Learning To Teach

11/6/18

ICLR 2018

By: Yang Fan, Fei Tian, Tao Qin, Xiang-Yang Li, Tie-Yan Liu

Presented by: Ronald Feng

UNIVERSITY OF WATERLOO
FACULTY OF MATHEMATICS

Overview

1. Intuition and Context
2. Framework Definition
    - Student Model
    - Teacher Model
3. Application to Data Teaching
4. Experimental Results
5. Critique

UNIVERSITY OF WATERLOO
FACULTY OF MATHEMATICS

Intuition and Context

- Teaching is a fundamental aspect of education systems
  - Self-learning is generally slower
- Current research focus in AI is on the *learner*
- L2T framework provides a <u>conceptual basis</u> for a system of *Teacher* and *Student* within a machine learning setting

L2T

- L2T framework consists of two intelligent agents:

    1. **Student model**: the “learner” in traditional ML algorithms

    2. **Teacher model**: goal to maximize speed and/or accuracy of student

- Once trained, teacher model generalizable

UNIVERSITY OF WATERLOO
FACULTY OF MATHEMATICS

Student Model

- Consider the supervised learning setting
- Student model takes input data and supervisor labels
  - Estimate a function, $f_\omega(x)$, which optimizes prediction of supervisor labels according to a given loss function L

$$\omega^* = \arg\min_{\omega \in \Omega} \sum_{(x,y)\in D} L(y, f_\omega(x)) \triangleq \mu(D, L, \Omega).$$

Teacher Model

- Goal is to improve student learning efficiency through modifying the following:
    1. D : Input Data
    2. L : Loss function
    3. Ω : Hypothesis Space

