Active Multi-task Learning via Bandits

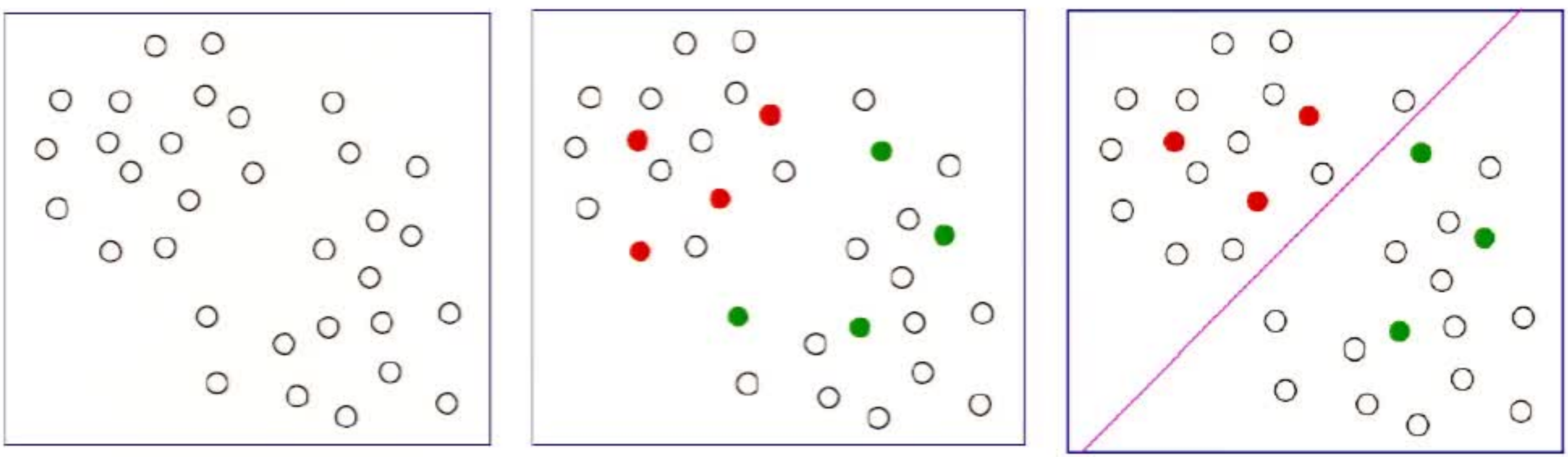
Meng Fang and Dacheng Tao
University of Technology, Sydney

Outline

- Motivation
- Connection between Active learning and Bandit
- Algorithm
- Experiments
- Conclusion

