

# 腰椎间盘突出早期退变智能诊断方法研究<sup>\*</sup>

汤阳杰 吴晓锋

欧阳林

(闽南师范大学数学与统计学院 漳州 363000)

(解放军第九〇九医院医学影像科 漳州 363000)

刘 群

罗爱芳

(闽南师范大学数学与统计学院 漳州 363000)

(解放军第九〇九医院医学影像科 漳州 363000)

〔摘要〕 分析腰椎间盘突出早期退变诊断现状及问题,提出基于代谢组学和标签分布学习方法可建立腰椎间盘突出早期退变分类器,并利用 PT-Bayes、AA-BP 和 SA-IIS 算法实施构建,通过实际计算分析各分类器的性能、应用意义及效果。

〔关键词〕 腰椎间盘突出早期退变;代谢组学;标签分布学习;分类器;智能诊断

〔中图分类号〕 R-058 〔文献标识码〕 A 〔DOI〕 10.3969/j.issn.1673-6036.2023.02.007

**Study on Intelligent Diagnostic Methods of Early Lumbar Disc Degeneration** TANG Yangjie, WU Xiaofeng, School of Mathematics and Statistics, Minnan Normal University, Zhangzhou 363000, China; OUYANG Lin, Department of Medical Imaging, PLA 909th Hospital, Zhangzhou 363000, China; LIU Qun, School of Mathematics and Statistics, Minnan Normal University, Zhangzhou 363000, China; LUO Aifang, Department of Medical Imaging, PLA 909th Hospital, Zhangzhou 363000, China

〔Abstract〕 The present situation and problems of diagnosis of early lumbar disc degeneration are analyzed. The classifiers for early lumbar disc degeneration (LDD) are proposed based on metabonomics and label distribution learning. The LDD classifiers based on PT-Bayes, AA-BP and SA-IIS algorithms are established, and the performance, application significance and effect of each classifier are analyzed by practical calculation.

〔Keywords〕 early lumbar disc degeneration; metabonomics; label distribution learning; classifier; intelligent diagnosis

## 1 引言

相关研究表明,腰椎间盘突出退变(lumbar disc degeneration, LDD)是造成腰背痛的重要原因<sup>[1]</sup>。据估计,全球约 84% 的人都曾经历腰背痛,部分患者会逐步转化为慢性残疾,给家庭和社会造成严重经济负担<sup>[2]</sup>。为了有效预防和治疗 LDD,应该准确诊断腰椎间盘的退变程度。

目前,国际普遍公认的 LDD 诊断标准仍是 Pfirrmann 分级系统,由 Pfirrmann C W A 等<sup>[3]</sup>在 2001

〔修回日期〕 2022-09-06

〔作者简介〕 汤阳杰,硕士研究生,发表论文 1 篇;通信作者:吴晓锋,教授,博士生导师,发表论文 136 篇。

〔基金项目〕 福建省区域发展项目“基于多模态功能影像的腰椎间盘突出退变代谢组学研究”(项目编号:2019Y3007)。

年提出。该分级系统通过磁共振成像 (magnetic resonance imaging, MRI) 技术得到椎间盘矢状位 T2 加权像 (T2 weighted image, T2WI), 将椎间盘退变由轻到重分为 I 至 V 5 个等级。影像医生根据 Pfirrmann 分级标准解读椎间盘的 MRI 图像, 然后对 LDD 等级进行主观判断。值得关注的是, 在对腰椎间盘的早期退变 (I 级和 II 级) 进行分级时, 不同影像医生往往会有较大分歧<sup>[3-4]</sup>。原因包括以下 3 方面。一是 Pfirrmann 分级系统描述 I 级和 II 级的指标存在不确定性, 例如, I 级和 II 级主要通过髓核结构是否同质光亮来划分<sup>[3]</sup>, 这种定性指标会导致判断的歧义。二是 Pfirrmann 分级系统只对 I 级和 II 级的腰椎间盘形态特征进行描述, 不能描述介于 I 级和 II 级之间的退变形态变化<sup>[4]</sup>。三是医生受自身经验等因素影响, 不可避免存在主观诊断误差。

已有研究表明<sup>[5]</sup>, 早期阶段的 LDD 可以通过基因、细胞和生长因子等干预疗法来修复和抑制, 而进入中后期的 LDD 将为不可逆退变。因此, 早期 LDD 是预防和治疗椎间盘退变的最佳时期。为了帮助医生在诊断早期 LDD 时消除歧义、纠正偏差, 有必要研究一种基于人工智能的 LDD 分类器<sup>[6]</sup>。

