Runpeng Yu

Home Page: https://yu-rp.github.io GitHub: yu-rp E-Mail: yrp19@mails.tsinghua.edu.cn
Address: Tsinghua University, University Town of Shenzhen, Nanshan District, Shenzhen, 518055, China

Education Background

Tsinghua University Shenzhen, China
MS, Data Science, Tsinghua-Berkeley Shenzhen Institute Sept. 2019 – Present
Courses: Statistical Learning, Bayesian Learning, Large Deviation Theory, Information Theory
GPA: 3.93 / 4.0

Tianjin University Tianjin, China
BMS, Information Management and Information Systems, College of Management and Economics Sept. 2015 – July 2019
Courses: General Artificial Intelligence, Data warehouse Technology, Data Mining
GPA: 3.86 / 4.0 (top 3%)

Nankai University Tianjin, China
BEcon, Finance (Minor), School of Finance Sept. 2016 – July 2019

Publications

1. **Runpeng, Y.** *et al.* Convergence and Robustness of Gradient Orthogonalization for Out-of-Distribution Generalization. *Under Review of CVPR* (2022).
2. **Runpeng, Y.** *et al.* Regularization Penalty Optimization for Addressing Data Quality Variance in OoD Algorithms. *Accepted by AAAI* (2022).

Research Experience

Noah's Ark Lab, Huawei Shenzhen, China
Research intern in machine learning and recommendation system Aug. 2020 - Aug. 2021

- On the Noisy and Long-tailed Training Data for OOD Algorithm
  - Out-of-Distribution (OOD) algorithms address the performance degradation when the training and testing data are not IID. We observe that the noisy and long-tailed training data significantly compromise OOD algorithms.
  - We propose RPO which integrates domain-wise and sample-wise reweighting to address this problem.
  - We derive the optimal weighting scheme by transforming the optimization of the neural network to an ordinary differential equation and solving it by the Green function.

- On the Orthogonality Constraint for OOD Algorithm
  - Orthogonality constraint (OC) has been proved empirically, without theoretical guarantee, promising for addressing the distribution shift between the training and testing data introduced by the non-causal feature.
  - Based on the neural tangent kernel theory, we prove the sufficiency of OC for filtering out the non-causal correlation in the training data.
  - We propose an updated version of OC and experimentally elaborate its capability of recognizing and removing the non-causal feature by latent feature visualization and input reconstruction.

- A Classification Framework of OOD Benchmarks and Algorithms
  - Recently, a plethora of OOD algorithms and OOD benchmarks have been proposed. Most of these algorithms perform well on part of the benchmarks but fail on the other part. We attribute this to the fact that there are two types of distribution shifts and OOD algorithms specialize in only one of them.
  - We propose the formal definitions of these two types of distribution shifts and the algorithms for empirically evaluating them.

Master Thesis Aug. 2021 - Present

- Low-bit-rate Image Compression with Semantic Information
  - It has been identified that the generative adversarial network (GAN) promotes the reconstruction quality of low-bit-rate image compression. We observe that GAN, though improves the perception metrics of reconstruction, omits important semantic information in the input.
  - We propose a semantic reconstruction loss for preserving the semantic label and obtaining good reconstruction quality, simultaneously.

Teaching Experience

Teaching Assistant for **Random Process** (32 hrs) given by Prof. Ercan Engin Kuruoglu. Fall, 2021
Teaching Assistant for **Random Process** (32 hrs) given by Prof. Ercan Engin Kuruoglu. Spring, 2021
Teaching Assistant for **Time Series Analysis** (32 hrs) given by Prof. Ercan Engin Kuruoglu. Fall, 2020

- I am responsible for marking assignment, answering question, holding presentation and organizing oral exam.

Awards

Scholarship of Tsinghua Shenzhen International Graduate School, Second Prize 2020
National Undergraduate Scholarship 2017
Merit Student of Tianjin University 2017

Services in Scientific Community

- Reviewer of ICASSP (2022) and CVPR (2022).
- Assist in organizing
  - The 2nd TBSI Workshop on Learning Theory 2020
  - The First China RISC-V Forum 2019
  - International Conference on Smart Manufacturing, Industrial & Logistics Engineering 2017
  - International Symposium on Semiconductor Manufacturing Intelligence 2017

Skills & Certifications

- Language: TOEFL (103), GRE (327+4)
- Programming: Python (Pytorch, Tensorflow, Numpy, Scipy, etc), Matlab, R, C++, C#, HTML, SQL
- Hobbies: Badminton, Swimming

Volunteer Experience

Qifeng Central Primary School Dali, Yunnan, China
Volunteer Teacher, Team Leader July 2017 – Aug. 2017