# Learning To Teach

**Yang Fan**[†]*, **Fei Tian**[‡]*, **Tao Qin**[‡], **Xiang-Yang Li**[†] **& Tie-Yan Liu**[‡]
[†]School of Computer Science and Technology, University of Science and Technology of China
fyabc@mail.ustc.edu.cn, xiangyangli@ustc.edu.cn
[‡]Microsoft Research
{fetia,taoqin,tyliu}@microsoft.com

ICLR-2018

# Introduction

- In human society, a good teacher will select teaching material according to student‘s level.
- In the field of machine learning. We can also feed appropriate data to the model such that it can improve faster.

--How to define appropriate? For example, curriculum learning.

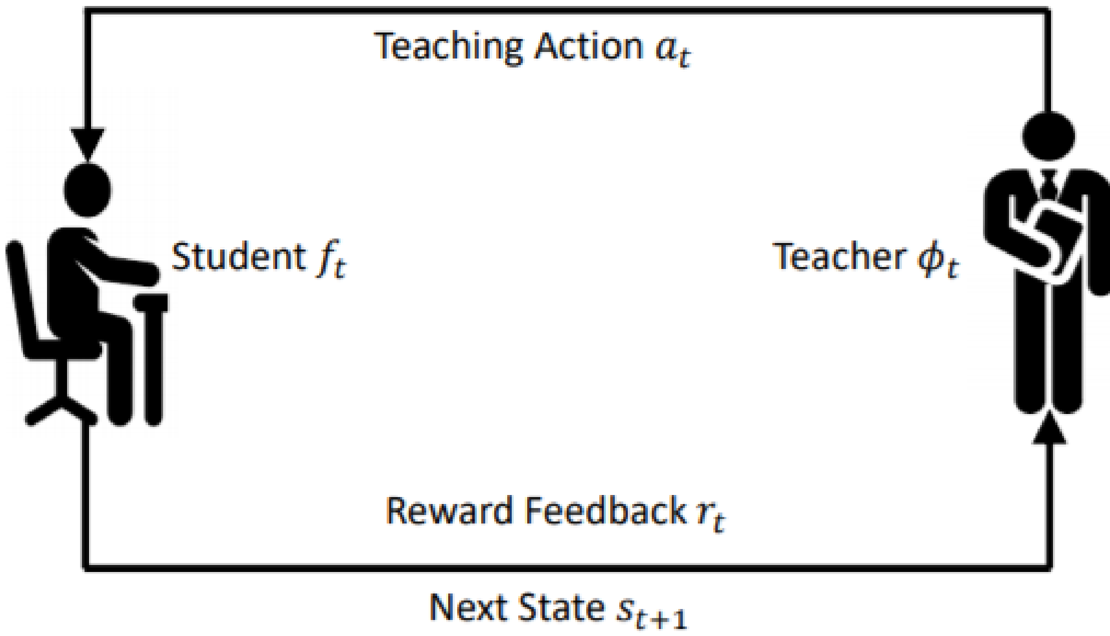
--However, data selected in this way may not best improve the model.

- We can train a teacher instead of designing criteria.

# Framework-based on RL

First, train a teacher using student A. Then, use teacher model to teach student B. Data with highest reward will be teach.

**Advantage:**
Receive information after the model changes. Instead of guessing which data is more helpful. Thus information is more accurate.
**Disadvantage:**
--The result depends on the quality of teacher.
--A good teacher to student A may not be a good teacher to student B. models are different.



# RL setting

State Features.

- Data features. (for texts)length of sentences, linguistic features.(for images) gradients histogram features.
- Student model features, signals reflecting how well current model is. Passed mini-batch number, the average historical training loss.
- Combination of both data and learner model, how important the arrived data is for current learner. Such as predicted probabilities of each class.

Action is denoted as $a = \{0,1\}^M$, M is the length of arrived batch.

