

# Online Transfer Learning for Differential Diagnosis Determination

Jie Xu<sup>\*†</sup>, Daby Sow<sup>†</sup>, Deepak Turaga<sup>†</sup>, Mihaela van der Schaar<sup>\*</sup>

<sup>\*</sup>University of California at Los Angeles, Los Angeles, California 90095

<sup>†</sup>IBM T. J. Watson Research Center, Yorktown Heights, New York 10598

## Abstract

In this paper we present a novel online transfer learning approach to determine the set of tests to perform, and the sequence in which they need to be performed, in order to develop an accurate diagnosis while minimizing the cost of performing the tests. Our learning approach can be incorporated as part of a clinical decision support system (CDSS) with which clinicians can interact. The approach builds on a contextual bandit framework and uses online transfer learning to overcome limitations with the availability of rich training data sets that capture different conditions, context, test results as well as outcomes. We provide confidence bounds for our recommended policies, which is essential in order to build the trust of clinicians. We evaluate the algorithm against different transfer learning approaches on real-world patient alarm datasets collected from Neurological Intensive Care Units (with reduced costs by 20%).

## Introduction

Recent advances in sensing and measurement technologies are enabling us to monitor complex human, engineered, physical, biological and chemical systems and processes in many sophisticated ways. This enables improved ability to understand the state of health of these systems, diagnose problems, and use this to design interventions to maximize health at varying timescales. However, while several such measurements can be made (e.g. by performing different tests on a patient), the decision on which test to perform and when to perform it remains a very challenging problem.

Challenges stem from multiple factors: i) There are complex relationships between different attributes that are being measured. ii) Tests have varying degrees of costs associated with them (e.g. some tests are very expensive). iii) Tests are significantly impacted by *context*, i.e. the *best* set of tests, measurements and interventions may be different depending on the context in which it takes place. iv) The determination of tests is often challenged by the limited access to relevant data. For instance, existing patients datasets often have distributions that do not necessarily capture the information needed for the accurate diagnosis in a

novel problem domain. Thus, the resulting diagnosis policies may perform poorly. This prompts the need for a system that can effectively perform context-specific diagnosis that maximizes diagnosis accuracy and minimizes test costs even when highly relevant data pertaining to the diagnosis decision is missing.

In this paper, we present a novel decision support system that addresses these challenges by transferring knowledge from multiple related problem domains and incrementally learning the best policies (i.e. sequences of test) to adopt depending on the context of the diagnosis problem. These contexts can be exogenous facts or meta-data about the problem. In the medical setting, they can be patient's age, gender and weight. Note that the contexts are different than the endogenous testing results. The use of multiple related problem domains enables transferring knowledge from the most relevant domains for different diagnosis contexts; it also creates a way to measure the semantic similarity between contexts: contexts are similar if their most related existing domains are the same. The learned semantic similarity is then used to develop context-specific solutions in the novel problem domain. The proposed approach is able to provide diagnosis confidence bounds which are important to ensure the trust of domain professionals.

## Related Work

Support systems for decision making have been extensively studied. A first strand of related research focuses on cost-sensitive learning (Turney 2000)(Greiner, Grove, and Roth 2002)(Zubek, Dietterich, and others 2004). A disadvantage of these approaches is that they rely on training datasets to learn the appropriate model. Our focus in this paper is on how to overcome the lack of initial training data by using transfer of knowledge from relevant datasets. The majority of the transfer learning literature assumes a single source from which knowledge can be transferred (Marx et al. 2005). Transfer learning from multiple sources is much more challenging; most works aiming to address this problem focus on classification problems (Duan et al. 2009)(Yao and Doretto 2010). In our considered setting, the target data arrives sequentially and the features are not given but need to be discovered (by performing various tests). Imitation-type transfer learning techniques are often adopted where source policies are applied to the target task initially, while the target task solution is learned gradually (Fernández and



are unknown a priori. Our algorithm begins by exploring the source policies for entities in the target domain. After accumulating sufficient data on entities for the target problem, it builds the target policy using the information extracted from applying the source policies. The algorithm is provided next in Algorithm 1. The parameter  $\rho^k \in [0, 1]$  is used to control when to adopt source policies and when to use the newly built target policies; it is decreasing in  $k$  and  $\lim_{k \rightarrow \infty} \rho^k = 0$ .

