基于模糊积分的分类器集成方法

Junhai Zhai

College of Mathematics and Information Science, Hebei University

mczjh@hbu.cn

2017 年 2 月 8 日

给定训练集T，$\Omega=\{\omega_1,\omega_2,\cdots,\omega_k\}$是类标集合，$D=\{D_1,D_2,\cdots,D_l\}$是从$T$训练出的$l$个分类器集合。对于任意的测试样例$x$，$D_i(x)=(\mu_{i1}(x),\mu_{i2}(x),\cdots,\mu_{ik}(x))$。其中，$\mu_{ij}(x)\in[0,1]$表示分类器$D_i(1\leq i\leq l)$将测试样例$x$分类到$j^{th}(1\leq j\leq k)$ 类的**支持度(隶属度)**，$\sum_{j=1}^{k}\mu_{ij}(x)=1$。

定义1. 给定测试样例x，称下面的$l \times k$阶的矩阵DM为x的**决策矩阵**。

$$DM(x) = \begin{bmatrix} \mu_{11}(x) & \cdots & \mu_{1j}(x) & \cdots & \mu_{1k}(x) \\ \vdots & & \vdots & & \vdots \\ \mu_{i1}(x) & \cdots & \mu_{ij}(x) & \cdots & \mu_{ik}(x) \\ \vdots & & \vdots & & \vdots \\ \mu_{l1}(x) & \cdots & \mu_{lj}(x) & \cdots & \mu_{lk}(x) \end{bmatrix}_{l \times k} \tag{1}$$

矩阵DM的i^{th}行表示分类器D_i将x分类为j^{th}类的支持度；矩阵DM的j^{th}列表示x被不同的分类器分类到j^{th}类的支持度。

定义2. 给定分类器集合$D=\{D_1, D_2, \cdots, D_l\}$，$P(D)$是$D$的幂集。$D$上的**模糊测度**$g$定义为满足如下两个条件的函数$g: P(D) \to [0,1]$。

(1) $g(\varnothing)=0$, $g(D)=1$;

(2) $\forall A, B \subseteq D$，若$A \subset B$，则$g(A) \leq g(B)$.。

如果$\forall A, B \subseteq D$，且$A \cap B = \varnothing$，下式成立，则称g为λ–模糊测度。

$$g(A \cup B) = g(A) + g(B) + \lambda g(A)g(B). \tag{2}$$

其中，$\lambda > -1$，且$\lambda \neq 0$，它的值由下式确定：

$$\lambda + 1 = \prod_{i=1}^{l}(1 + \lambda g_i). \tag{3}$$

其中，g_i表示在单个分类器上的模糊测度，称为**模糊密度**。
说明：理论上已经证明：不管集成几个分类器，即不论l等于几，满足条件的解λ只有一个。

确定g_i的方法通常有下列三种：

$$\begin{aligned}
&(1)\ g_i = p_i;\\
&(2)\ g_i = \frac{p_i}{2};\\
&(3)\ g_i = \delta\frac{p_i}{\sum_{j=1}^{l} p_j}, \delta \in [0,\ 1].
\end{aligned} \tag{4}$$

其中，p_i是分类器D_i在验证集的验证精度(或测试集的测试精度)。

说明：

(a) 模糊密度的3种取法，虽然值有较大的差异，但对最终结果影响不大(有人做过这方面的实验)。

(b) 文献中用第三种取法的较多，δ取值越大，越突出单个分类器的作用；δ取值越小，越突出集成分类器的作用。

定义3. 给定分类器集合$D=\{D_1,D_2,\cdots,D_l\}$，g是D上的模糊测度，函数$h:D\rightarrow R^+$关于g的**Choquet积分**定义为：

$$(C)\int hd\mu=\sum_{i=1}^{l}(h(D_i)-h(D_{i-1}))g(A_i). \tag{5}$$

其中，$0\leq h(D_1)\leq h(D_2)\leq\cdots\leq h(D_l)\leq 1$，$h(D_0)=0$，$A_i=\{D_1,D_2,\cdots,D_i\}\subseteq D$，$g(A_0)=0$。

说明：

(a) 定义中的排序也可以由大到小，但被积函数相应地变为$(h(D_{i-1})-h(D_i))$。即，要保证积分值非负。

(b) 类似地，下面算法中的第6步，也可以由小到大排序。相应地，第12步中的被积函数变为$\left(d_{i_tj}(x)\text{-}d_{i_{t-1}j}(x)\right)$

基于Choquet模糊积分的分类器集成算法如下：

算法 1: 基于模糊积分的分类器集成方法

输入: 训练集$T=\{(x_i,y_i)|x_i\in R^d, y_i\in Y\}$, $i=1,2,\cdots,n$, Y是类标的集合, $|Y|=k$, 测试样例x

输出: x的类标

用训练集T训练l个分类器(要求输出为后验概率);

用公式(4)中的某种方法确定l个基本分类器$D=\{D_1,D_2,\cdots,D_l\}$的模糊密度$g_1,g_2,\cdots,g_l$;

用公式(3)计算λ;

对测试样例x，用公式(1)计算决策矩阵$DM(x)$;

for $(j=1; j\le k; j=j+1)$ **do**

  // 对$DM(x)$的各列独立排序．即，对j^{th}列排序时和其他列无关；还需要注意分类器也随之排序

  对$DM(x)$的j^{th}列由大到小排序，记为$\left(d_{i_1 j}, d_{i_2 j,\cdots,} d_{i_l j}\right)$;

  令$g(A_1)=g_{i_1}$;

  for $(t=2; t\le l; t=t+1)$ **do**

    用公式(2)递归计算$g(A_t)=g_{i_t}+g(A_{t\text{-}1})+\lambda g_{i_t}g(A_{t\text{-}1})$;

  end

  // 因为对于不同的列，排序的结果可能不同，A_i也就不同，所以计算得到的$g(A_i)$也就不同．

  用公式(5)计算$\mu_j(x)=\sum_{t=2}^{l}\left(d_{i_{t-1}j}(x)\text{-}d_{i_t j}(x)\right)g(A_{t\text{-}1})$;

end

// 对每一类，都要计算一个隶属度。计算第j类的隶属度时，要对决策矩阵的第j列由大到小排序。所以，排序要做k次。k为类别数，l为基本分类器个数。决策矩阵是l行k列的矩阵，其阶数$l\times k$一般都不高。

用公式$j^*=argmax_{1\le j\le k}\{\mu_j(x)\}$确定$x$的类标$j^*$;

输出x的类标j^*。

The End