早期 LDD 伴随着椎间盘生化代谢物 (如水、蛋白多糖、糖胺多糖、椎体骨髓脂肪分数等) 的变化<sup>[7-10]</sup>。这些代谢物在 LDD 中是连续变化的, 并且可定量测量。因此, 利用这些代谢物指标去研究早期 LDD 可以避免采用图像形态特征描述退变而产生的不确定性问题。有学者<sup>[8-10]</sup>研究了 T1 $\rho$ 、T2\*、椎体脂肪分数 (fat fraction, FF) 等代谢指标与 Pfirrmann 分级之间的相关问题。本文将进一步考虑这些代谢指标与 Pfirrmann 分级之间关系的数学模型, 即应用机器学习方法建立关于 I 级和 II 级的 LDD 分类器。

采用在 MRI 上检测到的每个腰椎间盘的 T2\* 值和上、下位椎体 FF 值等代谢指标作为样本示例, 借助主观概率的区间对分法辅助医生生成每个腰椎间盘样本的标签分布概率, 尽量保留医生对腰椎间盘样本在 I 级和 II 级上的判断信息。通过引入标签分布学习<sup>[11]</sup>算法来处理早期 LDD (I 级和 II 级) 诊断的不确定性, 建立 PT - Bayes、AA - BP 和 SA - IIS 3 种

LDD 分类器。这些分类器可以对尚未确诊的腰椎间盘突出判断发生 I 级和 II 级退变的概率, 为医生诊断提供辅助。本文还在测试样本集上进一步分析 3 种分类器的性能, 并讨论所建立的 LDD 分类器如何辅助医生进行早期 LDD 诊断, 有助于精确诊断早期 LDD。

## 2 资料与方法

### 2.1 数据收集与处理

2.1.1 样本示例收集 2019 年 3—11 月, 经医院伦理委员会批准, 在解放军第九〇九医院医学影像科招募 110 名志愿者 (19 ~ 78 岁) 进行常规 MRI (选取的区域 Area  $\leq 200 \text{ mm}^2$ , 所得数值为均值 mean) 及 mDIXON - Quant (区域 Area 为椎体范围, 所得数值为均值 mean) 检查。由一名经验丰富的影像医生对每名志愿者 L1/L2 ~ L5/S1 (Contour 1 - 5) 等 5 节腰椎间盘的 T2\* 值和上、下位椎体 FF 值进行测量, 筛选得到满足样本采集条件的 78 名志愿者共计 390 个腰椎间盘的数据。每个腰椎间盘的 T2\* 值和上、下位椎体 FF 值构成 1 个样本示例。令向量  $x_i = (x_{i1}, x_{i2}, x_{i3})$  表示第  $i$  个样本示例, 其中  $x_{i1}$ 、 $x_{i2}$  和  $x_{i3}$  分别表示第  $i$  个椎间盘的 T2\*、FFu 和 FFd 值。

2.1.2 样本标签收集 根据影像医生的经验, 对腰椎间盘 Pfirrmann 分级的临床判断主要在相邻两个分级之间存在不确定性。采用决策论中量化主观概率的区间对分法<sup>[12]</sup>来辅助医生对每个腰椎间盘样本的标签分布进行赋值。第 1 步, 影像医生根据腰椎间盘样本的常规 T2WI 图像, 借助 Pfirrmann 分级标准进行判读; 第 2 步, 先确定最大可能的 Pfirrmann 分级  $c_1 \in \{I, II, III, IV, V\}$ , 然后确定另一个次可能的相邻退变等级  $c_2 \in \{I, II, III, IV, V\}$ ; 第 3 步, 对可能性较大的等级  $c_1$  做出如下判断: 等可能 (概率赋值为 0.5)、更可能 (概率赋值为 0.625)、非常可能 (概率赋值为 0.875)、完全可能 (概率赋值为 1); 第 4 步, 用 1 减去可能性较大等级的概率即可得出另外一个等级  $c_2$  的概率。由两名影像医生按照上述方法共同对 390 个腰椎间盘样本逐个赋予标签概率分布, 由此得到样本在每对相邻退变等