---

**Algorithm 1** Transfer Learning with Multiple Sources
 

---

- 1: **for** each entity  $k$  **do**
  - 2:     With probability  $\rho^k$ , select a source policy to apply
  - 3:     With probability  $1 - \rho^k$ , apply the target policy
  - 4:     After the current case is closed
  - 5:     Build the target policy using received data
  - 6: **end for**
- 

In Algorithm 1, there are two major questions that remain to be addressed: which source policy to apply (line 3) and how to build the target policy (line 7). We discuss them next. Let  $(k_1, k_2, \dots, k_t, \dots)$  be the subsequence of received entity cases where a source policy is adopted according to Algorithm 1. Without loss of generality, we normalize the entity context space to be  $\Theta \in [0, 1]^W$  where  $W$  is the context space dimension. We introduce some concepts of the algorithm as follows: 1) **Entity cluster**. An entity cluster is represented by the range of context information that is associated with entities in the cluster. In this paper, we will consider clusters with the form  $[i_w 2^{-(l-1)w}, (i_w + 1) 2^{-(l-1)w}]$  where  $i_w \in \{0, 1, \dots, 2^{(l-1)w} - 1\}$  for each context dimension  $w = 1, \dots, W$  for some positive integer  $l$ . Such a cluster is called a level- $l$  cluster. At each time  $k_t$  when source policies are applied, the algorithm keeps a set of mutually exclusive clusters that cover the entire context space. We call these clusters the *active* clusters, and denote this set by  $\mathcal{H}^t$ . Clearly, we have  $\cup_{C \in \mathcal{H}^t} = \Theta, \forall t$ . 2) **Counters**. For each active cluster  $C$ , the algorithm maintains several counters: for each source policy  $\pi \in \Pi$ ,  $M_C(\pi)$  records the number of entity cases so far in which  $\pi$  is applied. 3) **Diagnosis cost estimates**. For each active cluster  $C$ , the algorithm also maintains the sample mean diagnosis cost estimate  $\bar{r}_C(\pi)$  for each source policy  $\pi \in \Pi$ , using the observed diagnosis costs of cases that belong to  $C$  so far.

The algorithm is described in Algorithm 2. When an entity case  $k_t$  is received, the algorithm first checks which active cluster  $C \in \mathcal{H}^t$  it belongs to. Then it investigates counter  $M_C(\pi)$  for all  $\pi \in \Pi$  to see if there exists any under-explored source policy  $\pi$  such that  $M_C(\pi) \leq \gamma(t, l)$  where  $\gamma(t, l)$  is a time- and level-dependent control function. If there exists such an under-explored policy  $\pi$ , then the algorithm selects this policy for the current entity case. This is called an exploration step. If there does not exist any under-explored policy, then the algorithm selects the policy with the lowest cost estimate  $\arg \min_{\pi \in \Pi} \bar{r}_C(\pi)$ . This is called an exploitation step. After the diagnosis cost of the current entity case is observed, the cost estimate of the selected policy is updated. Moreover, if  $\sum_{\pi \in \Pi} M_C(\pi) \geq \zeta(l)$ , where  $\zeta(l)$