$$\min_{D,L,\Omega} \mathcal{M}(\mu(D, L, \Omega), D_{test}).$$

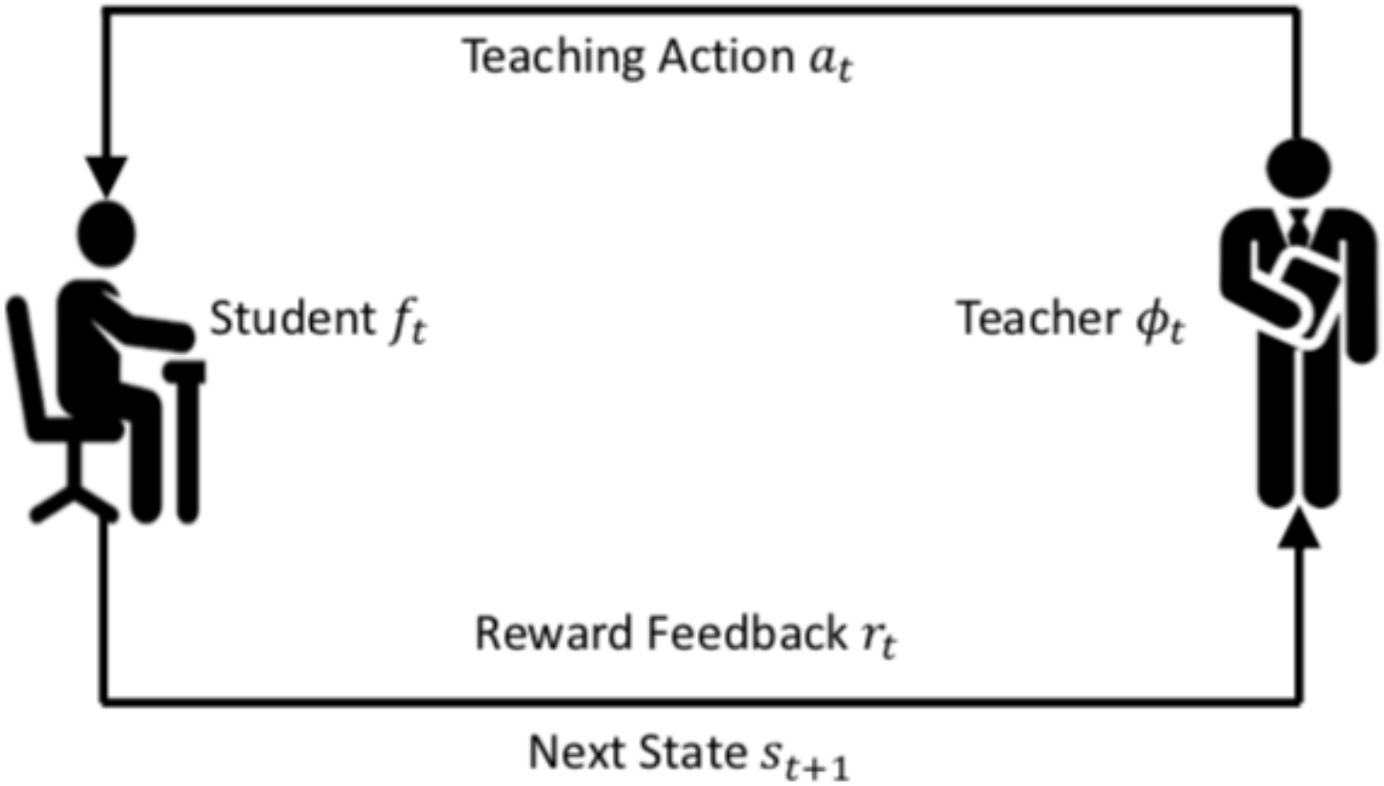
UNIVERSITY OF WATERLOO
FACULTY OF MATHEMATICS

Application: Data Teaching

- Data teaching: fixing loss function and hypothesis space
- Teaching model determines input data
- Teaching training approach: Reinforcement Learning
  - Reward is student convergence speed (maximize)
    - Given accuracy threshold, τ
  - Policy is teacher model action (learn parameters θ)

$$\max_{\theta} \sum_{t} r(\mu(\phi_{\theta}(s_t), L, \Omega)),$$

Application: Data Teaching



Application: Data Teaching

- Train teacher model by maximizing expected reward:

$$J(\theta) = E_{\phi_\theta(a|s)}[R(s,a)]$$

- Optimizer: REINFORCE (Williams, 1992)
  - Likelihood ratio policy gradient algorithm
  - Estimated empirically:

$$\nabla_\theta \approx \sum_{t=1}^{T} \nabla_\theta \log \phi(a_t|s_t) v_t$$

UNIVERSITY OF WATERLOO
FACULTY OF MATHEMATICS

Application: Data Teaching

- Student model learning rule: Mini-batch Stochastic Gradient Descent
  - Training data arrives in batches, sequentially in random order:

$$\{D_1, \cdots D_t, \ldots\}$$

  - Each mini-batch consists of *M* training instances

$$D_t = (d_1, \cdots, d_M)$$

  - Teacher determines which training instances to give students

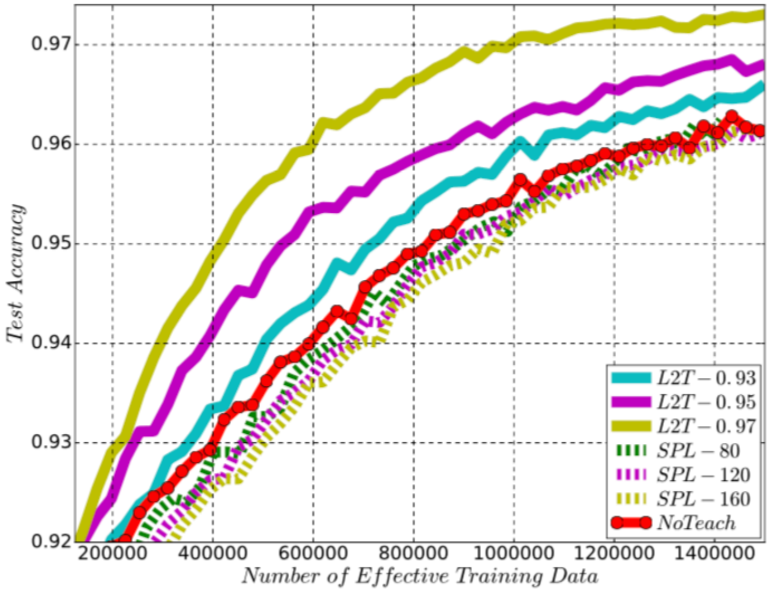
Experimental Results

- Tasks:
  1. MNIST: Image Classification
  2. CIFAR-10: Image Classification
  3. IMDB: Sentiment Classification
- Students:
  1. Multilayer Perceptron (MLP)
  2. Convoluted Neural Network (CNN): ResNet32 and ResNet110
  3. Recurrent Neural Network (RNN)
- Situations:
  1. Teaching new student with **same model architecture**
  2. Teaching new student with **different model architecture**

