# Journal of Applied Sciences

ISSN 1812-5654

science alert

ANSInet
an open access publisher
http://ansinet.com

Journal of Applied Sciences 13 (16): 3257-3262, 2013
ISSN 1812-5654 / DOI: 10.3923/jas.2013.3257.3262


# Piecewise Function Approximation and Vertex Partitioning Schemes for Multi-dividing Ontology Algorithm in Auc Criterion Setting (II)

[1]Li Yan, [2]Wei Gao and [1]JunSheng Li
[1]School of Engineer, Honghe University, Mengzi, 661100, People Republic of China
[2]School of Information Science and Technology, Yunnan Normal University, Kunming, 650500, People Republic of China

**Abstract:** As the extension of Gao *et al.* (2013), we present the partitioning scheme made of dyadic cubes, the approximation result is thereby obtained concern such partitioning. It is highlighted in this study that AUC criterion multi-dividing ontology algorithm with tree ontology structure and specific split rule has good statistic characteristics, and show that the algorithm under these partitioning schemes are convergent.

**Key words:** Ontology, multi-dividing, ROC optimization, AUC criterion, dyadic cube

## INTRODUCTION

This study is an extension of Gao *et al.* (2013). We continue discuss the theoretical problem of AUC criterion multi-dividing ontology algorithm. Notation and terminology used but undefined in this study can be found in Gao *et al.* (2013).

## PARTITIONING SCHEME MADE OF DYADIC CUBES

The first result we present in this study determines the multi-dividing ontology scoring function, constant on every cell of the initial partition $P = \{C_t\}_{1\le t\le T}$, that has maximum empirical AUC. This result immediately draws from empirical distribution and we skip the detail proof.

**Theorem 1:** Let $P = \{C_t\}_{1\le t\le T}$ be a partition of the ontology vertex space and:

$$\hat{f}^*(v)=2\cdot \amalg\{v\in L\}+\amalg\{v\in R\}$$

be the multi-dividing ontology score function determined by the partition-based splitting rule based on P and the sampling set S. Let subset $C\subset V$ be formed by a union of cells in P. Then, for any ontology score fuction $f = 2\cdot\amalg_C+\amalg_{V/C}$, we have:

$$\widehat{AUC}(f)\le \widehat{AUC}(\hat{f}^*)$$

For convenience, we assume that $V = [0, 1]^q$ and consider subpartitions of the partition P (j) made of dyadic cubes of side length $2^{-j}$, that is to say, the subsets have the form:

$$\prod_{i=1}^{q}[t_1/2^j,(t_1+1)/2^j]$$

for $0\le t_l<2^j$, $l\in\{1,\ldots, q\}$ and $|P(j)| = 2^{jq}$. Let $\hat{L}_j = L$ be the output of the partition-depend splitting ontology function from P (j) and:

$$\hat{f}_j^*(v)=2\prod\{v\in \hat{L}_j\}+2\prod\{v\in \hat{R}_j\}$$

be the related ontology score function. Formally, we get the next result for such partitioning scheme.

**Theorem 2:** For any $j\ge 1$, let $P_{2,j}$ be the set of partitions of ontology vertex set made of two non-empty sets obtained as unions of dyadic cubes of side length$2^{-j}$. Suppose that:

$$p^{a,b}\in[\underline{p}^{a,b},\overline{p}^{a,b}] \text{ with } 0<\underline{p}^{a,b}<\overline{p}^{a,b}<1$$

There exist constants $c^{a,b}<\infty$ depending on $\underline{p}^{a,b}$ b and $\overline{p}^{a,b}$ such that for any $\delta\in(0, 1)$, $j\ge 1$ and $n\ge 1$ large enough, we have with probability at least $1-\delta$:

$$AUC(f_1^*)-AUC(\hat{f}_j^*)\le\{AUC(f_1^*)-\max_{f\in F_{P_{2,j}}} AUC(f)\}+\sum_{a=1}^{k-1}\sum_{b=a+1}^{k} c^{a,b}\cdot\frac{2^{jq}}{\sqrt{n_a+n_b}}$$