---

**Algorithm 2** Policy Selection and Adaptive Clustering
 

---

- 1: Initialize  $\mathcal{H} = \Theta, \bar{r}_\Theta(\pi) = 0, M_\Theta(\pi) = 0, \forall \pi \in \Pi$ .
  - 2: **for** each entity  $k_t$  **do**
  - 3:     Determine active cluster  $C \in \mathcal{H}^t$  such that  $\theta^t \in C$
  - 4:     **Case 1:**  $\exists \pi \in \Pi$  such that  $M_C(\pi) < \gamma(t)$
  - 5:     Randomly select among such policies  $\sigma^t = \pi$
  - 6:     **Case 2:**  $\forall \pi \in \Pi, M_C(\pi) \geq \gamma(t)$
  - 7:     Select  $\sigma^t = \arg \min_{\pi \in \Pi} \bar{r}_C(\pi)$ .
  - 8:     Set  $M_C(\sigma^t) \leftarrow M_C(\sigma^t) + 1$
  - 9:     (The diagnosis reward  $r^t$  is observed.)
  - 10:     Update  $\bar{r}_C(\sigma^t)$
  - 11:     Update  $\sum_{\pi \in \Pi} M_C(\pi)$  using all past cases.
  - 12:     **if**  $\sum_{\pi \in \Pi} M_C(\pi) \geq \zeta(l)$  **then**
  - 13:         Uniformly partition  $C$  into  $2^W$  level- $(l+1)$  clusters.
  - 14:         Update the set of active clusters  $\mathcal{H}^t$ .
  - 15:         Update the counters and cost estimates for all new clusters using the entity cases received so far.
  - 16:     **end if**
  - 17:     Update the set of active clusters  $\mathcal{H}^t$ .
  - 18:     Update the counters and cost estimates for all new clusters using the entity cases received so far.
  - 19:     **end if**
  - 20: **end for**
- 

is a level-dependent control function, the current cluster  $C$  is partitioned in to  $2^W$  level- $(l+1)$  clusters. From the next entity case on,  $C$  is deactivated and the new level- $(l+1)$  clusters are activated. We will show how to select the control functions  $\gamma(t, l)$  and  $\zeta(l)$  in the next section.

Entity clusters for which the estimated best source policies are the same are considered to be similar and hence, they are grouped together to form a dataset from which the problem parameters can be estimated. Using these  $K$  set of parameters, we can produce  $K$  context-specific target policies.

**Confidence Bound**

We make the following widely adopted technical assumption below; however, this is not needed for running the algorithm.

**Assumption.** (Lipschitz) For each  $\pi \in \Pi$ , there exists  $L > 0, \alpha > 0$  such that for all  $\theta, \theta' \in \Theta$ , we have  $|V_\theta(\pi) - V_{\theta'}(\pi)| \leq L \|\theta, \theta'\|^\alpha$ .

The above assumption states that if the entity context information is similar, then the expected diagnosis cost by selecting the same diagnostic policy is also similar. We can derive a confidence level of the learned effectiveness of diagnosis policies as follows:

**Proposition.** For any active level- $l$  context cluster  $C$ , at any time when policy  $\pi$  has been adopted for  $M_C$  times on entities which belong to  $C$ , then for any individual context  $\theta \in C$ , the following confidence relation between the estimated diagnosis reward and the true diagnosis reward holds  $P(|\bar{r}_C - V_\theta| > L(\sqrt{W}/2^{l-1})^\alpha + \epsilon) < e^{-2\epsilon^2 M_C}$ .

**Experiments**
**Real-world patient dataset**

We test our proposed algorithm using an alarm data set obtained from the Columbia Medical Center neurological

Intensive Care Unit (ICU). This dataset contains over a million alarm events produced by patient monitoring systems for 581 patients.

## Methodology

We artificially treat the patient alarm dataset on each day as a separate dataset.  $K$  such datasets are picked as the source datasets and another one is picked as the target dataset. Moreover, the target data is made available to the system in sequence. We treat each alarm as a medical test. Hence, only when the test is performed, the corresponding alarm status is revealed. In the experiments, we focus on predicting whether the patient will have at least one of the two secondary complications: Pneumonia and Respiratory failure. The set of tests (alarms) that we consider includes Bradycardia, Tachycardia, High Blood Pressure, Low Blood Pressure, High Respiratory Rate. We assign different costs to different types of prediction errors. Specifically, we normalized the cost of a miss detection to be 1 and the cost of a false alarm to be  $c_1$ . A uniform cost  $c_2$  is assigned for the execution of any test. In the experiments, we use a single patient context: the APACHE II (“Acute Physiology and Chronic Health Evaluation II”) score that evaluates the severity of illness of our patients upon admission in the ICU.