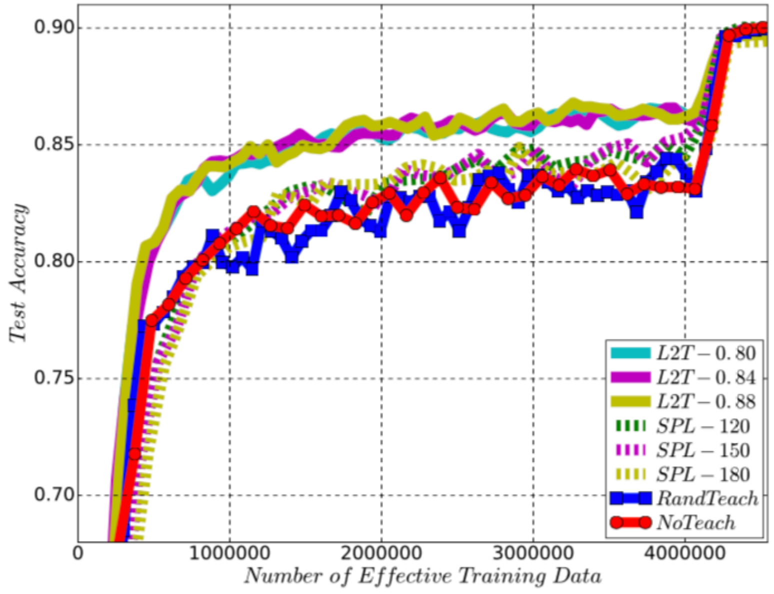
Experimental Results

- Teaching Strategies:
  - **L2T**
  - **NoTeach**
  - **Self-Paced Learning (SPL):** training student by data “hardness”
    - Hardness = Loss Value, Large loss = “hard”
    - Initially filter out “harder” data, slowly increase threshold
  - **RandTeach: Data instances are randomly filtered at each epoch**
    - Data-teaching baseline

