

## 基于代谢组学腰椎间盘突出变的计算机辅助诊断

<https://doi.org/10.12307/2021.082>江丽红<sup>1</sup>, 吴晓锋<sup>1</sup>, 欧阳林<sup>2,3</sup>, 罗爱芳<sup>2,3</sup>, 黄丽<sup>2,3</sup>

投稿日期: 2020-07-27

送审日期: 2020-07-29

采用日期: 2020-08-25

在线日期: 2021-02-10

中图分类号:

R459.9; R604; R319

文章编号:

2095-4344(2021)24-03796-08

文献标识码: A

## 文章快速阅读:

## 目的:

基于 MRI 获得的腰椎间盘突出数据建立自动识别腰椎间盘突出等级的计算机辅助诊断方法。

## 相关性分析:

采用 Spearman 相关分析检验 MRI 代谢指标与腰椎间盘突出退变的 Pfirrmann 等级显著相关。

## 划分训练集与测试集:

在 390 个样本上应用基于矢量均值相似性的样本分割法 (MSSS) 划分训练集和测试集。

## 结论:

采用 MRI 代谢指标建立腰椎间盘突出智能辅助诊断具有可行性, 为腰椎间盘突出退变诊断提供了新思路。

## 模型评价:

在测试集上比较 3 种诊断模型的性能。

## 在训练集上训练模型:

用 Softmax 回归、神经网络、支持向量机 3 种机器学习算法建立腰椎间盘突出退变诊断模型。

## 文题释义:

**MRI:** 是利用原子核在强磁场内发生共振产生的信号绘制出人体内部立体图像的技术, 可对人体某个部位进行各种角度和各种平面成像, 临床医生可根据腰椎间盘的磁共振成像评价腰椎间盘突出等级。MRI 定量技术可定量检测腰椎间盘的多种生化代谢物, 针对目前临床对腰椎间盘突出等级诊断有一定困难, 文章尝试基于 MRI 检测的生化代谢指标对腰椎间盘突出退变进行诊断。

**机器学习:** 是一种实现人工智能的方法, 包括逻辑回归、支持向量机、神经网络、决策树、朴素贝叶斯等多种算法, 可在大量样本中利用算法发现数据规律并训练分类器, 在得到新的样本时利用训练好的分类器对样本进行分类和预测。实验利用收集的腰椎间盘突出 MRI 代谢数据训练了 3 种机器学习分类器, 为腰椎间盘突出等级诊断提供了新方法。

## 摘要

**背景:** 腰椎间盘突出退变诊断对预防腰椎疾病意义重大, 但目前对其诊断主要依赖于影像医师的主观评价, 易因个人经验不足产生误判。

**目的:** 建立自动识别腰椎间盘突出等级的计算机辅助诊断方法, 为影像医师提供参考。

**方法:** 采用 Spearman 相关分析验证腰椎间盘的 MRI 代谢指标与腰椎间盘突出退变的 Pfirrmann 等级相关性, 并建立可用于腰椎间盘突出退变智能诊断的 Softmax 回归、神经网络和支持向量机等多种分类器。

**结果与结论:** 相关性分析结果表明, 椎间盘相邻上下位椎体脂肪分数 FF 值和 T2\* 值等 3 种生化代谢指标都与腰椎间盘突出退变显著相关, Softmax 回归、神经网络和支持向量机 3 种诊断模型的综合准确率分别为 0.477、0.515 和 0.523, kappa 系数分别为 0.311、0.300 和 0.330。实际分析结果表明, 采用 MRI 代谢指标建立腰椎间盘突出退变智能辅助诊断是可行的, 为腰椎间盘突出退变诊断提供了一种可行的途径。

**关键词:** 骨; 腰椎间盘突出退变; MRI; Pfirrmann 等级; Spearman 相关分析; softmax 回归; 神经网络; 支持向量机; 代谢指标

## Computer aided diagnosis of lumbar disc degeneration based on metabolomics

Jiang Lihong<sup>1</sup>, Wu Xiaofeng<sup>1</sup>, Ouyang Lin<sup>2,3</sup>, Luo Aifang<sup>2,3</sup>, Huang Li<sup>2,3</sup>

