Active Learning by Learning

Wei-Ning Hsu
Department of Electrical Engineering,
National Taiwan University
mhng1580@gmail.com

Hsuan-Tien Lin
Department of Computer Science and Information Engineering,
National Taiwan University
htlin@csie.ntu.edu.tw

AAAI-2015

Outline

- Introduction
- Methods
- Experiments
- Conclusion

Introduction

- Different AL strategy is unlikely to work on all scenario.
- Choosing strategy under different scenario is important but challenging practical task.
- let the machine adaptively “learn” from the performance of a set of given strategies on a particular data set.
- we design a learning algorithm that connects active learning with the well-known multi-armed bandit problem.
- The proposed approach, shorthanded ALBL for active learning by learning.

Outline

- Introduction
- Methods
- Experiments
- Conclusion

Multi-armed bandit

Given K bandit machines and a budget of T iterations.

The gambler is then asked to sequentially decide which machine to pull in each iteration $t = 1, \ldots, T$.

On being pulled, the bandit machine randomly provides a reward from a machine-specific distribution unknown to the gambler.

maximize the total rewards earned through the sequence of decisions.

Methods

Our key idea is to draw an analogy between our task and the multi-armed bandit problem.

The analogy faces two immediate difficulties:

- how to identify an appropriate multi-armed bandit method to solve the problem。
- how to design a reward scheme that connects the goal of active learning to the goal of the multi-armed bandit problem。

Choice of Multi-armed Bandit Method

- First, it is intuitive that the rewards are not independent random variables across the iterations, because the learning performance generally grows as Dl becomes larger.
- Second, the contributions to the learning performance can be time-varying because different algorithms may perform differently in different iterations.

One state-of-the-art method is called EXP4.P

EXP4.P

$\mathbf{w}(t)$	The weight vector in iteration t
$w_k(t)$	weight of the k-th active learning algorithm
$\mathbf{p}(t) \in [p_{\min}, 1]^K$	EXP4.P randomly chooses an expert (active learning algorithm in ALBL) based on p(t), and obtains the reward r of the choice
$\boldsymbol{\psi}^k(t) \in [0,1]^{n_u}$	The query vector for each algorithm.
$\psi_j^k(t)$	s the preference of the k-th algorithm on querying the label of xj ∈ Du in iteration t.

First, EXP4.P choose an active learning algorithm, and then, ALBL query the label of some x∗ ∈ Du

$q_j(t) = \sum_{k=1}^{K} p_k(t)\psi_j^k(t)$	the probability of querying the j-th instance in the t-th iteration

update the $w_k(t)$

Choice of Reward Function

- test accuracy- not suitable due to the costliness of label.
- training accuracy- it suffers from the inevitable training bias. it suffers from the sampling bias when using active learning to strategically query the unlabeled instances.
- First assume that the data pool D is fully labeled and each example in D is generated i.i.d. from some distribution that will also be used for testing.

$\frac{1}{n}\sum_{i=1}^{n}[\![y_i = f(x_i)]\!]$ an unbiased estimator of the test accuracy of f.

IMPORTANCE-WEIGHTED-ACCURACY

$q_i > 0$ for each example $(x_i, y_i) \in D$,

sample one (x_*, y_*)

$s_i \in \{0, 1\}$ denote the outcome of the sampling.

$c_i = [\![y_i = f(x_i)]\!]$

expected value of $s_i \frac{c_i}{q_i}$ over the sampling process is simply c_i.

$\frac{1}{n} \sum_{i=1}^{n} s_i \frac{c_i}{q_i} = \frac{1}{n} \frac{c_*}{q_*}$ is also an unbiased estimator of the test accuracy of f

D_l can be requried since $q_i > 0$.

IMPORTANCE-WEIGHTED-ACCURACY

RANDOM that randomly selects one instance from the entire data pool.
First, no modification of other active learning algorithms is needed.
Second, RANDOM is sometimes competitive to active learning algorithms.

$$\text{IW-ACC}(f,\tau) = \frac{1}{nT}\sum_{t=1}^{\tau} W_t[\![y_{i_t} = f(x_{i_t})]\!]. \quad W_t = (q_{i_t}(t))^{-1}$$

an unbiased estimator of the test accuracy of f

Theorem 1. *For any* τ, $\mathbb{E}[\text{IW-ACC}(f,\tau)] = \frac{1}{n}\sum_{i=1}^{n}[\![y_i = f(x_i)]\!]$, *where the expectation is taken over the randomness of sampling independently in iteration* $1, 2, \ldots, \tau$.