**Proof:** For any $j\ge 1$, let $C_j$ be the set of non-empty subsets of ontology vertex set such that it can be formed from the $2^{jq}$ dyadic cubes of side length $2^{-j}$, except $V = [0, 1]^q$ itself. For each $j\ge 1$, let $\tilde{L}_j^*$ be the true left cell depended on the partition $C_j$ of the set V and $\tilde{L}_j^*$ be the empirical counterpart. The related multi-dividing ontology score functions are denoted by:

**Correspnding Author:** Li Yan, School of Engineer, Honghe University, 661100, Mengzi, People Republic of China

$$\tilde{f}_j^*(v) = 2 \cdot \amalg\{v \in \tilde{L}_j^*\} - 1$$

and:

$$\tilde{f}_j^*(v) = 2 \cdot \amalg\{v \in \tilde{L}_j^*\} - 1$$

Specially, we have:

$$\begin{aligned} & AUC(f_1^*) - AUC(\hat{f}_j^*) \\ & = \{AUC(f_1^*) - AUC(\tilde{f}_j^*)\} \\ & + \{AUC(\tilde{f}_1^*) - \widehat{AUC}(\tilde{f}_j^*)\} \\ & + \{\widehat{AUC}(\tilde{f}_1^*) - \widehat{AUC}(\hat{f}_j^*)\} \\ & + \{\widehat{AUC}(\hat{f}_1^*) - AUC(\hat{f}_j^*)\} \\ & \leq \{AUC(f_1^*) - AUC(\tilde{f}_j^*)\} \\ & + 2 \sup_{f \in F_{P_{2,j}}} \left| \widehat{AUC}(f) - AUC(f) \right| \end{aligned}$$

Let symmetric kernel:

$$\begin{aligned} & h_f((v_i, y_i), (v_j, y_j)) \\ & = \amalg\{(y_1 - y_2)(f(v_1) - f(v_2)) > 0\} \\ & + \frac{1}{2} \amalg\{y_1 \neq y_2, f(v_1) = f(v_2)\} \end{aligned}$$

and U-statistic:

$$\hat{U}_n(f) = \frac{2}{n(n-1)} \sum_{1 \leq i < j \leq n} h_f((V_i, Y_i), (V_j, Y_j))$$

with expectation:

$$U(f) = AUC(f) \sum_{a=1}^{k-1} \sum_{b=a+1}^{k} 2p^{a,b}(1 - p^{a,b})$$

The empirical $\widehat{AUC}(f)$ can be expressed as:

$$\widehat{AUC}(f) = \hat{U}_n(f) \sum_{a=1}^{k-1} \sum_{b=a+1}^{k} \frac{(n_a + n_b)(n_a + n_b - 1)}{2 n_a n_b}$$

By virtues of Hoeffding's exponential inequality for U-statistics and the union bound, we infer that for any $\delta \in (0, 1)$ and $n \geq 1$, with probability at least $1-\delta$:

$$\sup_{f \in F_{P_{2,j}}} \left| \hat{U}_n(f) - U(f) \right| \leq \sqrt{\frac{\log(\delta / (2|P_{2,j}|))}{2n}}$$

Then, the desired bound follows from the fact that:

$$\left| \widehat{AUC}(f) - AUC(f) \right| \leq \sum_{a=1}^{k-1} \sum_{b=a+1}^{k} \{ \frac{1}{2\underline{p}^{a,b}(1 - \overline{p}^{a,b})} \left| \hat{U}_n(f) - U(f) \right|$$

$$+ \frac{1}{2} \{ \left| \frac{1}{p^{a,b}} - \frac{n_a + n_b}{n_a} \right| + \left| \frac{1}{1 - p^{a,b}} - \frac{n_a + n_b}{n_b} \right| \} \}$$

and using the standard Hoeffding probability inequality to control the fluctuations of:

$$\frac{n_a}{n_a + n_b}$$

around:

$$p^{a,b} \in [\underline{p}^{a,b}, \overline{p}^{a,b}]$$

## PARTITIONING SCHEME FOR ONTOLOGY WITH TREE STRUCTURE

In this section, based on a training sample set S, we consider that ontology graph is a tree with depth D. Using a cost-complexity modifying scheme, we select a certain tree-structured subpartition of P ($\Gamma$) = $\{C_{D,t}, 0 \leq t < 2^D\}$. Intuitively, we expect that the expected AUC of the resulting multi-dividing ontology score function is larger than the $f_D$ (v).

