Multi-Domain Active Learning for Semi-Supervised Anomaly Detection

Vincent Vercruyssen, Lorenzo Perini, Wannes Meert, Jesse Davis

DTAI Research Group @ KU Leuven, Belgium

KU LEUVEN Spottr LEUVEN.AI DTAI

Multi-domain active learning as a challenging problem

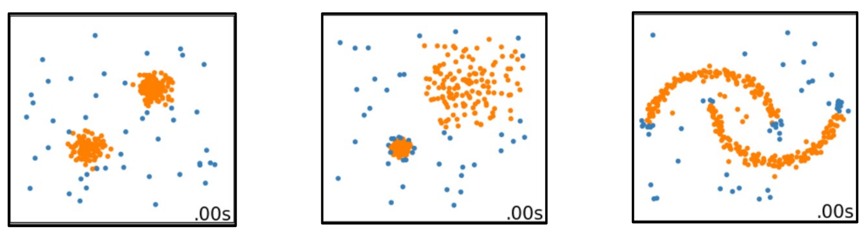
Scenario: "*You own a retail company and manage several retail stores and each store is quite different. You measure the water consumption in each store, and you want to use machine learning to detect periods of abnormal water consumption in each store.*"

GIVEN: a multi-domain dataset M consisting of K datasets, a fixed query budget T, and an expert oracle O

DO: learn a classifier for each of the K datasets with active learning

Challenges

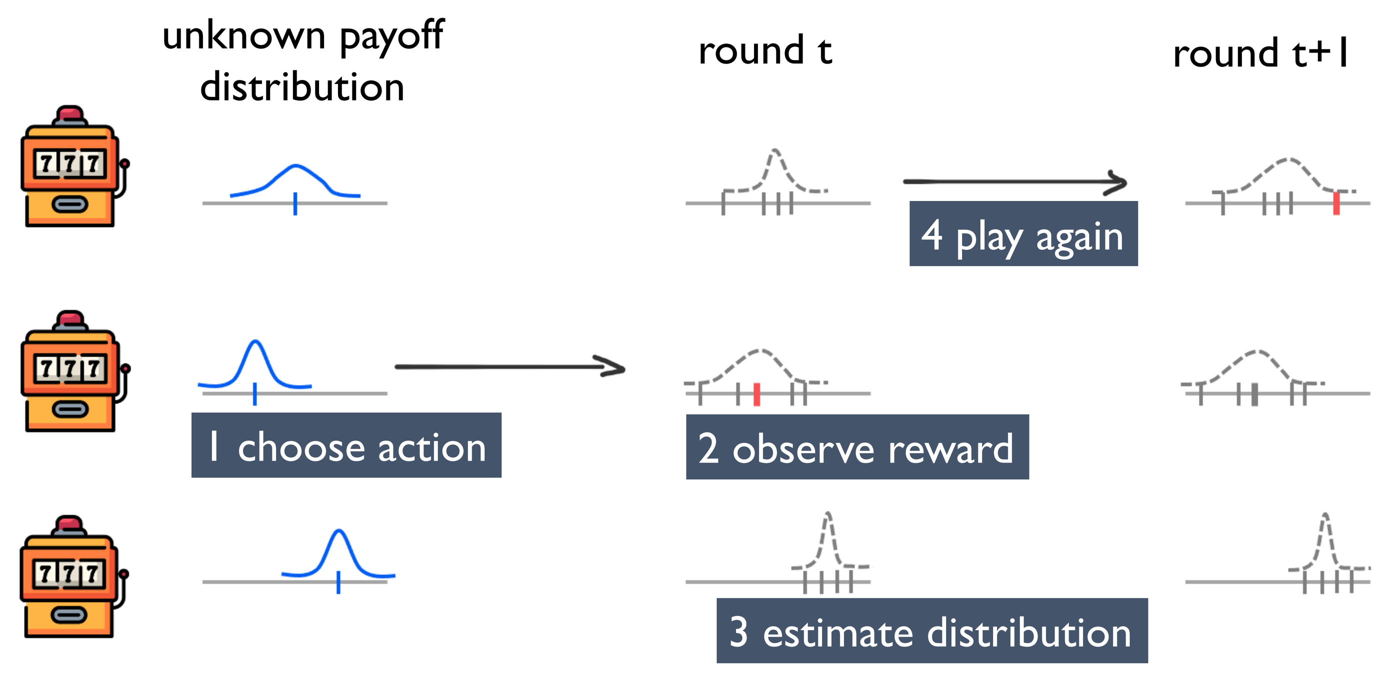
1. Marginal gain of acquiring a new label for a dataset = labeling payoff
2. Diminishing labeling payoff upon acquiring new labels from the oracle
3. Some datasets are more interesting to label than others