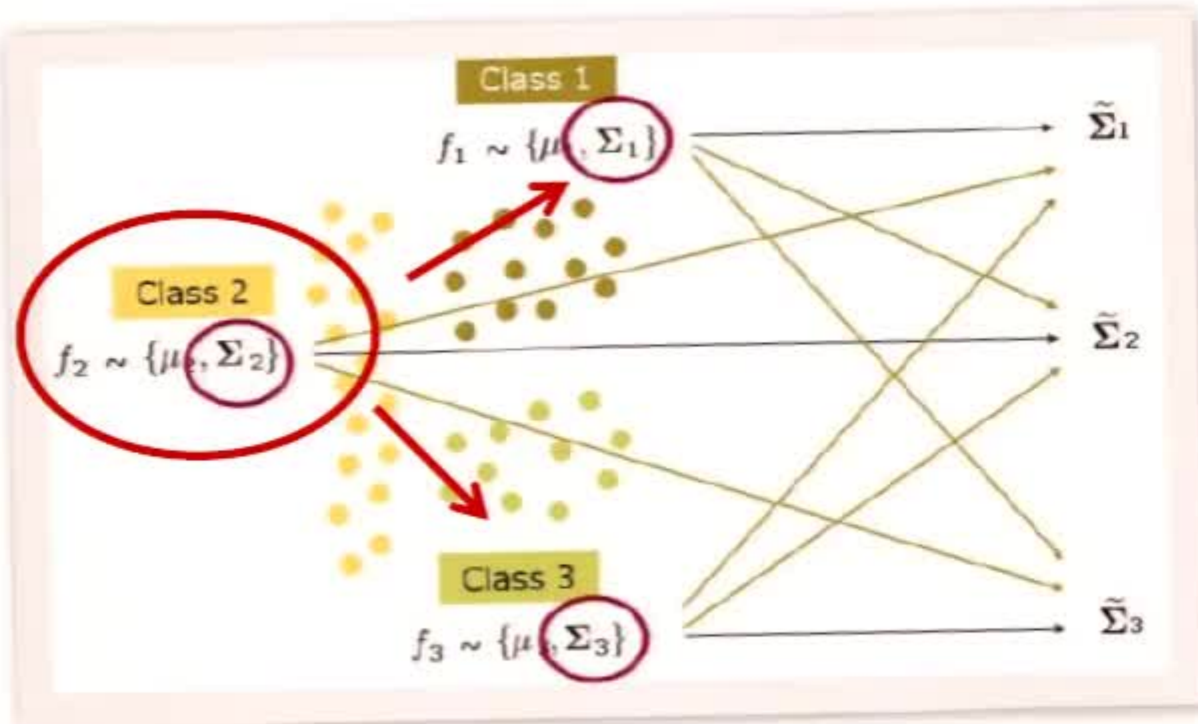
Motivation

- Labeling is expensive. The most time-consuming and costly part is usually the collecting of data.



Motivation

- Labeling instances for one task can also affect the other tasks especially when the task has a small number of labeled data.