In what follows, we propose two technologies for modifying an ontology tree. For $0 \leq d \leq D$ and $0 \leq t < 2^D$, to each cell $C_{d,t}$, we assign a weight $\omega$ ($C_{d,t}$) in order that the following constraints are satisfied:

- For all $d \in \{0, ...c, D\}$ and $t \in \{0, ...c, 2^D-1\}$, we have $\omega$ $(C_{d,t}) \in [0, 1]$
- If $\omega$ $(C_{d,t})$ = 1, then we have $\omega$ $(C_{d1,t1})$= 1 for each cell $C_{d1,t1}$ satisfies $C_{d,t} \subset C_{d1,t1}$

The set of weights $\omega$ satisfying above two conditions said admissible and determines the vertices of a subtree $\Gamma(\omega)$ of the original ontology tree structure $\Gamma$. A cell $C_{d',t'}$ is said terminal if $\omega$ $(C_{d,t})$ = 1 and $\omega$ $(C_{d',t'})$ = 0 for each cell $C_{d',t'} \subset C_{d',t'}$. Terminal cells correspond to the outer leaves of the ontology tree $\Gamma$ ($\omega$) and form a partition P ($\Gamma(\omega)$) of the feature space V. For two fixed admissible weights $\omega_1$ and $\omega_2$, P ($\Gamma(\omega_1)$) is a subpartition of P ($\Gamma(\omega_2)$) if $\{C_{d,t}: \omega_1 (C_{d,t}) = 0\} \subset \{C_{d,t}: \omega_2 (C_{d,t}) = 0\}$ and we denote as $\Gamma(\omega_1) \subseteq \Gamma(\omega_2)$. The algorithm modifying stage consists of choosing those terminal leaves and building the muliti-dividing ontology score function:

$$f_\omega(v) = \sum_{C_{d,t} \in P(\Gamma(\omega))} (2^D - 2^{D-d} t) \cdot \prod\{v \in C_{d,t}\}$$

When distributions $G^{a,b^*}$, $H^{a,b^*}$, $\bar{G}^{a,b,*}$ and $\bar{H}^{a,b,*}$ are known, the optimal modifying tree in the AUC sense can be described by:

$$\omega^* = \underset{\omega}{\arg\max} AUC(f_\omega)$$

where, the maximum is taken over all admissible sets of weights ω. In particular reality, however, the distributions are unknown. Thus, we replace AUC ($f_\omega$) by an estimate:

$$\widehat{AUC}'(f_\omega)=\sum_{a=1}^{k-1}\sum_{b=a+1}^{k}\{\frac{1}{n'_a n'_b}\sum_{i:Y'_i=a}\sum_{j:Y'_j=b}\prod\{f_\omega(V_i)>f_\omega(V_j)\} + \frac{1}{2n'_a n'_b}\sum_{i:Y'_i=a}\sum_{j:Y'_j=b}\prod\{f_\omega(V_i)=f_\omega(V_j)\}\}$$

depended on a sample set $S'_{n'} = \{S'_1,\dots,S'_k\}$ formed of i.i.d. copies of the pair (V, Y ), where:

$$n'_a=\sum_{i=1}^{n'}\prod\{Y'_i=a\} \text{for} 1\le a\le k$$

Our basic idea of dealing with practical implement is to add to the optimistic training performance estimate:

$$\widehat{AUC}(f_\omega)$$

with penalizes term. That is:

$$\widehat{CPAUC}(f_\omega,\lambda)=\widehat{AUC}(f_\omega)-\lambda\cdot|P(\Gamma(\omega))|$$