ALBL

Algorithm 1 ACTIVE LEARNING BY LEARNING

Input: $D = (D_u, D_l)$: data pool; T: query budget; $A = \{a_1, \ldots, a_K\}$: active learning algorithms, one of which is RANDOM

Initialize:

set $t = 1$, $budget\_used = 0$
while $budget\_used < T$ **do**
    run EXP4.P for one iteration and obtain the choice vector $\mathbf{p}(t)$
    for all x_i in D, calculate $q_i(t)$ using $\mathbf{p}(t)$ from EXP4.P and $\boldsymbol{\psi}^k(t)$ from all the a_k's
    sample an instance x_{i_t} based on $q_i(t)$, and record $W_t = (q_{i_t}(t))^{-1}$
    if $x_{i_t} \in D_u$ (i.e., has not been queried) **then**
        query y_{i_t}, move (x_{i_t}, y_{i_t}) from D_u to D_l, and train a new classifier f_t with the new D_l
        $budget\_used = budget\_used + 1$
    else
        $f_t = f_{t-1}$ because D_l is not changed
    end if
    calculate reward $r = \text{IW-ACC}(f_t, t)$
    feed r to a modified EXP4.P that updates the weights of all the algorithms (experts) that suggest x_{i_t}
    $t = t + 1$
end while

Outline

- Introduction
- Methods
- Experiments
- Conclusion

Experiment

Baseline:

RANDOM

UNCERTAIN (Tong and Koller 2002)

PSDS (Donmez and Carbonell 2008)

QUIRE (Huang, Jin, and Zhou 2010)

take SVM (Vapnik 1998) as the underlying classifier

take six real-world data sets, liver, sonar, vehicle, breast, diabetes, heart) from the UCI Repository

Experiment

We first compare ALBL with the four algorithms it incorporates.