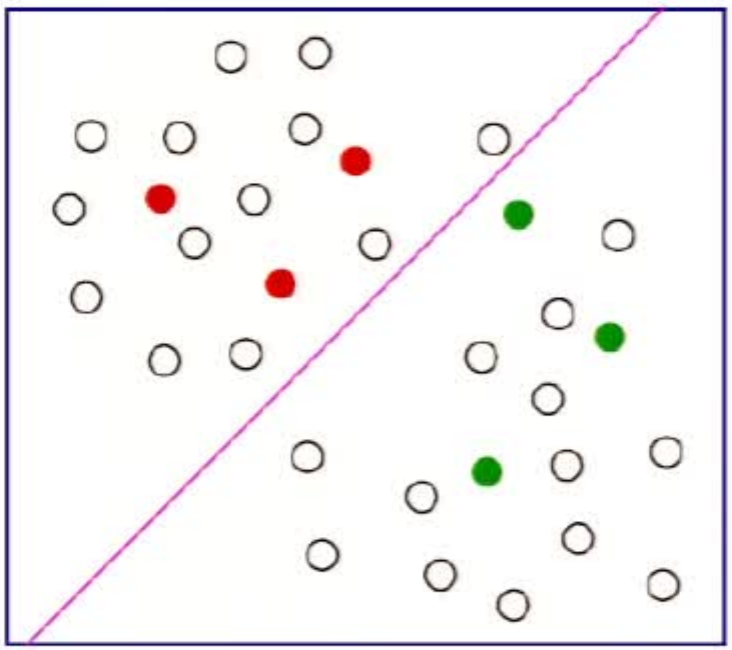
Related work

- EER: expected error reduction
- VIO: summarize uncertainties for each task

Related work

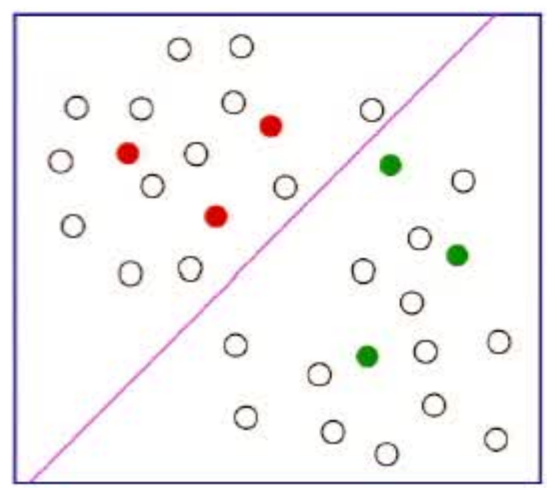
- EER: expected error reduction
- VIO: summarize uncertainties for each task
- Ours: use bandit framework

Active learning vs. Multi-armed bandit



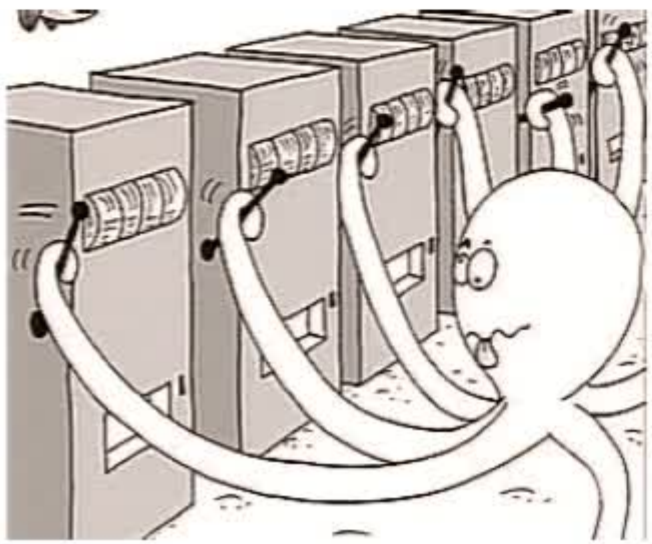
Active learning

- Select an instance from a pool
- Query the label of the selected instance
- Train a new classifier based on new labeled data
- The goal: obtain a classifier with good performance



Bandit – Multi-armed bandit

- Select an arm from a set of arms
- Get the payoff of the selected arm
- Update the historical payoff records for each arm
- The goal: obtain the arm with high payoffs



The similar things between active learning and multi-armed bandits

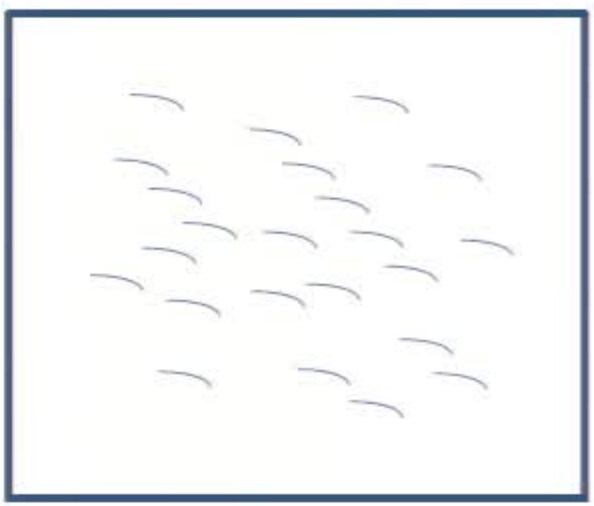
Active learning	Multi-armed bandits
Hypothesis (classifier)	Arm
Performance	Payoff
Make the query for a instance and get the performance based on the new labeled train dataset	Pull the arm and obtain the payoff