where, balance parameter $\lambda\ge 0$ is regarded as a trade-off factor between model complexity and training performance accuracy. We look for the sub-tree reaching the maximal complexity-penalized empirical AUC:

$$\omega^*_\lambda=\arg\max_\omega \widehat{CPAUC}(f_\omega,\lambda)$$

We consider the choice of λ and discuss this issue with cross-validation technology. The next result reveals that there exist finite sequences of ontology sub-trees of the original ontology tree Γ containing all Γ ($\omega_\lambda^*$). We omit the detail proof since it similarly as Theorem 3.10 in Breiman *et al.* (1984) and mail result in Ripley (1996).

**Theorem 3:** For a given ontology tree structure Γ, there exist a finite increasing sequence of constants $0=\lambda_0<\lambda_1<\dots c<\lambda_m=\infty$ satisfies that root = Γ ($\omega^*_{\lambda m}$)⊆...⊆Γ ($\omega^*_{\lambda 0}$) = Γ and for any j∈{1,…, m} and any $\lambda\in[\lambda_{j-1},\lambda_j]$, we get:

$$T(\omega^*_\lambda)=T(\omega^*_{\lambda j})$$

Using the non-parametric model selection approach, we now propose a strategy for selecting an ontology sub-tree Γ (ω) in a sample-driven version with largest possible AUC. Here the AUC criterion multi-dividing ontology scheme consists of maximizing:

$$\widehat{CPAUC}(f_\omega)=\widehat{AUC}(f_\omega)-\text{pen}(|P(\Gamma(\omega))|,n)$$

where, pen (T, n) is a fixed and explicit penalty term satisfies that no cross validation or resampling is demanded by the choosing procedure. Let:

$$\tilde{f}^*_n=f_{P(\Gamma(\tilde{\omega}^*_n))}$$

with:

$$\tilde{\omega}^*_n=\arg\max_{\omega \text{ admissible}}\widehat{CPAUC}(f_\omega)$$

Our ideal to an adequate choice for the penalty term based on establishing a distribution-free bound for the quantity:

$$E[\sup_{\omega||P(\Gamma(\omega))|=T}|\widehat{AUC}(f_\omega)-AUC(f_\omega)|]$$

with T ∈{1,..., $2^D$}.

In what follows, we consider two situations, corresponding to different ways of performing the optimization step in the growing stage and obtaining differ penalties for model selection:

- $R_1$: Splits are yielded through the ontology procedure with at most κ perpendicular cuts, κ≥1
- $R_2$: Let $V = [0, 1]^q$ be feature space and splits are obtained by partition-based function from the set of dyadic cubes:

$$\prod_{m=1}^{q}[t_m 2^{-J},(t_m+1)2^{-J}]$$

for all m∈{1, ..., q}, where $0\le t_m<2^J$

The result we stated following describes the performance of the ontology score function $F^*_n$ based on structural AUC maximization in each of these cases.

**Theorem 4:** Assume that $p^{a,b}\in[\underline{p}^{a,b},\overline{p}^{a,b}]$ with $0<\underline{p}^{a,b}<\overline{p}^{a,b}<1$ and for all each pair of (a, b), T∈{1,… , $2^D$} and n≥1 the penalty term is selected as follows:

- If splits are optimized using the $R_1$ rule, then for any $(T,\kappa)\in N^{*2}$:

$$\text{pen}(T,n)=\sum_{a=1}^{k-1}\sum_{b=a+1}^{k}\{\frac{1}{\underline{p}^{a,b}(1-\overline{p}^{a,b})}\cdot\sqrt{32\frac{\log(16((n_a+n_b+1)q)^{2T\kappa})+T}{n_a+n_b}}\}$$

- If splits are optimized using the $R_2$ rule, then for any $(T,J)\in N^{*2}$:

$$pen(T,n)=\sum_{a=1}^{k-1}\sum_{b=a+1}^{k}\frac{1}{p^{a,b}(1-\bar{p}^{a,b})}\sqrt{\frac{\log(4T^{2Jq})+T}{2(n_a+n_b)}}$$