Reward reflects how fast the student model learns. So $r_t < 0, \forall t < T$. $r_T = -\log(i_\tau/T')$. $i_\tau$ is the index of the first batch when the accuracy exceeds $\tau$, where T' is pre-defined maximum iteration number

# Experiment Setup

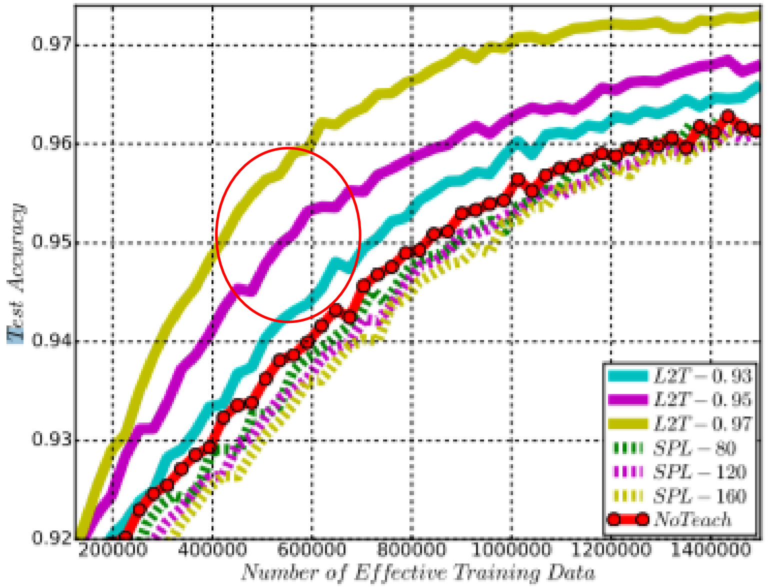
Different teaching strategies

- NoTeach. Conventional machine learning process.
- Self-paced Learning. Teaching data from easy to hard.
- Learning to teach. The teacher model in L2T framework.
- RandTech. Random select data.

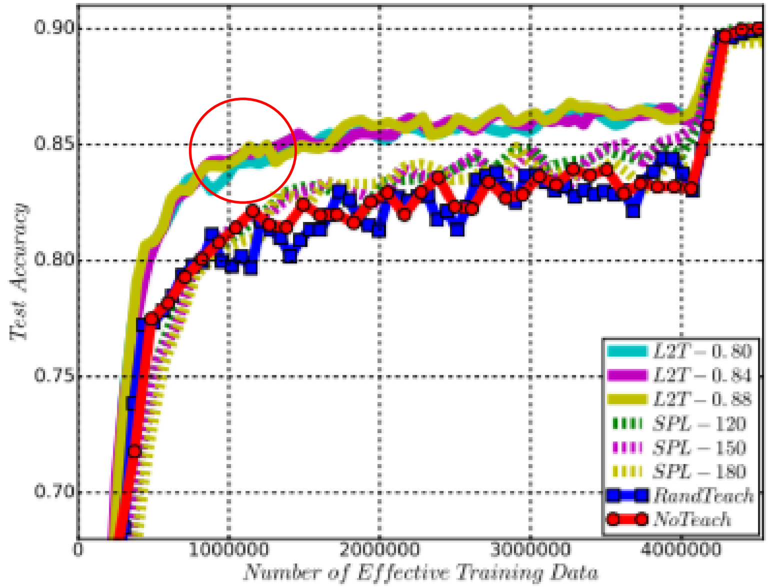
To test generalization ability of the teacher model learnt, consider two settings:

- Teaching a new student with the same model architecture. Train the teacher model using a student model. Then fixed the teacher model to train a student model with the same architecture.
- Teaching a new student with different model architecture. Two student models are of different architecture.