Exploration-exploitation with multi-armed bandits

Trade-off

1. Exploration = finding out *which* datasets have a high labeling payoff
2. Exploitation = focusing on labeling *those* datasets with high payoff



Intuition:

Choosing a group of examples to label = choosing a machine to play

An *Active Learning Bandits* (ALBA) algorithm for multi-domain active learning

1. ALBA chooses between a fixed set of *actions* each query round

action = choose which of the K datasets to query an example from

$$|actions| = K$$

OR action = choose which cluster in a dataset to query from

$$|actions| = K \times C$$

2. ALBA selects *one* final query from a group to ask to the oracle

select an example from the group randomly

✗ not necessarily query most informative examples
✓ unbiased estimate of the reward distribution

OR select an example from the group heuristically

✓ select examples with highest estimated labeling payoff
✗ distorts the reward distribution estimate

3. Each *action* (asked query) leads to a reward

reward = entropy reduction in the predictions of the classifier

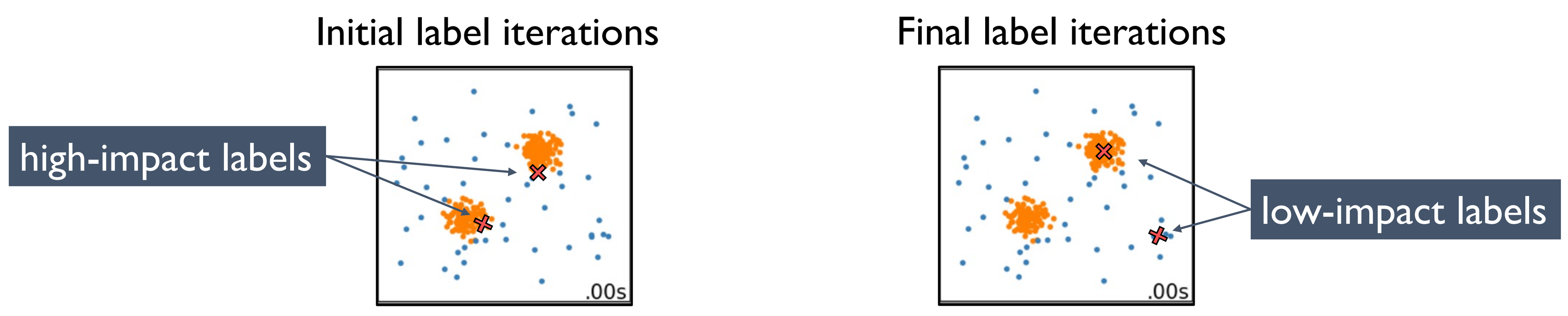
$$r = \sum_{x \in D^k} \left[H_{f_+^k}(x) - H_{f^k}(x) \right]$$

OR reward = number of examples for which the predicted label changes

$$r = 1 - \frac{Y_{f_+^k} \cdot Y_{f^k}}{\left\| Y_{f_+^k} \right\| \, \left\| Y_{f^k} \right\|}$$