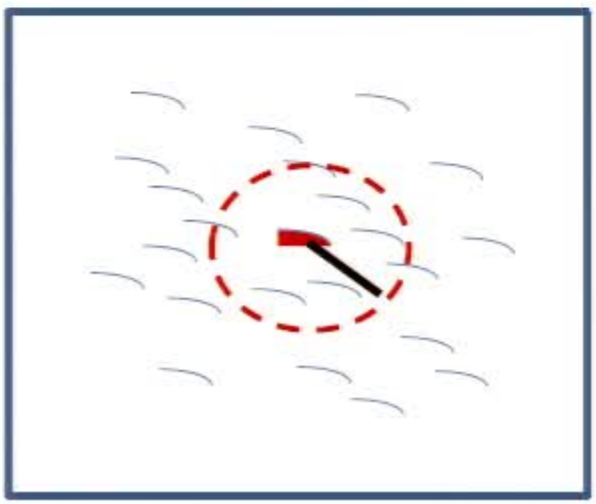
Under the bandit framework

- We formalize the active learning algorithm for multi-task learning under the bandit framework.
- Hypothesis - arm
- Risk - payoff
- Trade-off between exploration and exploitation: confidence bound of hypothesis

Trade-off: confidence of hypothesis

- Confidence: distance to the ground truth.





Algorithm

- Risk and confidence
- Trade-off between risk and confidence
- Two goals: lower risk and lower confidence
- Provide an implementation of our approach based on multi-task learning with trace-norm regularization method.

Arm - hypothesis

- In the multi-task learning, we consider the hypothesis as the arm.
- Given a dataset, we solve the optimization problem:

$$h = \arg\min_{h \in H} \hat{R}(h) + \mu \|W\|_*$$

Payoff - Risk

- The risk

$$R(h) = \frac{1}{M} \sum_{m=1}^{M} \mathbb{E}_{(x,y)\sim\mu_m}[\ell(h(x), y)]$$

- Average empirical risk.

Confidence bound

- Confidence bound

$$CB = \sqrt{\frac{ln(1/\sigma)}{2\overline{n}M}} + 2LB\left(\sqrt{\frac{\|\hat{C}\|_\infty}{\overline{n}}} + \sqrt{\frac{2(\ln(\overline{n}M)+1)}{\overline{n}M}}\right)$$

- It equates the excess risk of multi-task learning algorithm with trace-norm regularization.

Two criterions

- Consider both the risk and the corresponding confidence, we want to find a hypothesis which can be

$$h = \arg\min_{h \in H} R(h) + C(h)$$

- Then we want to minimize both the risk and the upper confidence bound.

Trade-off between risk and confidence

- For the multi-task learning problem, firstly, we must learn a large enough candidate set to contain hypothesis set with low risk.
- Then we should also learn a small enough hypothesis set that we can find such hypothesis close to true hypothesis.

Active learning algorithm

Experiments

- We evaluate our algorithm on a synthetic dataset and three real multi-task datasets: Restaurant & Consumer dataset, Dermatology dataset and School dataset.

Baselines:

- ERR: expected error reduction based method.

Baselines:

- VIO value of information algorithm, which summaries the uncertainty of each task using traditional uncertainty strategy, defined as

