Proceedings of the Twenty-Sixth AAAI Conference on Artificial Intelligence

Active Learning from Oracle with Knowledge Blind Spot

Meng Fang, Xingquan Zhu and **Chengqi Zhang**
Centre for Quantum Computation & Intelligent Systems,
FEIT, University of Technology Sydney, NSW 2007, Australia
Meng.Fang@student.uts.edu.au, xqzhu@it.uts.edu.au, chengqi@it.uts.edu.au

Abstract

Active learning traditionally assumes that an oracle is capable of providing labeling information for each query instance. This paper formulates a new research problem which allows an oracle admit that he/she is incapable of labeling some query instances or simply answer "I don't know the label". We define a unified objective function to ensure that each query instance submitted to the oracle is the one mostly needed for labeling and the oracle should also has the knowledge to label. Experiments based on different types of knowledge blind spot (KBS) models demonstrate the effectiveness of the proposed design.

Introduction

Recently, several works argue that it is too strong to assume that oracles may always behave perfectly. Some studies focus on multiple annotators scenario where multiple oracles provide labels with varying expertise (Donmez and Carbonell 2008; Yan et al. 2011). While works in this category assume oracles are subject to different levels of expertise, they inherently disregard whether an oracle can label an instance or not, and will still require all oracles to label instances which may be out of their domain knowledge.

In this paper, we formulate a new active learning setting, which allows an oracle to admit that he/she is incapable of labeling some query instances and will not provide labels for those instances. Meanwhile, if we assume that each query process is subject to a certain amount of cost, answering "I don't know the label" may also involve necessary cost because the oracle has spent effort and time to investigate the instances.

Motivated by the above observations, we propose, in this paper, a mutual information theory based framework to query the instance which has the maximum mutual information according the knowledge blind spot of the oracle. To characterize the oracle's knowledge blind spot, we use diverse density to transform instances into a new feature space, through which we can accurately assess the likelihood of each unlabeled instance belonging to the oracle's KBS.[1]

[1]For more information about our approach and experiments, see http://mengfang.wordpress.com/2012/02/03/alokbs/

Information-theoretic Model

Consider a data set with n instances $\{x_1, x_2, \ldots, x_n\}$, where the label for the i^{th} instance is denoted by y_i. Our objective is to query the instance that has the most information, given the labeled data, as defined in Eq.(1), where $\mathcal{U}$ denotes the set of unlabeled instances and H represents the entropy/uncertainty of instance x_i with respect to the class labels predicted from a classifier $\hbar(.)$ trained from labeled set $\mathcal{L}$.

$$\underset{x_i \in \mathcal{U}}{argmax}\, H(y_i; \hbar(\mathcal{L})) \tag{1}$$

Definition A Knowledge Blind Spot (KBS) is defined as a collect of instances to which a labeler (or an oracle) does not have knowledge assigning class labels.

Denote $\mathcal{B}^+$ the set of KBS of an oracle, and $\mathcal{B}^-$ is the set of knowledge that the oracle has already acquired. Then the expected entropy of an unlabeled instance x_i with respect to sets $\mathcal{B}^+$ and $\mathcal{B}^-$ is given in by

$$\begin{aligned} H(y_i; \hbar(\mathcal{L})) = &P(x_i \in \mathcal{B}^+)H(y_i|x_i \in \mathcal{B}^+; \hbar(\mathcal{L})) \\ &+P(x_i \in \mathcal{B}^-)H(y_i|x_i \in \mathcal{B}^-; \hbar(\mathcal{L})) \end{aligned} \tag{2}$$

It is clear that knowledge base $\mathcal{B} = \mathcal{B}^+ \cup \mathcal{B}^-$, and

$$P(x_i \in \mathcal{B}^+) + P(x_i \in \mathcal{B}^-) = 1 \tag{3}$$

Combining KBS and instance uncertainty, the objective function in Eq.(1) can be rewrite as

$$\underset{x_i \in \mathcal{U}}{argmax}(1 - P(x_i \in \mathcal{B}^+))H(y_i|x_i \in \mathcal{B}^-; \hbar(\mathcal{L})) \tag{4}$$

Eq.(4) represents the trade-off between minimizing the probability of falling into an oracle's KBS and maximizing the conditional mutual information.

Diverse Density for Characterizing KBS

To model an oracle's knowledge blind spot, we employ the diverse density concept (Maron and Lozano-Pérez 1998). We assume that some regions or concept set C exist to represent an oracle's KBS. We then define the diverse density of the target concept C as the probability that concept C is the target concept given the observed knowledge blind set ($\mathcal{B}^+$) and the acquired knowledge set ($\mathcal{B}^-$) of the oracle.

$$DD(C) = P(C|\mathfrak{b}_1^+, \mathfrak{b}_2^+, \ldots, \mathfrak{b}_p^+, \mathfrak{b}_1^-, \mathfrak{b}_2^-, \ldots, \mathfrak{b}_q^-) \tag{5}$$

Assume target concept set C consists of a number of small concepts $C = \{c_1, c_2, \ldots, c_m\}$, the conditional probability of each small concept c_k, given an instance $\mathfrak{b}_\tau$ in the knowledge base $\mathcal{B}$, can be defined as a feature value of $\mathfrak{b}_\tau$ (Chen, Bi, and Wang 2006). So we form a new set of feature for $\mathfrak{b}_\tau$ as

$$\mathbf{f}_C(\mathfrak{b}_\tau) = [f_{c_1}(\mathfrak{b}_\tau), \ldots, f_{c_m}(\mathfrak{b}_\tau)]^T = [P(c_1|\mathfrak{b}_\tau), \ldots, P(c_m|\mathfrak{b}_\tau)]^T \quad (6)$$