Then, there exist a positive constant C such that the expected deficit of AUC of the multi-dividing ontology sub-tree maximizing the complexity-penalized area under the ROC curve is bounded by:

$$AUC^* - E(AUC(\tilde{f}_n^*)) \le \inf_{1\le T\le 2^D}\{C\cdot pen(T,n) + \{AUC^* - \sup_{\omega|P(\Gamma(\omega))|=T} AUC(f_\omega)\}\}$$

Let $T\ge 1$, we denote by $P_\Gamma$ (T) the set of all tree-structured partitions of the feature space $V\subset\mathbb{R}^q$ with $T\ge 1$ non empty cells and let $F\Gamma(T) = \cup_{p\in Pt(T)}F_P$ be the set of piecewise constant ontology score functions associated to such partitions. Denote the empirical AUC maximizer over $F_\Gamma$ (T) as:

$$\hat{f}_{n,T}^* = \arg\max_{f\in F_\Gamma(T)} \widehat{AUC}(f)$$

The desired inequality of Theorem 4 is heavily depended on the lemma below.

**Lemma 1:** Assume that the hypotheses of Theorem 4 are satisfied:

- If splits are optimized with $R_1$ rule and the penalization is selected accordingly, then for any $(T, \kappa)\in N^{*2}$, we have:

$$P\{\sup_{f\in F_\Gamma(T)} AUC(f) - AUC(\tilde{f}_n^*)\ge\varepsilon\}\le \sum_{a=1}^{k-1}\sum_{b=a+1}^{k}\{16((n_a+n_b+1)q)^{2T\kappa}e^{-n(p^{a,b})^2(1-\bar{p}^{a,b})^2\varepsilon^2/512} + e^{-n(p^{a,b})^2(1-\bar{p}^{a,b})^2\varepsilon^2/128}\}$$

- If splits are optimized with $R_2$ rule and the penalization is selected accordingly, then for any $(T, J)\in N^{*2}$, we get:

$$P\{\sup_{f\in F_\Gamma(T)} AUC(f) - AUC(\tilde{f}_n^*)\ge\varepsilon\}\le \sum_{a=1}^{k-1}\sum_{b=a+1}^{k}\{4T^{2Jq}e^{-n(p^{a,b})^2(1-\bar{p}^{a,b})^2\varepsilon^2/8} + e^{-n(p^{a,b})^2(1-\bar{p}^{a,b})^2\varepsilon^2/2}\}$$

**Proof of Lemma 1:** In terms of main result of Breiman *et al.* (1984) and Ripley (1996), we confirm that for each $\varepsilon>0$ and each $T\ge 1$:

$$P\{\sup_{f\in F_\Gamma(T)} AUC(f) - AUC(\tilde{f}_n^*)\ge\varepsilon\}\le P\{\sup_{l\ge 1}\overline{CPAUC}(\hat{f}_{n,l}^*) - AUC(\tilde{f}_n^*)\ge\frac{\varepsilon}{2}\} + P\{\sup_{l\ge 1}\overline{CPAUC}(\hat{f}_{n,l}^*) - AUC(f)\ge\frac{\varepsilon}{2}\}$$

Thus, the first term on the right side of the above inequality can be represented and bounded as follows:

$$\begin{aligned} &P\{\overline{CPAUC}(\hat{f}_{n,l}^*) - AUC(\tilde{f}_n^*)\ge\frac{\varepsilon}{2}\}\\ &\le P\{\inf_{l\ge 1}\left|\overline{CPAUC}(\hat{f}_{n,l}^*) - AUC(\hat{f}_n^*)\right|\ge\frac{\varepsilon}{2}\}\\ &\le \inf_{l\ge 1} P\{\left|AUC(\hat{f}_{n,l}^*) - \widehat{AUC}(\hat{f}_{n,l}^*)\right|\ge\frac{\varepsilon}{2} + pen(l,n)\}\\ &\le \inf_{l\ge 1} P\{\sup_{f\in F_\Gamma(T)}\left|AUC(\hat{f}_{n,l}^*) - \widehat{AUC}(\hat{f}_{n,l}^*)\right|\ge\frac{\varepsilon}{2} + pen(l,n)\} \end{aligned} \qquad (1)$$