级的分布情况，见表 1。

表 1 样本在每对相邻退变等级的分布

| 退变等级     | L1/L2 | L2/L3 | L3/L4 | L4/L5 | L5/S1 | 合计  |
|----------|-------|-------|-------|-------|-------|-----|
| I - II   | 63    | 55    | 48    | 34    | 35    | 235 |
| II - III | 5     | 7     | 5     | 7     | 11    | 35  |
| III - IV | 7     | 8     | 8     | 15    | 11    | 49  |
| IV - V   | 3     | 8     | 17    | 22    | 21    | 71  |
| 合计       | 78    | 78    | 78    | 78    | 78    | 390 |

由表 1 可以看出，在 390 个样本中，属于 I - II 级退变的样本有 235 个。令  $d_i = (d_i^I, d_i^{II})$  表示第  $i$  个腰椎间盘的标签分布，其中  $d_i^I$  和  $d_i^{II}$  分别表示第  $i$  个腰椎间盘退变等级为 I 级和 II 级的概率。由此可得到早期腰椎间盘标签分布样本集  $D = \{(x_1, d_1), (x_2, d_2), \dots, (x_{235}, d_{235})\}$ 。

2.1.3 训练集与测试集划分 本文基于矢量均值相似性的样本分割法<sup>[13]</sup>，提出了一种标签分布样本集的分解方法，可以使训练集和测试集在最大概率退变等级为 I 级或 II 级的样本比例大致相同，方法步骤如下。第 1 步，将采集到的标签分布样本集  $D = \{(x_1, d_1), (x_2, d_2), \dots, (x_{235}, d_{235})\}$  转化为单标签样本集  $D' = \{(x_1, y_1), (x_2, y_2), \dots, (x_{235}, y_{235})\}$ ，其中  $y_i$  表示第  $i$  个样本的标签分布中概率最大的退变等级；第 2 步，求得样本集  $D'$  中样本示例  $x_i = (x_{i1}, x_{i2}, x_{i3})$  的均值向量  $\mu = (\mu_1, \mu_2, \mu_3)$ ，并计算每个样本示例  $x_i$  与均值向量  $\mu$  之间的余弦相似度  $\cos\theta_i = \frac{\mu^T x_i}{\sqrt{\mu^T \mu} \sqrt{x_i^T x_i}}$ ，当  $\cos\theta_i$  越接近于 1，说明样本示例  $x_i$  与  $\mu$  越相似；第 3 步，按照样本退变等级先 I 级、后 II 级的原则将  $D'$  中的样本进行排序，其后再按照样本余弦相似度从小到大对 I 级和 II 级的样本进行排序；第 4 步，将  $D'$  中排序好的所有样本分成 78 段，前 77 段中每段包含 3 个样本，第 78 段包含 4 个样本，抽取每段的第 2 个样本组成单标签测试集  $T'$ ，剩下的样本组成单标签训练集  $S' = D' - T'$ ；第 5 步，通过对照每个样本示例，把单标签的训练集  $S'$  和测试集  $T'$  还原为由 157 个样本（约占总样本数的 2/3）组成的标签分布训练集  $S$  和 78

个样本组成的标签分布测试集  $T$ 。

2.1.4 样本示例归一化 为了消除样本示例  $x_i = (x_{i1}, x_{i2}, x_{i3})$  中不同代谢指标的量纲和值域差异，通过如下归一化公式分别将训练集  $S$  和测试集  $T$  中的每个样本示例的 3 个代谢指标变换为在 0 与 1 之间取值：

$$z_{ij} = \frac{(x_{ij} - \min_i x_{ij})}{(\max_i x_{ij} - \min_i x_{ij})}, x_i \in S \text{ 或 } T, j = 1, 2, 3 \quad (1)$$

其中， $x_{ij}$  表示第  $i$  个样本  $x_i$  中的第  $j$  个属性值， $\max_i x_{ij}$  和  $\min_i x_{ij}$  分别表示训练集  $S$  或测试集  $T$  中第  $j$  个属性的最大值和最小值。

2.2 LDD 分类器

2.2.1 概述 在标签分布学习中，每个样本拥有一个多标签集上的概率分布，而标签分布分类器对样本的标签预测也表示为多标签集上的概率分布。这个概率分布反映了样本属于每个标签的合理性程度，因此可用来描述在多标签上的不确定性。本文选择 PT - Bayes 分类器等 3 种基于不同策略的标签分布学习器。