## Baseline approaches

1) **Empirical Diagnosis (EM)**: All medical tests are executed and all alarms are revealed. It predicts the complication if any of the alarms are positive. 2) **Average Transfer (AT)**: In this approach, we combine all source datasets to estimate the average problem parameters and construct a new diagnostic policy. This average source policy is applied to the target cases while the target policy is gradually learnt. 3) **MultiSourceTriAdaBoost (MSTAB)**: This is a modified version of the state-of-the-art MultiSourceTriAdaBoost algorithm in (Yao and Doretto 2010) for transfer learning with multiple sources. We modified the weight update structure to incorporate the diagnosis cost instead of a plain diagnosis accuracy. Since the original algorithm is an offline algorithm that assumes a training set for the target problem, we also extended it to produce an online version using batch updates as more target cases are received.

## Results

We report the diagnosis cost results for three sets of experiments in Table 1 for various parameters. The diagnosis cost is computed by averaging the diagnosis cost of patient cases from the target patient case set. In all experiments, the proposed transfer learning algorithm significantly outperforms the baseline approaches by reducing the diagnosis cost up to 20% against the best baseline approach. We also investigate the impact of different choices of contexts on the diagnosis performance. Table 2 shows the achieved diagnosis cost by using different contexts for  $K = 2$ ,  $c_1 = 0.3$  and  $c_2 = 0.005$ . The experiment results indicate that the ApacheII score is the best context in our problem. Nevertheless, the diagnosis cost by using any context is lower than those achieved by the baseline approaches.

	$c_2$	EM	AT	MSTAB	Proposed
$K = 2,$ $c_1 = 0.3$	0.01	0.299	0.259	0.246	<b>0.212</b>
	0.005	0.268	0.241	0.231	<b>0.193</b>
	0.001	0.245	0.233	0.224	<b>0.187</b>
$K = 2,$ $c_1 = 0.5$	0.01	0.357	0.309	0.312	<b>0.292</b>
	0.005	0.332	0.290	0.282	<b>0.267</b>
	0.001	0.302	0.277	0.262	<b>0.260</b>
$K = 3,$ $c_1 = 0.3$	0.01	0.274	0.260	0.265	<b>0.218</b>
	0.005	0.258	0.239	0.232	<b>0.191</b>
	0.001	0.250	0.226	0.224	<b>0.192</b>

Table 1: Diagnosis cost comparison

Context	Age	GCS	ApacheII	ApachePhys
Diag. cost	0.222	0.225	0.193	0.205

Table 2: Impact of contexts

## Conclusion

In this paper, we proposed an online transfer learning approach for differential diagnosis determination and showed how it can be incorporated as part of a clinical decision support system to improve the diagnosis performance. We envision that the proposed methodology can also be used in other complex diagnosis systems besides clinical diagnosis, such as cyber-security, biological and mechanical diagnosis.

## References

- Duan, L.; Tsang, I. W.; Xu, D.; and Chua, T.-S. 2009. Domain adaptation from multiple sources via auxiliary classifiers. In *Proceedings of the 26th Annual International Conference on Machine Learning*, 289–296. ACM.
- Fernández, F., and Veloso, M. 2006. Probabilistic policy reuse in a reinforcement learning agent. In *Proceedings of the fifth international joint conference on Autonomous agents and multiagent systems*, 720–727. ACM.
- Greiner, R.; Grove, A. J.; and Roth, D. 2002. Learning cost-sensitive active classifiers. *Artificial Intelligence* 139(2):137–174.
- Marx, Z.; Rosenstein, M. T.; Kaelbling, L. P.; and Dietterich, T. G. 2005. Transfer learning with an ensemble of background tasks. *Inductive Transfer* 10.
- Slivkins, A. 2009. Contextual bandits with similarity information. *arXiv preprint arXiv:0907.3986*.
- Turney, P. 2000. Types of cost in inductive concept learning. In *Workshop on Cost-Sensitive Learning at 17th International Conference on Machine learning*.
- Yao, Y., and Doretto, G. 2010. Boosting for transfer learning with multiple sources. In *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*, 1855–1862. IEEE.
- Zubek, V. B.; Dietterich, T. G.; et al. 2004. Pruning improves heuristic search for cost-sensitive learning. Technical report, Corvallis, OR: Oregon State University, Dept. of Computer Science.