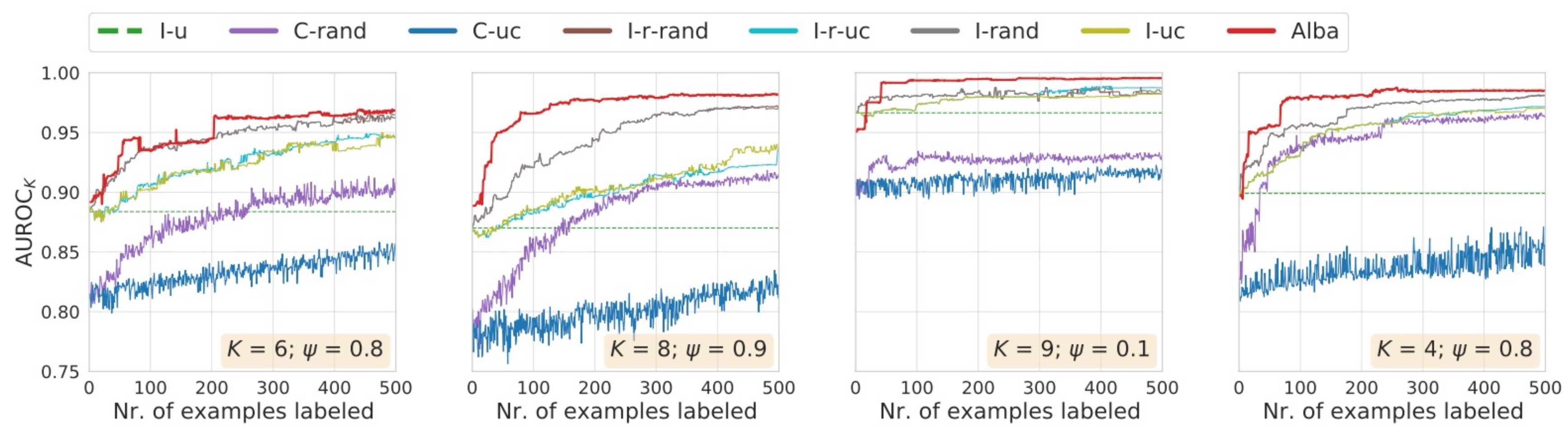
4. ALBA uses *rotting bandits* to handle diminishing labeling payoff

use the *Sliding Window Average* (SWA) algorithm



The *Active Learning Bandits* (ALBA) algorithm outperforms active learning baselines on several real-world datasets

ALBA is significantly better than all baselines @ 100 query rounds



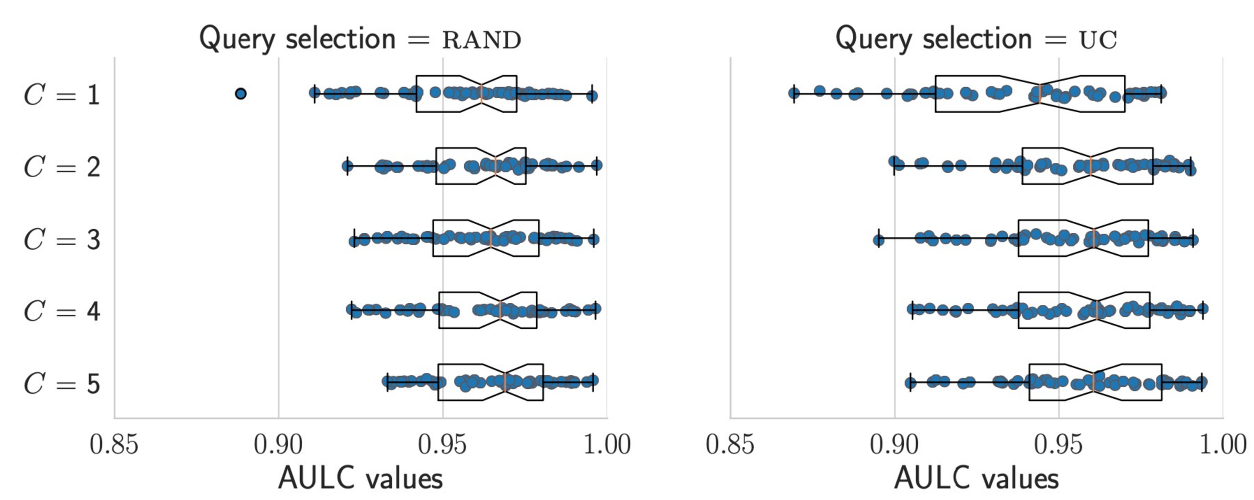
Cosine reward > entropy reward
Random query selection > heuristic query selection

Reward function	Query sel. strategy	Ranking Avg. ± SD
cosine	rand	**1.806 ± 0.813**
cosine	uc	2.944 ± 0.926
entropy	rand	2.241 ± 0.843
entropy	uc	3.009 ± 0.825

$$\text{reward} + \text{query selection} \sim \text{AULC}$$

	Coeff.	Std. error	t	$P > \|t\|$
Intercept	0.9568	0.003	371.522	0.000
reward	0.0007	0.003	0.220	0.826
query sel.	0.0080	0.003	2.681	0.002

Increasing the number of clusters increases ALBA's performance



Conclusions

1. Multi-armed bandit algorithms can be an effective tool in active learning problems
2. Our contribution, the ALBA algorithm, outperforms existing active learning algorithms for multi-domain active learning
3. Constructing a large benchmark helped gain insight in the algo's

Code & data: https://github.com/Vincent-Vercruyssen