$$VOI(Y, x) = \sum_{y} p(Y = y|x) R(p, Y = y, x)$$

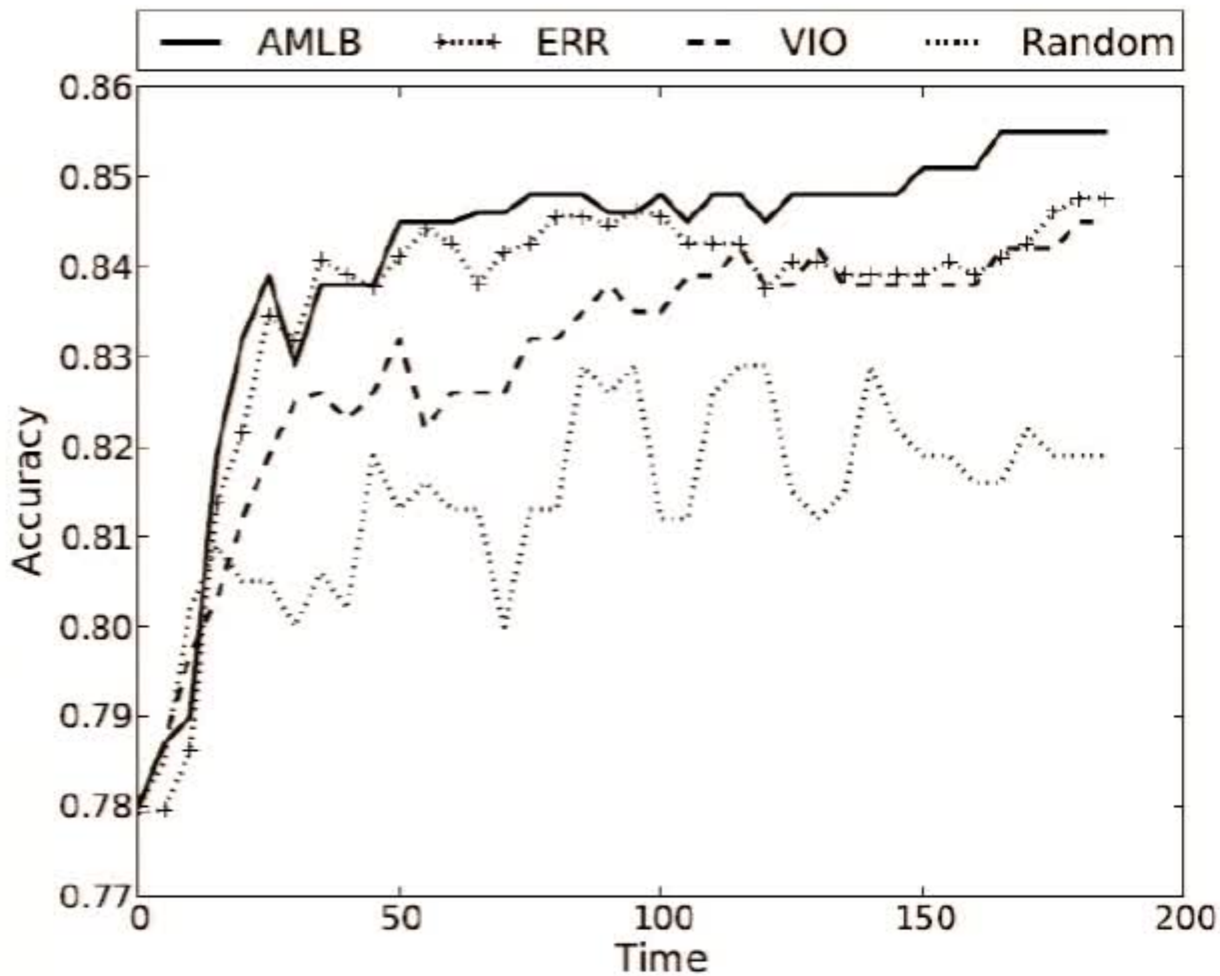
- where R is the rewards function and we use $R(p, Y = y, x) = -\log p(Y = y|x)$. This strategy is to select the instance which has the most uncertainty information over all tasks;

Baselines:

- Random: passive learning algorithm, which randomly selects instances from dataset.