2.2.2 PT - Bayes 分类器 选择问题转换（problem transformation, PT）策略中的 PT - Bayes 分类器<sup>[11]</sup>，通过样本随机重采样方法把标签分布训练集  $S$  转化成单标签训练集  $S_M$ ，再利用集合  $S_M$  训练朴素 Bayes 分类器，得到分类函数为：

$$P(c|x_p) = \frac{P(x_p|c)P(c)}{P(x_p)} = \frac{|S_M^c| \prod_{l=1}^3 P(x_{pl}|c)}{\sum_{k=I,II} |S_M^k| \prod_{l=1}^3 P(x_{pl}|c)} \quad (2)$$

其中， $c \in \{I, II\}$ ，腰椎间盘突出代谢指标  $l = 1, 2, 3$ ， $x_p$  为待预测样本示例， $S_M^c$  为训练集  $S_M$  中退变等级为  $c$  的样本子集，条件概率  $P(x_{pl}|c)$  满足均值  $\mu_{cl} = \left(\frac{1}{|S_M^c|}\right) \sum_{x_l \in S_M^c} x_{l1}$ 、方差  $\sigma_{cl} = \sqrt{\left(\frac{1}{|S_M^c|}\right) \sum_{x_l \in S_M^c} (x_{l1} - \mu_{cl})^2}$  的正态分布， $S_M^c$  为训练子集  $S_M$  中各样本的属性  $l$  值组成的集合。

2.2.3 AA - BP 分类器 选择适应性修改（algorithm adaptation, AA）策略中的 AA - BP 分类器<sup>[11]</sup>，是一个由输入层、隐层和输出层组成的 3 层

前馈神经网络分类器, 见图 1。分类函数为:

$$d_p^c = \frac{\exp(\sum_{h=1}^7 v_{hc} (1 + \exp(\gamma_h - \sum_{l=1}^3 \omega_{lh} x_{pl}))^{-1} - \theta_c)}{\sum_{j=I, II} \exp(\sum_{h=1}^7 v_{hj} (1 + \exp(\gamma_h - \sum_{l=1}^3 \omega_{lh} x_{pl}))^{-1} - \theta_j)} \quad (3)$$

其中,  $c \in \{I, II\}$ ,  $x_{pl}$  为待预测样本示例  $x_p$  的第  $l$  个代谢指标值,  $l = 1, 2, 3$ ;  $\omega = (\omega_{lh}) \in R^{21}$  为输入层与隐层之间的连接权重组成的向量,  $\gamma = (\gamma_h) \in R^7$  为 7 个隐层神经元的阈值组成的向量;  $v = (v_{hj}) \in R^{14}$  为隐层与输出层之间的连接权重组成的向量,  $\theta = (\theta_j) \in R^2$  为 2 个输出层神经元的阈值组成的向量。隐层神经元的激活函数选用了 sigmoid 函数<sup>[13]</sup>, 输出层神经元的激活函数选用了 softmax 函数<sup>[11]</sup>。AA-BP 的优化目标是使训练集  $S$  中所有样本的真实标签分布与预测标签分布之间的累积方差为最小。由梯度下降法求解最小方差可优化得到神经元连接权重及其阈值。

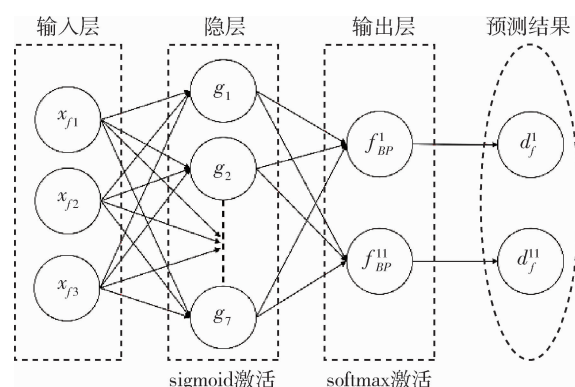


图 1 AA-BP 神经网络结构

2.2.4 SA-IIS 分类器 选择专门化设计 (specialized algorithms, SA) 策略中的 SA-IIS 分类器<sup>[11]</sup>, 是利用训练集  $S$  训练得到一个分类函数  $p(c|x;\theta)$ , 使得训练集  $S$  中所有样本的真实标签分布与预测标签分布之间的累积库尔贝克-莱布勒散度 (kullback-leibler, KL)<sup>[11]</sup> 为最小。考虑分类函数  $p(c|x;\theta)$  为如下熵函数:

$$p(c|x;\theta) = \frac{1}{Z} \exp(\sum_{l=1}^3 \theta_{cl} g_l(x)) \quad (4)$$

其中  $Z = \sum_{k=I}^{II} \exp(\sum_{l=1}^3 \theta_{kl} g_l(x))$  为归一化因

子,  $\theta = (\theta_{k1}, \theta_{k2}, \theta_{k3})$  为待训练 (优化) 的参数向量,  $g_l(x)$  为样本示例  $x$  在第  $l = 1, 2, 3$  个属性下的取值,  $c \in \{I, II\}$ 。由高斯牛顿法求解最小累积 KL 散度来优化得到参数向量  $\theta$ 。

## 2.3 评估指标

为了评估上述 LDD 标签分布分类器的预测性能, 需要选择可度量所预测的测试样本标签分布与其实际标签分布之间差距的指标作为 LDD 分类器的评估指标。Cha S H<sup>[14]</sup> 提出了 41 个可度量不同概率分布之间距离或相似性的指标。本文从中选出了 4 个指标作为评估 LDD 分类器的指标, 公式如下。这些规则是: 每个指标之间的相关性较小; 每个指标不应出现被零除或取零对数等不可计算的情况; 所选指标较为常用。Chebyshev 和 Squaredchored 为距离度量指标, Cosine 和 Intersection 为相似性度量指标。↓ 表示数值越小性能越好, ↑ 表示数值越大性能越好。 $d_i^c$  表示第  $i$  个测试样本的退变等级为  $c$  的实际概率,  $\hat{d}_i^c$  表示由分类器预测的第  $i$  个测试样本退变等级为  $c$  的概率, 其中  $c \in \{I, II\}$ 。

$$\text{Chebyshev } \downarrow = \sum_{i=1}^{|T|} \max_{c \in Y} |d_i^c - \hat{d}_i^c| \quad (5)$$

$$\text{Squaredchored } \downarrow = \sum_{i=1}^{|T|} \sum_{c \in Y} (\sqrt{d_i^c} - \sqrt{\hat{d}_i^c})^2 \quad (6)$$

$$\text{Cosine } \uparrow = \sum_{i=1}^{|T|} \frac{\sum_{c \in Y} d_i^c \hat{d}_i^c}{\sqrt{\sum_{c \in Y} (d_i^c)^2} \sqrt{\sum_{c \in Y} (\hat{d}_i^c)^2}} \quad (7)$$

$$\text{Intersection } \uparrow = \sum_{i=1}^{|T|} \sum_{c \in Y} \min(d_i^c, \hat{d}_i^c) \quad (8)$$

## 3 结果与分析

### 3.1 LDD 分类器指标对比

利用 157 个训练样本分别训练 PT-Bayes、AA-BP 和 SA-IIS 分类器, 得到 3 个 LDD 分类器。再将 78 个测试示例输入到训练好的分类器中, 可以得到各分类器对每个测试示例的预测标签分布。由评估指标公式分别求出 3 个 LDD 分类器的 Chebyshev、Squaredchored、Cosine 和 Intersection 指标, 见表 2。

表 2 3 个 LDD 分类器的指标对比

| 指标              | PT - Bayes | AA - BP  | SA - IIS |
|-----------------|------------|----------|----------|
| Chebyshev ↓     | 0. 290 2   | 0. 288 9 | 0. 287 2 |
| Squaredchored ↓ | 0. 174 6   | 0. 165 8 | 0. 157 6 |
| Cosine ↑        | 0. 850 0   | 0. 858 2 | 0. 859 8 |
| Intersection ↑  | 0. 709 8   | 0. 711 1 | 0. 712 8 |

由表 2 的计算结果可以看出，PT - Bayes、AA - BP 和 SA - IIS 3 个分类器在相似性指标 Cosine 上都达到 0. 85 以上，在距离指标 Squaredchored 上都达到 0. 18 以下，说明 3 个分类器对早期 LDD（I 级和 II 级）都能达到一定的预测精度。

