — *Sept. 2018-Present*
Major course: *Machine learning, Computer vision, Information theory*
Co-advised by: Prof. Shao-Lun Huang, TBSI and Prof. Khalid M. Mosalam, UC-Berkeley

Tongji University — **Shanghai, China**
B.Eng., Structural Engineering, School of Civil Engineering — *Sept. 2014-Jun. 2018*
GPA: 4.4/5.0 (19/168); Advised by: Prof. Suzhen Li

PAPERS

- **Zifeng Wang**, Xi Chen, Rui Wen, Shao-Lun Huang and Ercan E. Kuruoglu. *Information Theoretic Counterfactual Learning from Missing Not At Random Feedback.* **In progress**.
- **Zifeng Wang**, Xi Chen, Rui Wen and Shao-Lun Huang. *Finding Influential Instances for Distantly Supervised Relation Extraction.* **In progress**.
- **Zifeng Wang**, Hong Zhu, Zhenhua Dong, Xiuqiang He and Shao-Lun Huang. *Less Is Better: Unweighted Data Subsampling via Influence Function.* **AAAI 2020**. [pdf][code]
- **Zifeng Wang**, Yuyang Zhang, Khalid M. Mosalam, Yuqing Gao and Shao-Lun Huang. *Deep Fusion Network with RGB-Depth Image for Pixel-level Semantic Segmentation on Construction Sites.* **Automation in Construction**. (under review)
- **Zifeng Wang** and Suzhen Li. *Data-driven Risk Assessment on Urban Pipeline Network Based on a Cluster Model.* **Reliability Engineering and System Safety**.

INDUSTRY EXPERIENCE

Jarvis Lab, Tencent — **Shenzhen, China**
Research intern in Machine learning and NLP — *Dec. 2019-Present*

- Research about counterfactual learning from missing-not-at-random data for recommender system, based on information bottleneck theory.
- Research in instances selection in distant supervision via an influence subsampling paradigm.
- Research in heterogeneous GCN for automatic disease diagnosis.

Noah's Ark Lab, Huawei — **Shenzhen, China**
Research intern in Recommender system — *Apr. 2019-Oct. 2019*

- Join in counterfactual learning for improving ad-click rates by exploiting non-displayed events, and publish in CIKM 2019.
- Join in one class field-aware factorization machine for queries recommendation with implicit feedback, submit to JMLR.
- Lead research in guiding data subsampling for better model with less data, based on influence function and robust supervised learning theory, which is applicable for highly sparse data ($>$ 10M dims), publish in AAAI 2020 (ML track, score 8,7,6; accept rate: 18%).

TEACHING

- TA for Bayesian Learning and Data Analysis in TBSI, given by Ercan E. Kuruoglu — *Spring, 2020*
- TA for Machine Learning in TBSI, given by Shao-Lun Huang and Yang Li — *Fall, 2019*

PROJECTS

Smart Robot Development [code] — **Shenzhen, China**
Work with Prof. Shao-Lun Huang and Prof. Lin Zhang, in TBSI Lab 2C — *Feb. 2019- Aug. 2019*

- Develop face recognition module on MTCNN and pretrained InceptionV1 for 1:N face verification.
- Develop indoor object detection module based on YOLO-V3, and object tracking module on Siamese FCN.

AWARDS & ACHIEVEMENTS & OTHERS

- Reviewer for 23rd IEEE International Conference on Intelligent Transportation Systems (ITSC) — *Apr. 2020*
- Best Student Research Runner-up, in 2019 TBSI workshop in Data Science — *Dec. 2019*
- Outstanding graduate student (4/40), graduate thesis (3/168) of Tongji University — *Jun. 2018*
- Merit student scholarship of Tongji University — *2015/2016/2017*
- Kaggle: KKBox Music Recommendation Challenge (151/1081,15%) — *Dec. 2017*
- Meritorious winner (1st class prize,$\approx$ 7%) in USA Mathematical Contest in Modeling — *Apr. 2017*

SKILLS & CERTIFICATION

- English: CET-6 (615), IELTS (7.0)
- IT: Linux, Python, C++ and Python packages including Pytorch, Tensorflow, Numpy, Scipy, Pandas, Sklearn, keras, etc.