About the second term, by pen $(T, n)\le\varepsilon/4$, we deduce:

$$\begin{aligned} &P\{\sup_{f\in F_\Gamma(T)} AUC(f) - \sup_{l\ge 1}\overline{CPAUC}(\hat{f}_{n,l}^*)\ge\frac{\varepsilon}{2}\}\\ &\le P\{\sup_{f\in F_\Gamma(T)} AUC(f) - \overline{CPAUC}(\hat{f}_{n,l}^*)\ge\frac{\varepsilon}{2}\}\\ &\le P\{\sup_{f\in F_\Gamma(T)} AUC(f) - \widehat{AUC}(\hat{f}_{n,l}^*)\ge\frac{\varepsilon}{4}\}\\ &\le P\{\sup_{f\in F_\Gamma(T)}\left|\widehat{AUC}(f) - AUC(f)\right|\ge\frac{\varepsilon}{4}\} \end{aligned} \qquad (2)$$

We first deal with the $R_1$ rule, where optimization steps are performed using at most κ perpendicular splits. Note that the n-th VC shattering coefficient of the class:

$$\Phi = \cup_{P\in P_\Gamma(T)}\{A\times B : (A,P)\in P^2\}$$

of subsets of V×V is thus bounded as follows:

$$F(\Phi, n)\le((na+nb+1)q)^{2T\kappa}$$

Using Vapnik-Chervonenkis inequality and characteristics of the kernel set $\{h_f\text{-}U(f)\}_{f\in F\Gamma(T)}$, this obtains that for each ε and $n\le 1$:

$$P\{\sup_{f\in F_\Gamma(T)}\left|\hat{U}_n(f) - U(f)\right|\ge\varepsilon\}\le 8((n+1)q)^{2K\kappa}e^{-n\varepsilon^2/32}$$

Thus, for large number n, we get:

$$P\{\sup_{f\in F_\Gamma(T)}\left|\widehat{AUC}(f) - AUC(f)\right|\ge\varepsilon\}\le\sum_{a=1}^{k-1}\sum_{b=a+1}^{k}\{16((n_a+n_b+1)q)^{2K\kappa}\cdot e^{-n(p^{a,b})^2(1-\bar{p}^{a,b})^2\varepsilon^2/32}\} \qquad (3)$$

Combined with (1), we deduce:

$$P\{\overline{CPAUC}(\hat{f}^*_{n,1}) - AUC(\tilde{f}^*_n) \ge \frac{\varepsilon}{2}\}$$
$$\le \sum_{a=1}^{k-1}\sum_{b=a+1}^{k}\sum_{l=1}^{2^D}\{16((n_a+n_b+1)q)^{2T\kappa}$$
$$\cdot e^{-n(\underline{p}^{a,b})^2(1-\overline{p}^{a,b})^2(\frac{\varepsilon}{2}+pen(T,n))^2/32}\}$$
$$\le \sum_{a=1}^{k-1}\sum_{b=a+1}^{k}\{e^{-n(\underline{p}^{a,b})^2(1-\overline{p}^{a,b})^2\varepsilon^2/128}$$
$$\cdot\sum_{l=1}^{2^D}16((n_a+n_b+1)q)^{2\kappa}e^{-n(\underline{p}^{a,b})^2(1-\overline{p}^{a,b})^2 pen(T,n)^2/32}\}$$
$$\le \sum_{a=1}^{k-1}\sum_{b=a+1}^{k}\{e^{-n(\underline{p}^{a,b})^2(1-\overline{p}^{a,b})^2\varepsilon^2/128}\sum_{l\ge 1}e^{-T}\}$$
$$\le \sum_{a=1}^{k-1}\sum_{b=a+1}^{k}e^{-n(\underline{p}^{a,b})^2(1-\overline{p}^{a,b})^2\varepsilon^2/128}$$