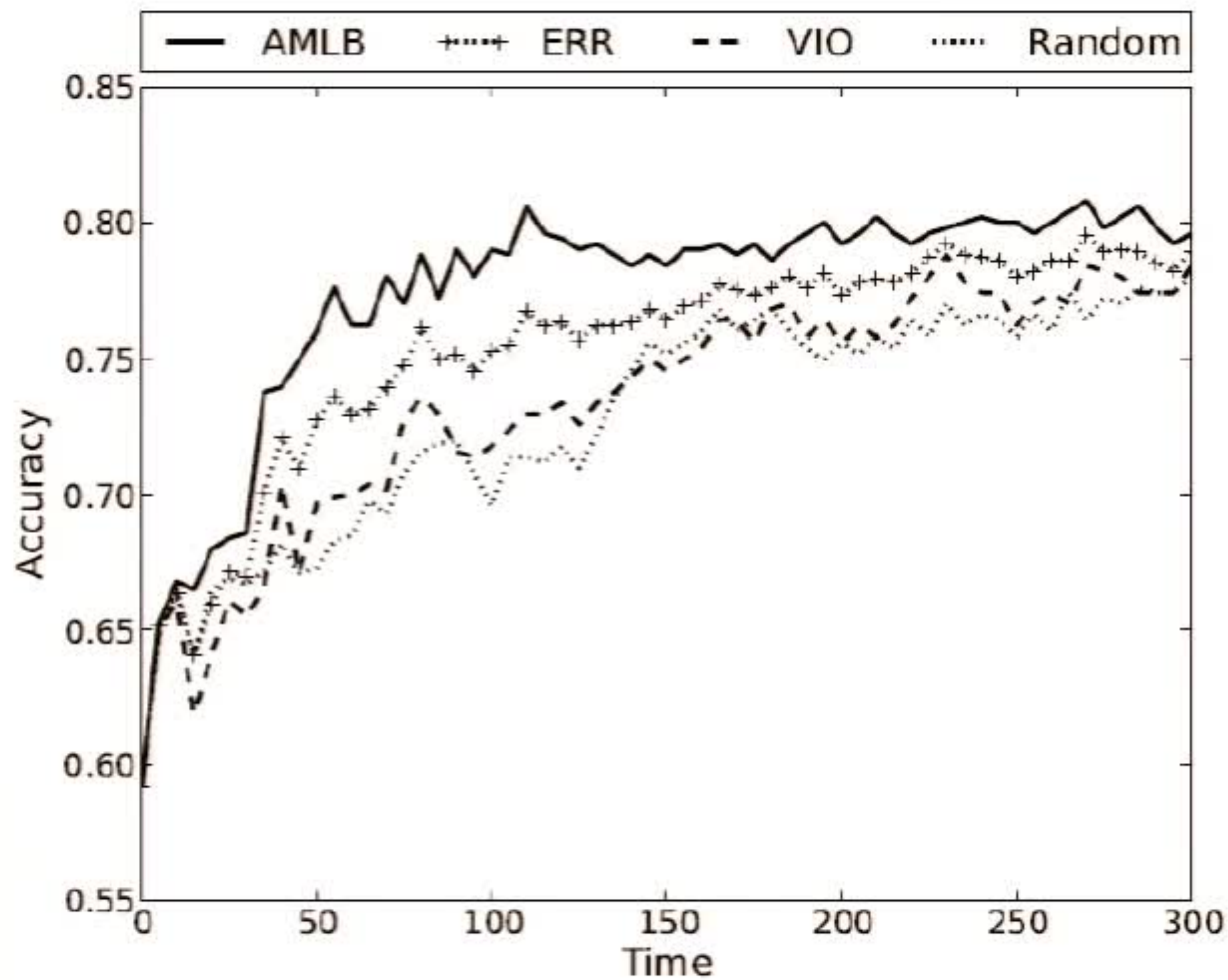
Synthetic data

- Performance comparison:



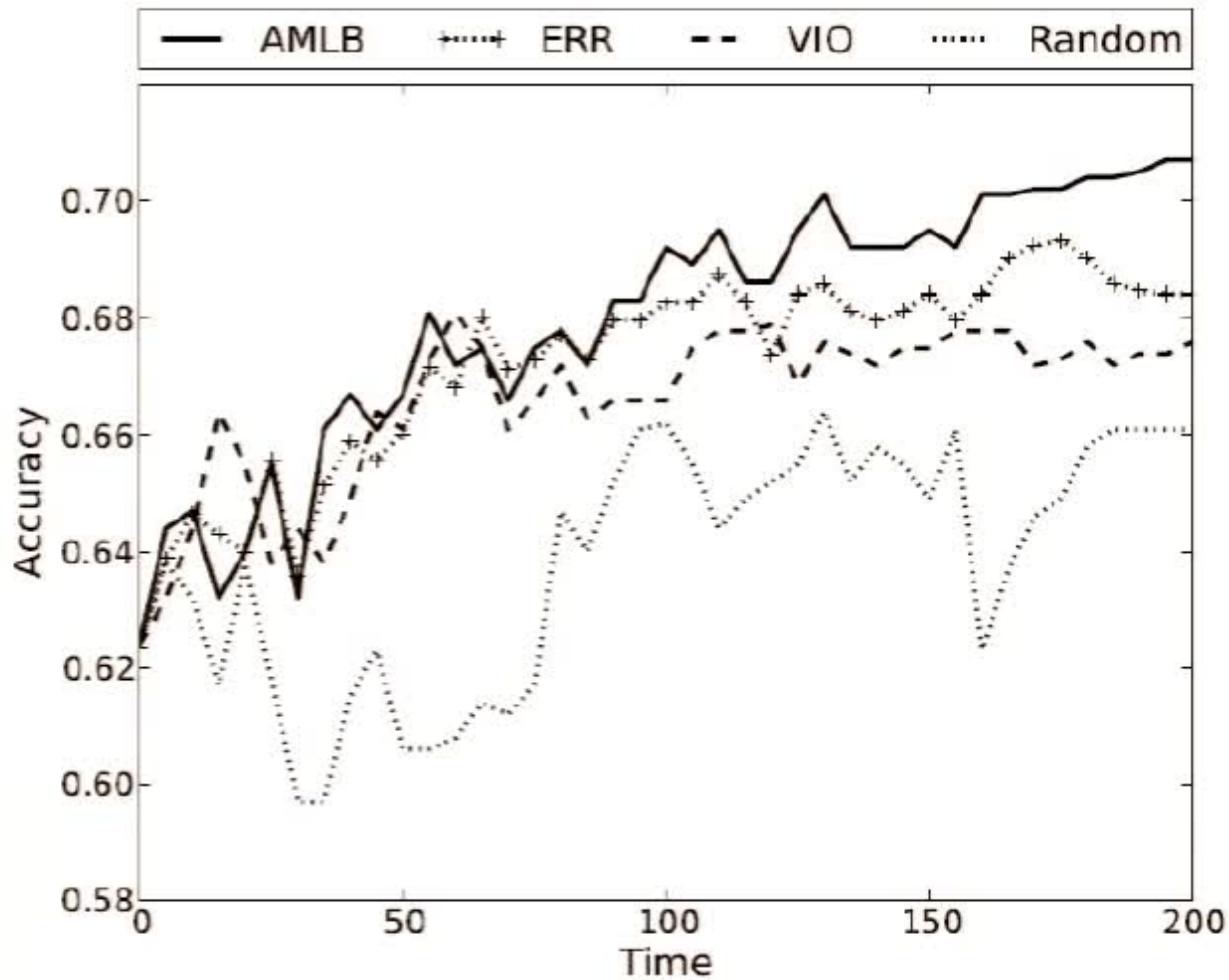
Restaurant & consumer data

- Performance comparison



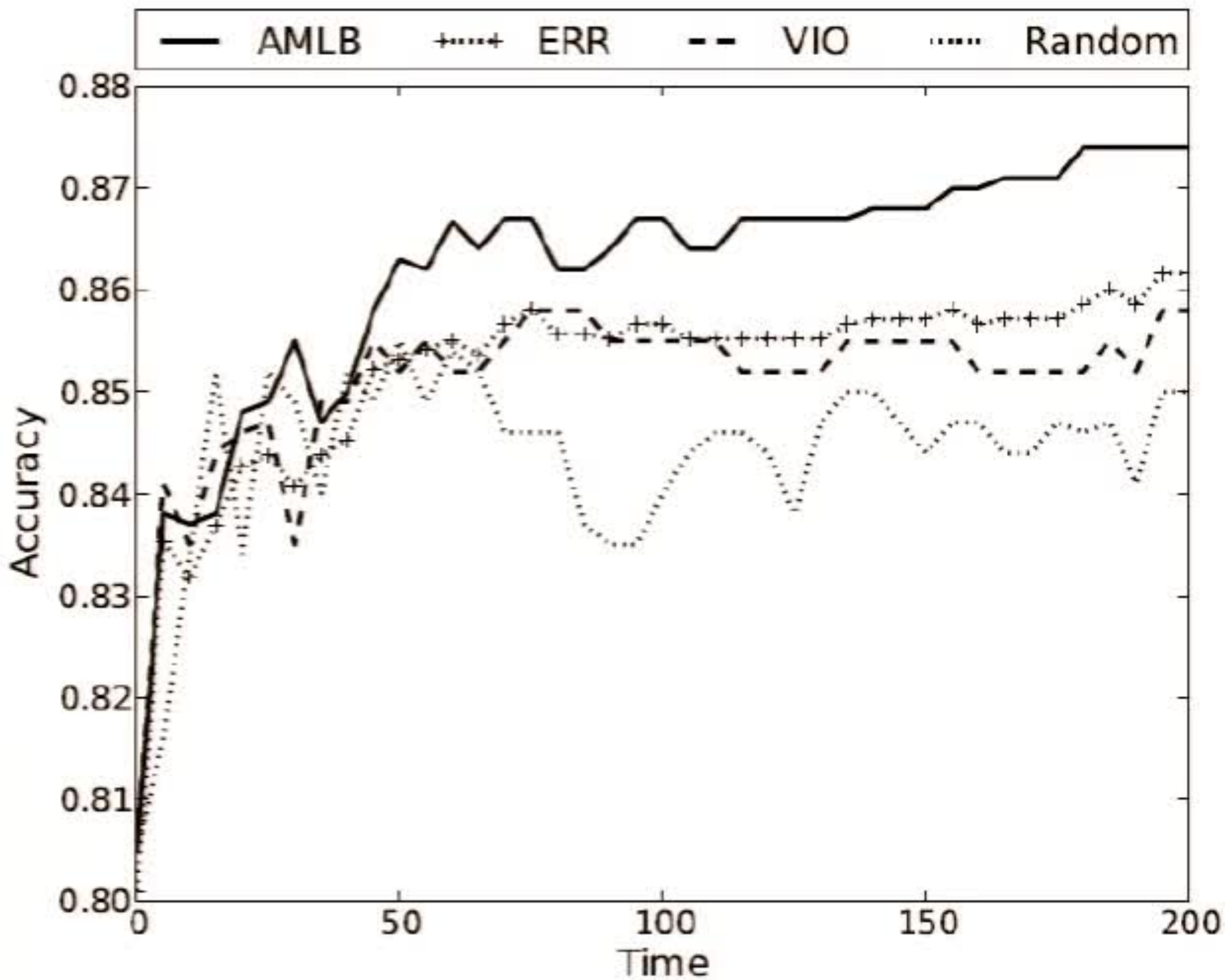
Dermatology data

- Performance comparison



School data

- Performance comparison



Conclusion

- Propose a new active learning framework for multi-task learning, named active multi-task learning via bandits.
- Consider the trade-off between minimizing the risk and improving the confidence bounds for the hypothesis.
- Provide an implementation of our approach based on multi-task learning with trace-norm regularization method.

Q & A

- Thanks.

- Finding a job in academia or industry.
- Email: Meng.Fang@student.uts.edu.au