Experiment

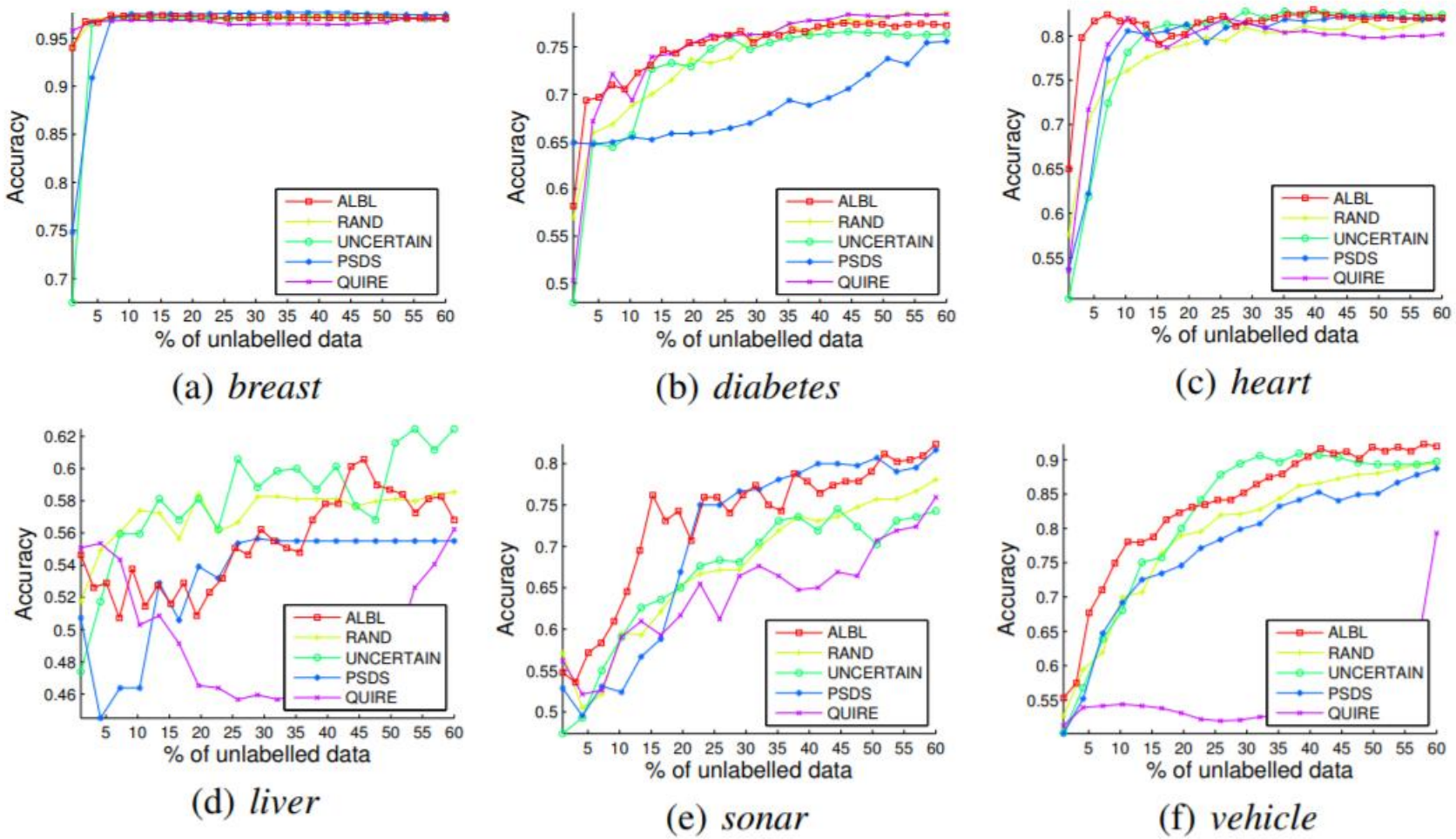

Figure 1: Test accuracy of ALBL and underlying algorithms

ALBL is usually close to the best curves of the four underlying algorithms, except in liver

Experiment

ALBL versus fixed combination

The performance of ALBL was compared to that of FIXEDCOMB when incorporating two active learning algorithms, one of which reaches the best performance and the other reaches the worst performance on each data set.

Further, we consider sampling weight ratios: 10:0, 8:2, 6:4,5:5, 4:6, 2:8, 0:10 in FIXEDCOMB.

Experiment

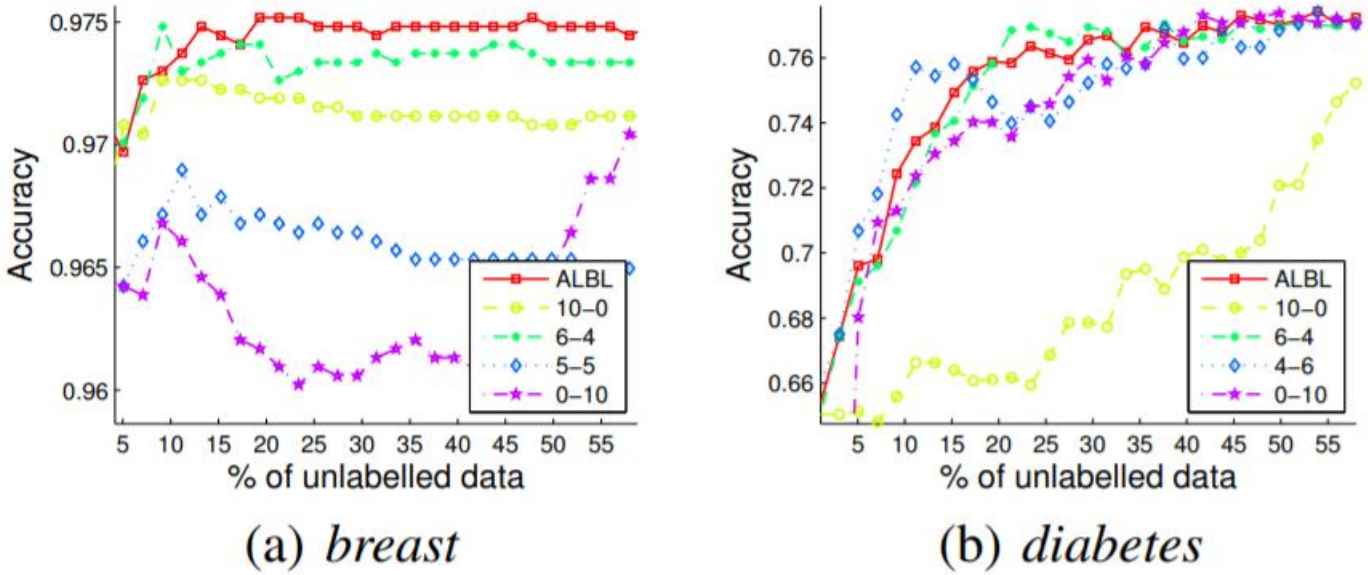

Figure 2: Test accuracy of ALBL versus FIXEDCOMB

Two drawback:
First, deciding the best weight ratio beforehand is a very challenging endeavor.
The second drawback is that FIXEDCOMB cannot capture the time varying behavior of the underlying algorithms.