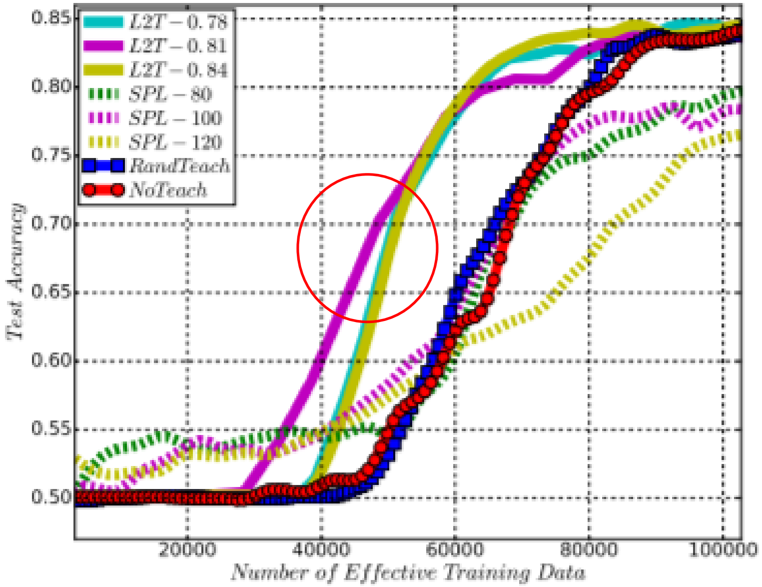
# Same architecture




(a) MNIST



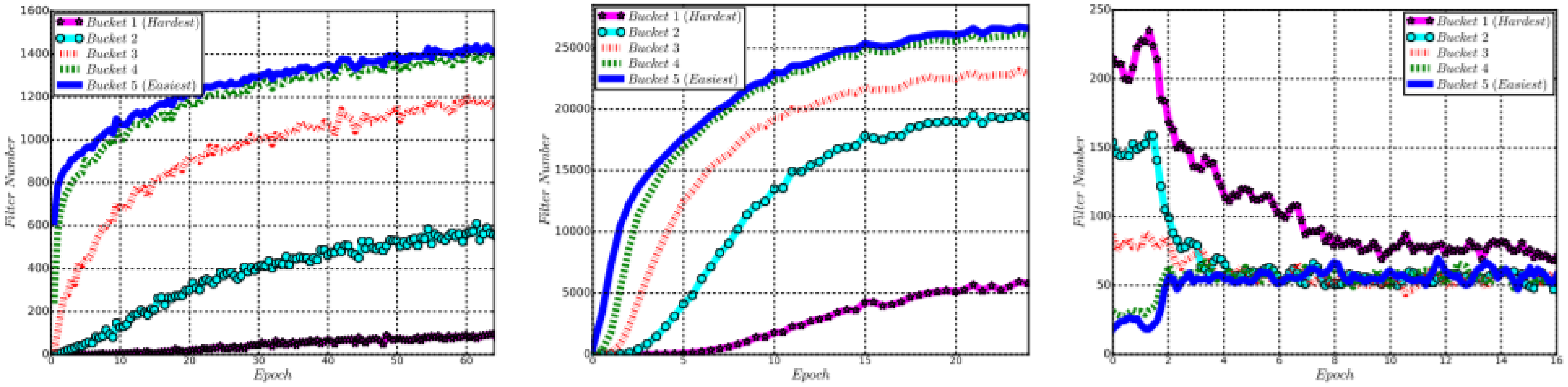

(b) CIFAR-10




(c) IMDB

X: number of effective training data.
Y: test accuracy.