Combining with (2) and (3), we obtain:

$$P\{\sup_{f\in F_\Gamma(T)} AUC(f) - \sup_{l\ge 1}\overline{CPAUC}(\hat{f}^*_{n,1}) \ge \frac{\varepsilon}{2}\}$$
$$\le \sum_{a=1}^{k-1}\sum_{b=a+1}^{k}\{16((n_a+n_b+1)q)^{2T\kappa}$$
$$\cdot e^{-n(\underline{p}^{a,b})^2(1-\overline{p}^{a,b})^2\varepsilon^2/32}\}$$

The first assertion of the Lemma 1 is thus proved.

Now, we assume that $V = [0, 1]^q$ and cells are yielded as unions of dyadic cubes of side length $2^J$ , $J\in\mathbb{N}$. In the $R_2$ case, in view of Hoeffding's inequality for U-statistics combined with the union bound and the fact that $|h_f: f\in F\Gamma (T)|\le T^{2Jq}$, for any $\varepsilon$ and $n\ge 1$, we have:

$$P\{\sup_{f\in F_\Gamma(T)}\left|\hat{U}_n(f) - U(f)\right| \ge \varepsilon\} \le 2T^{2Jq}e^{-2n\varepsilon^2}$$

For n large enough, we infer:

$$P\{\sup_{f\in F_\Gamma(T)}\left|\widehat{AUC}(f) - AUC(f)\right| \ge \varepsilon\}$$
$$\le \sum_{a=1}^{k-1}\sum_{b=a+1}^{k}4T^{2Jq}e^{-2n(\underline{p}^{a,b})^2(1-\overline{p}^{a,b})^2\varepsilon^2}$$

We omit the remainder discussion since it is totally similar to the one in the $R_1$ case.

**Proof of Theorem 4:** We deduce:

$$AUC^* - E[AUC(\tilde{f}^*_n)]$$
$$= \inf_{T\ge 1}\{(AUC^* - \sup_{f\in F_\Gamma(T)} AUC(f))$$
$$+(\sup_{f\in F_\Gamma(T)} AUC(f) - E[AUC(\tilde{f}^*_n)])\}$$

Hence:

$$(\sup_{f\in F_\Gamma(T)} AUC(f) - E[AUC(\tilde{f}^*_n)])^2$$
$$\le u + \int_{t=u}^{\infty} P\{(\sup_{f\in F_\Gamma(T)} AUC(f)$$
$$- AUC(\tilde{f}^*_n))^2 > t\}dt$$

By integrating the tail bounds stated in Lemma 1 and taking $u = C\,(pen\,(K, n))^2$, we get the desire result.

The next result is immediate corollary of Theorem 4. It is stated that under mild assumptions, AUC-consistent ontology sub-trees are exist. We skip the proof.

**Corollary 1:** Assume that assumptions of Theorem 4 are satisfied and that there exist a sequence $\Gamma_n\ (\omega_n)$ of subtrees of the ontology trees $\Gamma_n$ produced such that $E\ (AUC(f^*_n))\to AUC^*$, as $n\to\infty$. In addition, we assume that:

- If $\Gamma_n$ is yielded by the $R_1$ rule with $\kappa = \kappa\ (n)$ axis-parallel splits, then as $n\to\infty$, we have:

$$\kappa\ (n).E[|P(T_n(\omega_n))|] = o\ (n/\log n)$$

- if $\Gamma_n$ is yielded by the $R_2$ rule based on dyadic hypercubes of side length $2^{\Sigma J}$ with $J = J\ (n)$, then as $n\to\infty$, we have:

$$E[|P\ (\Gamma n(\omega n))|] = o\ (n)$$

and:

$$J\ (n) = o\ (n/\log n)$$

Hence:

$$\lim_{n\to\infty} E[AUC(\tilde{f}^*_n)] = AUC*$$

In the $R_2$ case, particularly, we have following approximation conclusion.