<sup>1</sup>School of Mathematics and Statistics, Minnan Normal University, Zhangzhou 363000, Fujian Province, China; <sup>2</sup>Institute of Medical Imaging, School of Medicine, Xiamen University, Zhangzhou 363000, Fujian Province, China; <sup>3</sup>Department of Medical imaging, PLA 909<sup>th</sup> Hospital, Zhangzhou 363000, Fujian Province, China

**Corresponding author:** Wu Xiaofeng, Professor, Doctoral supervisor, School of Mathematics and Statistics, Minnan Normal University, Zhangzhou 363000, Fujian Province, China

**Co-corresponding author:** Ouyang Lin, MD, Chief physician, Institute of Medical Imaging, School of Medicine, Xiamen University, Zhangzhou 363000, Fujian Province, China; Department of Medical imaging, PLA 909<sup>th</sup> Hospital, Zhangzhou 363000, Fujian Province, China

<sup>1</sup> 闽南师范大学数学与统计学院, 福建省漳州市 363000; <sup>2</sup> 厦门大学医学院医学影像研究所, 福建省漳州市 363000; <sup>3</sup> 解放军第九〇九医院医学影像科, 福建省漳州市 363000

**第一作者:** 江丽红, 女, 1996 年生, 广东省广州市人, 硕士, 主要从事机器学习与健康医疗大数据研究。

**通讯作者:** 吴晓锋, 教授, 博士生导师, 闽南师范大学数学与统计学院, 福建省漳州市 363000

**通讯作者:** 欧阳林, 博士, 主任医师, 厦门大学医学院医学影像研究所, 福建省漳州市 363000; 解放军第九〇九医院医学影像科, 福建省漳州市 363000

<https://orcid.org/0000-0002-7568-9741> (江丽红)

**基金资助:** 福建省科技计划项目 (2019Y31010067), 项目负责人: 欧阳林; 第九〇九医院青年苗圃基金 (18Y021), 课题负责人: 罗爱芳

**引用本文:** 江丽红, 吴晓锋, 欧阳林, 罗爱芳, 黄丽. 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.



## Abstract

**BACKGROUND:** The diagnosis of lumbar disc degeneration is of great significance for the prevention of lumbar disease, and the diagnosis of lumbar disc degeneration mainly relies on the subjective evaluation of the imaging physician, which is likely to misjudge because of insufficient experience.

**OBJECTIVE:** To propose a computer-aided diagnosis technique for classification on the lumbar disc degeneration, and to provide reference for imaging doctors.

**METHODS:** Spearman correlation analysis is used to verify the correlation between magnetic resonance imaging metabolic indices of lumbar intervertebral disc and the Pfirrmann grades of lumbar disc degeneration. Several classifiers for the intelligent diagnosis of lumbar disc degeneration are developed by means of machine learning strategies such as the Softmax regression, the neural network and the support vector machine.

**RESULTS AND CONCLUSION:** The result of correlation analysis showed that three metabolic indices such as fat fraction (FF) of adjacent upper and lower vertebral bodies of degenerative disc, T2\* values were significantly correlated with lumbar disc degeneration. The classification accuracy of the softmax regression, the neural network and the support vector machine respectively was 0.477, 0.515 and 0.523, and kappa's coefficient of these three diagnostic models was 0.311, 0.300 and 0.330, respectively. The actual analysis indicates that it is feasible to establish a computer-aided intelligent diagnosis of lumbar disc degeneration by using the MRI metabolic indices, showing a promising approach for the diagnosis of lumbar disc degeneration.

**Key words:** bone; lumbar disc degeneration; MRI; Pfirrmann level; Spearman correlation analysis; softmax regression; neural network; support vector machine; metabolic indices

**Funding:** the Science and Technology Project of Fujian Province, No. 2019Y31010067 (to OL); the Fund of the 909<sup>th</sup> Hospital Youth Nursery, No. 18Y021 (to LAF)

**How to cite this article:** JIANG LH, WU XF, OUYANG L, LUO AF, HUANG L. Computer aided diagnosis of lumbar disc degeneration based on metabolomics.

Zhongguo Zuzhi Gongcheng Yanjiu. 2021;25(24):3796-3803.