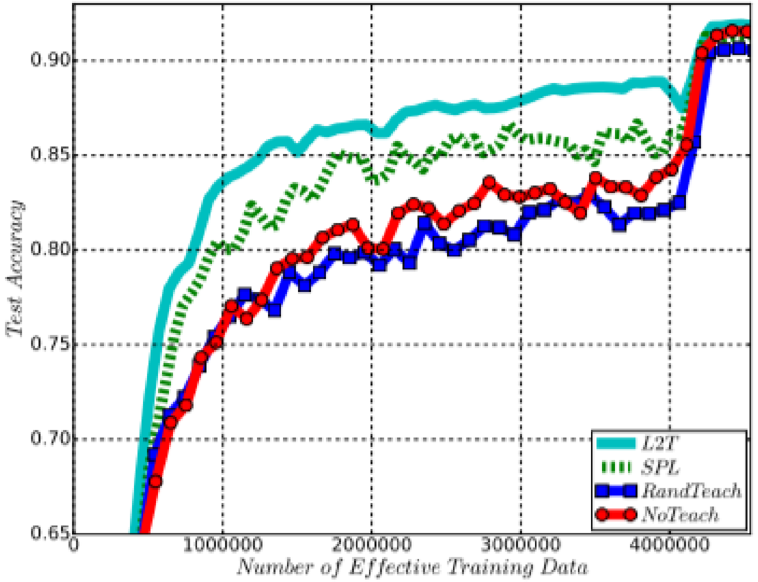
# Filtration number analysis



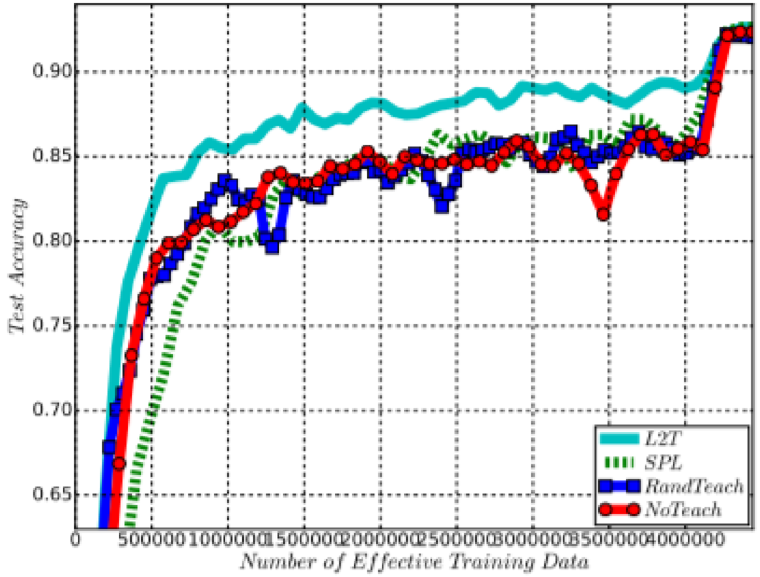

(a) MNIST (b) CIFAR-10 (c) IMDB

Different colors reflect different hardness levels.
In a and b, teacher favors harder data. In c, teacher is similar to SPL.
In IMDB, harder instances are more likely to contain noise.

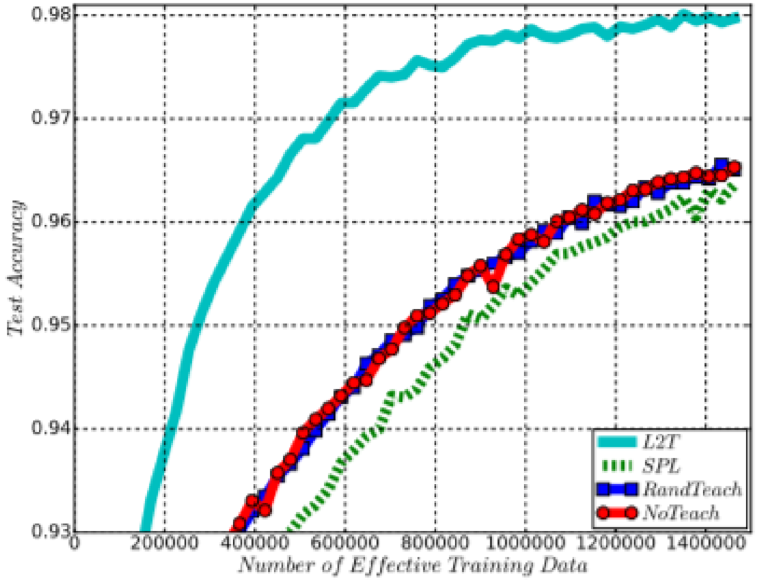
# Different architecture




(a) ResNet32 → ResNet110




(b) MNIST → CIFAR-10




(c) CIFAR10→MNIST

- A: Apply the teacher trained based on ResNet32 to teach ResNet110 on CIFAR-10.
- B: Apply the teacher trained based on MLP for MNIST to train CNN for CIFAR-10.
- C: Apply the teacher trained based on CNN for CIFAR-10 to train MLP for MNIST.

# Accuracy improved

| Teaching Policy | NoTeach | SPL | L2T |
|---|---|---|---|
| Accuracy | 88.54% | 88.80% | **89.46%** |

IMDB dataset

- NoTeach. Conventional machine learning. Converged
- SPL. Self-paced learning. Converged
- L2T. Train teacher using half of the dataset. Then teach student model using  full dataset until model converge.