**Theorem 5:** Suppose that assumptions of Theorem 4 are fulfilled and that the ontology tree $\Gamma_n$ is given via the ontology algorithm with the $R_2$ rule depended on dyadic hypercubes of side length $2^{-J}$ , $J = J\ (n) \ge 1$ and depth $D_n$. If, in addition, as $n\to\infty$, $J\ (n) \to\infty$ and the sizes of the cells $\{C\ (n)\!: t = 0,..., 2^{D_n+1}-1\}$ of the related partition $P\ (\Gamma_n)$ uniformly contract to zero such that:

$$\max_{0\le t<2^{D_n+1}} \mu(C_t^{(n)}) \to 0$$

then, the modified ontology trees $\Gamma_n\ (\tilde{\omega}^*_n)$ yielded from $\Gamma_n$ are AUC consistent.

**Proof of Theorem 5:** We consider the $R_2$ case. By Corollary 1, it suffices to show that:

$$\lim_{n\to\infty}\sup_{f\in F_{\Gamma_n}} AUC(f) = AUC^*$$

Let $\{C_{dn,t}\}_{0\le t<2Dn}$ be the cells of the partition $P_{Dn}$ corresponding to the ontology tree $\Gamma_n$ output by multi-dividing algorithm and $f_{Dn}$ be the related multi-dividing ontology score function. Since, in the $R_2$ case, $f_{Dn}$ and $\hat{\eta}_{Pdn}$ produce the same performance, we obtain:

$$\begin{aligned} &AUC^* - \sup_{f\in F_n} AUC(f) \\ &\le \sum_{a=1}^{k-1}\sum_{b=a+1}^{k}\{\frac{E[\left|\eta^{a,b}(V) - \hat{\eta}^{a,b}_{P_{D_n}}(V)\right|]}{2p^{a,b}(1-p^{a,b})} \\ &+\frac{1}{4p^{a,b}(1-p^{a,b})}\sum_{t=1}^{D_n} G^{a,b}(C_{n,t})\} \end{aligned} \quad (4)$$

where:

$$G^{a,b}(C_{n,t}) = E[\left|\eta^{a,b}(V) - \eta^{a,b}(V')\right| \cdot \mathbb{I}\{(V,V')\in C^2_{n,t}\}]$$

Furthermore, we observe that:

$$\begin{aligned} &\sum_{t=1}^{D_n} G(C_{n,t}) \le \sum_{t=1}^{D_n} \mu(C_{n,t})^2 \\ &\le \max_{0\le t<D_n} \mu(C_{n,t}) \end{aligned}$$

Note that the term on the right hand side of (4) vanishes as $n\to\infty$. In view of the argument of Theorem 6.1 in Breiman *et al.* (1984) ensures that the term on the left hand side goes to 0 as $n\to\infty$. Thus, we complete the proof of Theorem 5.

## CONCLUSION

In this study, we display two theoretical performances for AUC criterion multi-dividing ontology problem: (1) Partitioning scheme made of dyadic cubes manifest a good approximation properties; (2) If ontology graph has tree structure, then there exist good vertex partition scheme under two split rules, and we yield the convergence results for such setting. To the best of our knowledge, these results are the first to state in ontology field.


## ACKNOWLEDGMENT

First we thank the reviewers for their constructive comments in improving the quality of this paper. This work was supported in part by the Yunnan Province basic research for application (2013fz127). We also would like to thank the anonymous referees for providing us with constructive comments and suggestions.


## REFERENCES


Breiman, L., J. Friedman, R.A. Olshen and C.J. Stone, 1984. Classification and Regression Trees. Wadsworth, USA., ISBN-10: 0534980546, Pages: 368.

Gao, W., L. Yan and L. Liang, 2013. Piecewise function approximation and vertex partitioning schemes for multi-dividing ontology algorithm in AUC criterion setting. Int. J. Comput. Appl. Technol.,

Ripley, B.D., 1996. Pattern Recognition and Neural Networks. Cambridge University Press, Cambridge.