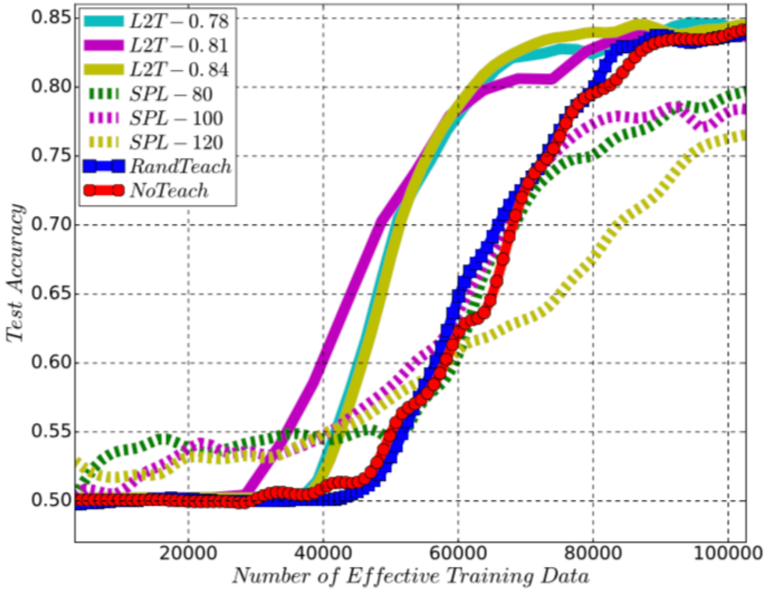
Convergence Speed



(a) MNIST

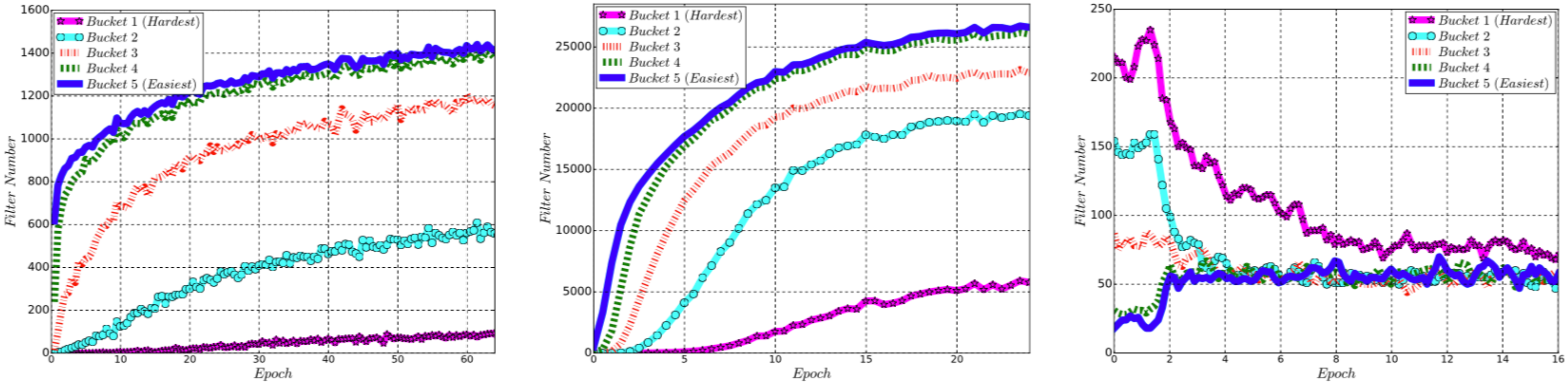


(b) CIFAR-10



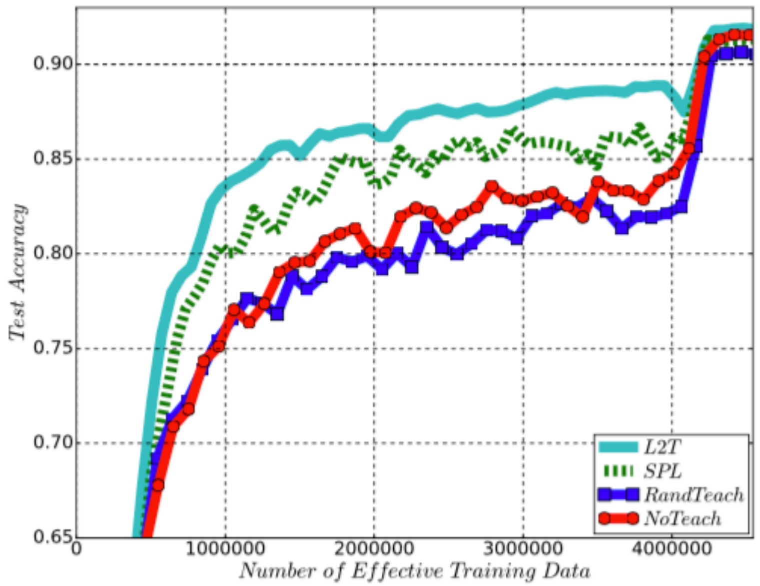
(c) IMDB

Analysis of Filtered Data



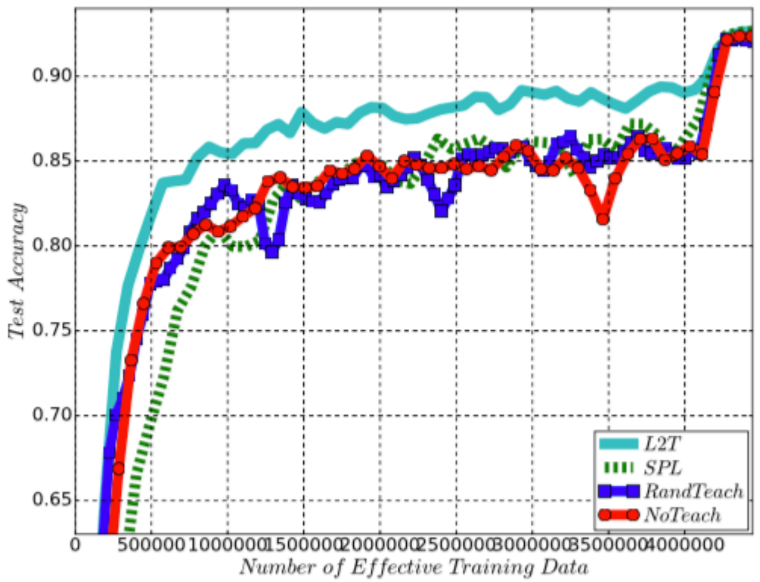
(a) MNIST (b) CIFAR-10 (c) IMDB

Teaching New Student: Different Architecture



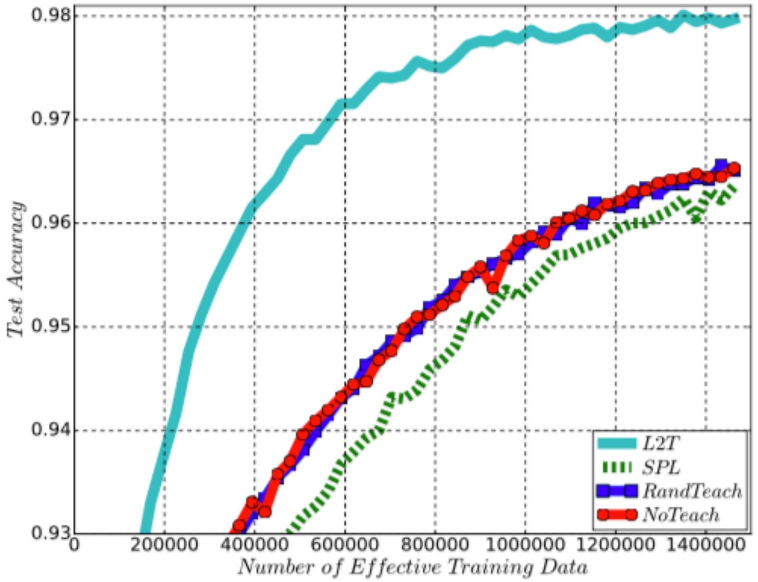