3. 2 LDD 分类器应用价值

3.2.1 对腰椎间盘突出早期退变进行准确预测 LDD 分类器对腰椎间盘突出早期退变（I 级和 II 级）进行准确的预测，可以辅助医生解决 LDD 的诊断难题。如果医生对某一腰椎间盘突出早期退变的判断结果与分类器的预测结果大致相同（如在最大概率的等级上判断一致），可进一步确信自己的判断。如果医生判断结果与分类器的预测结果出现较大分歧（如在最大概率等级上判断不一致），分类器的预测结果可以提示医生是否存在误判的可能，再通过一些可解释性说明帮助医生找到判断偏差的原因。

3.2.2 为早期 LDD 诊断提供新的思路 基于代谢组学的 LDD 分类为早期 LDD 诊断提供了一种新的思路。由于代谢指标可以感知到腰椎间盘突出从 I 级向 II 级的连续退变过程，可以根据代谢指标变化（而不是形态学改变）来描述腰椎间盘的早期退变，与早期退变的修复和治疗手段相结合，更细致地定义腰椎间盘突出早期退变的等级。随着医院信息化的不断发展，采集和储存更多人群的腰椎间盘突出代谢组学数据已成为可能。利用人工智能和大数据方法充分挖掘这些组学数据的潜在规律，将为腰椎间盘的临床精确诊断提供更多帮助。

4 结语

本文基于代谢组学和标签分布学习研究了适用于早期 LDD（I 级和 II 级）的分类器，采用主观概

率的区间对分法将医生对早期 LDD 的定性判断转化成定量的概率标签分布，给出一种基于矢量均值相似性的标签分布样本集分解方法，建立基于 PT - Bayes、AA - BP 和 SA - IIS 算法的 LDD 分类器。实际计算表明，所建立的 LDD 分类器对早期退变可达到一定的预测精度。

参考文献

1 茅峰, 刘建文, 陈静云, 等. 椎间盘退变中自噬与凋亡的分子关系研究进展 [J]. 中国中医骨伤科杂志, 2021, 29 (9): 81 - 84.

2 ZHU J, ZHANG X, GAO W, et al. lncRNA /circRNA - miRNA - mRNA ceRNA network in lumbar intervertebral disc degeneration [J]. Molecular medicine reports, 2019, 20 (4): 3160 - 3174.

3 PFIRRMANN C W A, METZDORF A, ZANETTI M, et al. Magnetic resonance classification of lumbar intervertebral disc degeneration [J]. Spine, 2001, 26 (17): 1873 - 1878.

4 GRIFFITH J F, WANG Y, ANTONIO G, et al. Modified Pfirrmann grading system for lumbar inter - vertebral disc degeneration [J]. Spine, 2007, 32 (24): e708 - e712.

5 余城墙, 张宇, 谢程欣, 等. 椎间盘退变分子生物学机制及再生治疗的优势与未来 [J]. 中国组织工程研究, 2019, 23 (30): 4889 - 4896.

6 姚泽阳, 谢稳, 邱海龙, 等. 人工智能在临床医学中的应用与展望 [J]. 医学信息学杂志, 2020, 41 (3): 39 - 43.

7 苏树燕, 刘源. 椎间盘退变的磁共振功能成像研究进展 [J]. 汕头大学医学院学报, 2019, 32 (1): 49 - 52.

8 李丹, 王莉霞, 吴文骏, 等. MRI T1rho 对下腰痛者腰椎间盘退变的定量评估 [J]. 磁共振成像, 2019, 10 (12): 904 - 907.

9 曾菲菲, 查云飞, 邢栋, 等. 磁共振扩散峰度成像和 T2 \* - mapping 技术定量检测腰椎间盘突出退变的对比研究 [J]. 放射学实践, 2018, 33 (10): 1087 - 1092.

10 李勉文, 崔运能, 赵银霞, 等. 腰椎骨髓脂肪沉积与椎间盘退变 Pfirrmann 分级的相关性研究 [J]. 临床放射学, 2017, 36 (12): 1838 - 1842.

11 GENG X. Label distribution learning [J]. IEEE transactions on knowledge and data enginee - ring, 2016, 28 (7): 1734 - 1748.

12 岳超源. 决策理论与方法 [M]. 北京: 科学出版社, 2004.

13 江丽红, 吴晓锋, 欧阳林, 等. 基于代谢组学腰椎间盘突出变的计算机辅助诊断 [J]. 中国组织工程研究, 2021, 25 (24): 3796 - 3803.

14 CHA S H. Comprehensive survey on distance/similarity measures between probability density functions [J]. International journal of mathe - matical models and methods in applied sciences, 2007, 1 (4): 300 - 307.