And its conditional probability is defined as

$$P(c_k|\mathfrak{b}_\tau) \propto \text{đ}(c_k, \mathfrak{b}_\tau) = exp(-\frac{||c_k - \mathfrak{b}_\tau||^2}{\sigma^2}) \quad (7)$$

We can use a sign function to define new labeling information for all instances in $\mathcal{B}$ as $\text{ł}(\mathcal{B}) = [sign(\mathfrak{b}_1^+), \ldots, sign(\mathfrak{b}_p^+), sign(\mathfrak{b}_1^-), \ldots, sign(\mathfrak{b}_q^-)]^T$. Then we form a well defined binary classification problem as

$$P(x_i \in \mathcal{B}^+) = \hbar(\mathbf{f}_C(\mathcal{B}), \text{ł}(\mathcal{B}))[\mathbf{f}_C(x_i); 1] \quad (8)$$

In Eq.(8), $\hbar(.)[\mathbf{f}_C(x_i); 1]$ denotes the class distribution of the classifier $\hbar(.)$ in classifying $\mathbf{f}_C(x_i)$ into class "1" and one can use any learning algorithm to train $\hbar(.)$.

Active Learning with Knowledge Blind Spot

Algorithm 1 lists major steps of the proposed framework for active learning with knowledge blind spot.

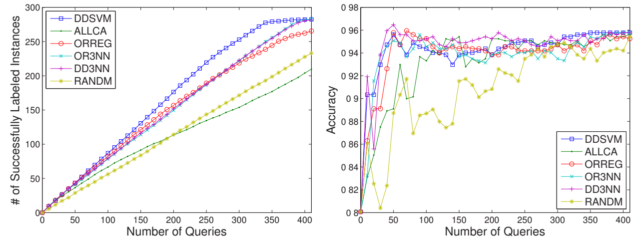
Algorithm 1 Active Learning with Knowledge Blind Spot

Input: (1) Unlabeled instances set: $\mathcal{U}$; (2) An oracle O; (3) A learner $\hbar(.)$; and (4) The number (or the percentage) of instances required to be labeled by the oracle O (*reqLabeled*)
Output: Labeled instance set $\mathcal{L}$

```
L ← Randomly label a tiny potion of instances from U
numLabeled ← |L|;  numQueries ← 0
B⁻ ← L;  B⁺ ← ∅;  B ← B⁺ ∪ B⁻
while numLabeled ≤ reqLabeled do
    i* ← argmax_{x_i∈U} H[x_i]  Use Eq.(4) and (8) select instance
        with the maximum utility value in U
    y_{i*} ← Query the label of x_{i*} from the oracle O
    if the oracle answers "I don't know the label" then
        B⁺ ← B⁺ ∪ x_{i*};
    else
        L ← L ∪ (x_{i*}, y_{i*});   B⁻ ← B⁻ ∪ x_{i*}
        numLabeled ← numLabeled + 1
    end if
    U ← U \ x_{i*};  B ← B⁺ ∪ B⁻  Update knowledge base B
    numQueries ← numQueries + 1
end while
```

Experiments

We use 10-fold cross validation for our experiments based on Wisconsin Breast Cancer Diagnostic Data Set and report the average results. To the best of our knowledge, there is no algorithm available for modeling KBS of the oracle for active learning. So we implement following KBS modeling approaches based on the framework in Algorithm 1: ALLCA(a case-based mining algorithm to model KBS of the oracle), ORREG(train a regression model for KBS in the original feature space), OR3NN(calculate $P(x_i \in \mathcal{B}^+)$ by using 3-NN in original feature space), DD3NN(calculate $P(x_i \in \mathcal{B}^+)$ by using 3-NN but in new feature space $\mathbb{R}_c$ transformed by using diversity density) and DDSVM(calculate $P(x_i \in \mathcal{B}^+)$ by using probability SVMs classifier trained on new feature space $\mathbb{R}_c$ by using diverse density). We also compare our framework with a baseline method: RANDM(randomly select and send instances to the oracle for labeling).



(a) Wisconsin Diagnostic (b) Wisconsin Diagnostic

Figure 1: The number of successfully labeled instances (a) and accuracies of the classifiers (b) *vs.* the number of active learning iterations (x-axis) trained from the data set $\mathcal{L}$ labeled by different methods.

In Figure 1(a), DDSVM outperforms others in avoiding the oracle's KBS and results in the most number of instances to be labeled by the oracle. Figure 1(b) presents the learning curves of the classifiers trained from instance sets labeled by different methods and indicates that almost all methods result in better performance than random labeling, this is mainly because that all approaches except Random involve the active learning component in the labeling process.

Conclusion

In this paper, we formulated a new active learning paradigm where the oracle, used for labeling, may be incapable of labeling some query instances. We used diverse density to model an oracle's KBS, and combined the uncertainty of each unlabeled instance and its likelihood of belonging to the KBS to select instances for labeling. Empirical results demonstrate the effectiveness of our approach.

Acknowledgments: This research is supported by Australian Research Council Future Fellowship under Grant No. FT100100971.

References

Chen, Y.; Bi, J.; and Wang, J. Z. 2006. MILES: Multiple-Instance Learning via Embedded Instance Selection. *IEEE PAMI* 28(12):1931–1947.

Donmez, P., and Carbonell, J. G. 2008. Proactive learning: cost-sensitive active learning with multiple imperfect oracles. In *Proc. of CIKM*, CIKM '08, 619–628.

Maron, O., and Lozano-Pérez, T. 1998. A framework for multiple-instance learning. In *Proc. of NIPS*, NIPS '97, 570–576.

Yan, Y.; Rosales, R.; Fung, G.; and Dy, J. 2011. Active learning from crowds. In *Proc. of ICML*, ICML '11, 1161–1168.