(a) ResNet32 → ResNet110

CIFAR-10



(b) MNIST → CIFAR-10

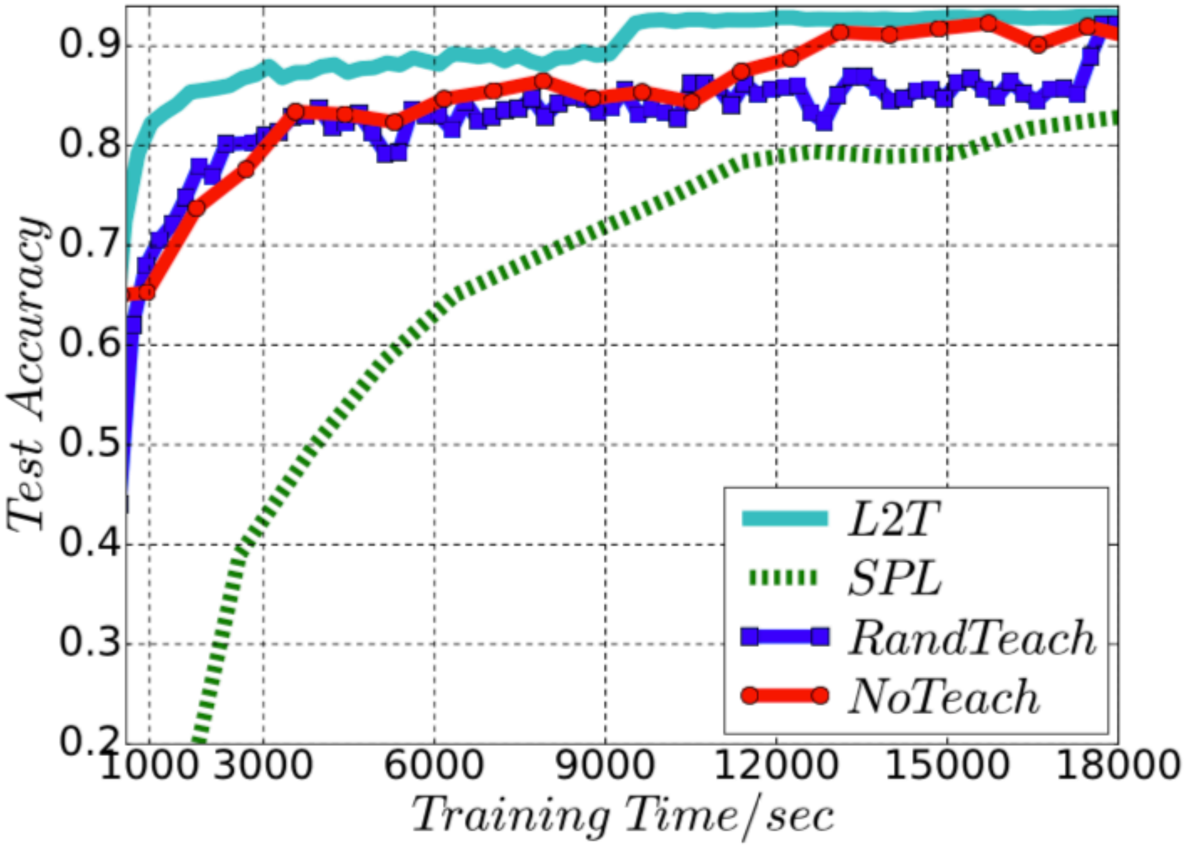
MLP -> ResNet32



(c) CIFAR10→MNIST

ResNet32 -> MLP

Wall-Clock Training Time



- Teaching ResNet32 on CIFAR-10
- Fastest training time compared to benchmark teaching strategies
  - Despite needing to train the teacher in tandem

Critique

QUESTIONS

Thank you for your time