Experiment

Finally, we demonstrate the benefits of using the unbiased estimator in ALBL by comparing it with two related approaches:

COMB- the unlabeled examples as the bandit machines instead. takes a human-designed criterion called CLASSIFICATION ENTROPY MAXIMIZATION(CEM) as the reward. defined as the entropy of f_t-predicted labels in D_u

ALBL-TRAIN that takes the training accuracy as the reward.

Experiment

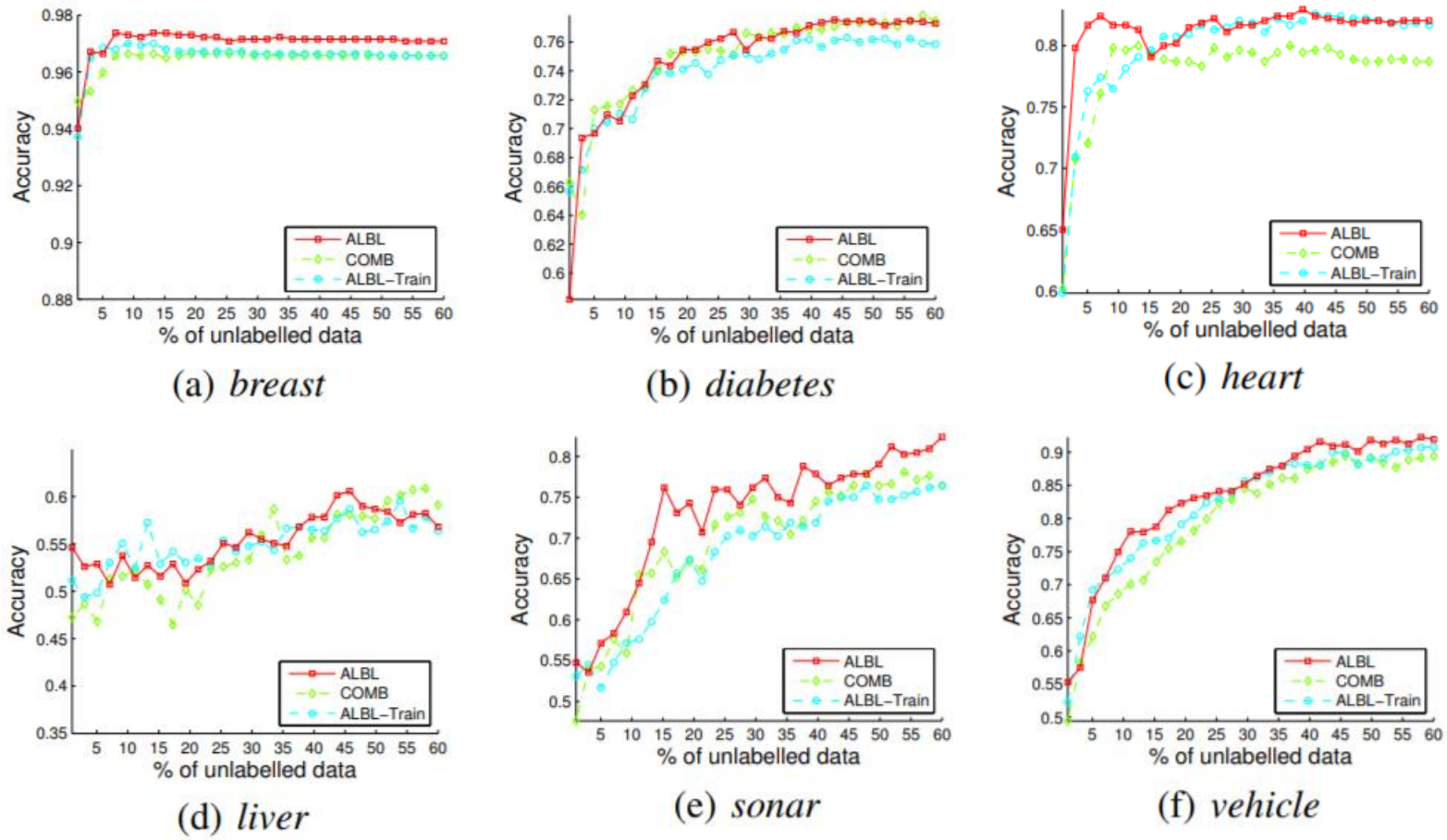

Figure 3: Test accuracy of ALBL, COMB, and ALBL-TRAIN

On most of the other data sets, ALBL achieves superior performance to those of the COMB and ALBL-TRAIN

Experiment

We further analyze the superior performance by evaluating IW-ACC, CEM, the training accuracy, and the true test accuracy at each iteration of ALBL, and depict two representative results in Figure 4.

Experiment

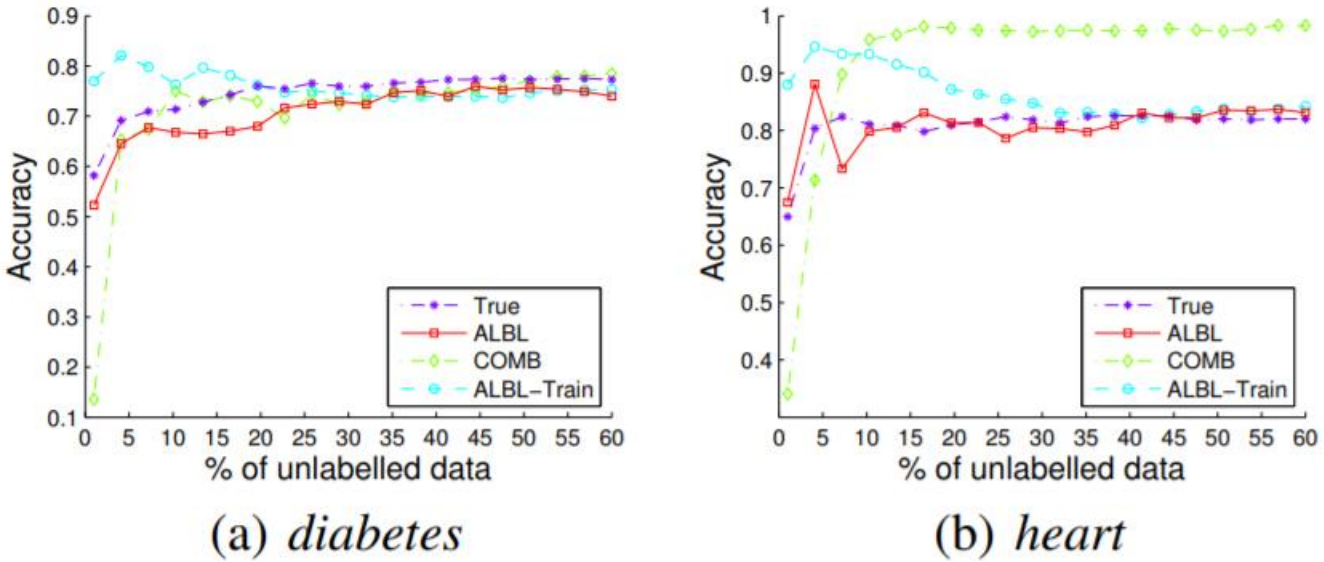

Figure 4: Different estimations of true test accuracy

Outline

- Introduction
- Methods
- Experiments
- Conclusion

Conclusion

- We propose a pool-based active learning approach ALBL. ALBL adaptively and intelligently chooses among various existing active learning strategies by using their learning performance as feedback
- We utilize the famous EXP4.P algorithm from the multi-armed bandit problem for the adaptive choice, and estimate the learning performances with IMPORTANCE-WEIGHTED-ACCURACY
- First, ALBL is effective in making intelligent choices, and is often comparable to or even superior to the best of the existing strategies.
- Second, ALBL is effective in making adaptive choices, and is often superior to naive blending approaches that randomly choose the strategies based on a fixed ratio.
- Third, ALBL is effective in utilizing the learning performance, and is often superior to the human-criterion-based blending approach COMB.