## 0 引言 Introduction

腰背痛是影响社会公共健康的一个重要问题，而腰椎间盘突出退变是腰背痛的主要原因之一<sup>[1-2]</sup>。临床研究表明，人体腰椎间盘突出退变可能从16-20岁开始，并且是一个随增龄而加重的不可逆过程<sup>[3]</sup>。在传统影像诊断中，影像医师根据椎间盘形态、信号改变等评价腰椎间盘突出退变，但此时腰椎间盘突出退变通常已发展到中晚期，临床疗效较差，难以有效阻止腰椎间盘突出退变进展<sup>[4]</sup>。如能够实现腰椎间盘突出退变诊断将有利于指导临床及时干预和治疗，这是预防腰椎间盘突出发生更严重退变的关键步骤，也是当前亟待解决的难题。另一方面，医师根据影像对腰椎间盘突出退变等级进行评价时结果可能因经验不足产生差异，具有一定的主观性，故亟需建立一个自动检测腰椎间盘突出退变等级的计算机辅助诊断系统，为客观评价腰椎间盘突出退变提供依据。

Pfirrmann标准将腰椎间盘突出退变分为I-V5个逐级加重的退变等级，已被公认为判断腰椎间盘突出退变程度的金标准<sup>[5]</sup>。在该标准中，医生可根据椎间盘的形态学变化、MRI信号降低程度等指标判断受检者的椎间盘退变等级。而腰椎间盘突出退变会伴随生化结构的改变<sup>[6]</sup>。过去的研究都是基于影像形状特征或生物力学指标等对腰椎间盘突出进行分类，很少有文献研究基于代谢组学的腰椎间盘突出退变等级分类。实验基于MRI检测的椎间盘代谢组学数据，应用一定的机器学习方法建立分类器自动识别腰椎间盘突出退变等级，为腰椎间盘突出退变提供辅助诊断。

目前，机器学习方法中的回归算法<sup>[7-8]</sup>、支持向量机算法<sup>[9-11]</sup>、神经网络算法<sup>[12-14]</sup>、集成算法等已被广泛应用于各种疾病诊断或代谢组学分析<sup>[15-16]</sup>。这些研究表明引入机器学习有助于分析代谢物与疾病间的关系，挖掘代谢数据中隐含的信息。在基于机器学习的人体腰椎间盘突出退变诊断领域，BOUNDS等<sup>[17]</sup>提出通过训练多层感知机来诊断4种类型的腰痛；KARABULUT等<sup>[18]</sup>使用生物力学指标和过采样技术(SMOTE)构建逻辑模型树(LMT)来识别脊柱滑脱和椎间盘突出；CASTRO-MATEOS等<sup>[19]</sup>从MRI图像中提取椎间盘形状和强度特征，基于Adaboost算法自动识别腰椎间盘突出退变4个

Pfirrmann等级；HE等<sup>[20]</sup>基于465个腰椎间盘的梯度直方图特征，运用有监督距离度量学习方法建立了4个等级腰椎间盘突出退变的自动识别算法；EBRAHIMZADEH等<sup>[21]</sup>用强度和形状特征比较了多层感知机、k-最近邻和支持向量机等机器学习方法对腰椎间盘突出进行“诊断”的能力；ALOMARI等<sup>[22]</sup>仅利用形状特征提出了基于贝叶斯分类器自动“诊断”腰椎间盘突出退变的方法。

一些文献是针对腰椎间盘的疾病诊断，而非腰椎间盘突出退变等级的诊断<sup>[17-18, 21-22]</sup>。还有一些文献研究了腰椎间盘突出退变等级的诊断，但用的是椎间盘图像特征而非代谢组学数据<sup>[19-20]</sup>。已有研究表明，腰椎间盘中的水、乳酸、糖等多种生化代谢物与腰椎间盘突出退变存在相关性<sup>[23-25]</sup>。因此，从腰椎间盘的代谢组学(营养源)层面研究其退变机制，应用机器学习建立腰椎间盘突出退变等级分类器是值得深入探索的科学问题，目前仍鲜见相关成果。实验基于MRI成像技术测量得到的人群腰椎间盘生化代谢物数据，采用Softmax回归、BP神经网络和支持向量机等机器学习方法建立3种腰椎间盘突出退变等级分类器，并通过所获得的MRI椎间盘测试样本比较了3种分类器的性能。

## 1 材料和方法 Materials and methods

### 1.1 设计 前瞻性实验。

1.2 时间及地点 实验于2019年3至11月在解放军第九〇九医院医学影像科完成。

### 1.3 实验方法

#### 1.3.1 数据预处理

(1)数据集：实验数据由解放军第九〇九医院和厦门大学东南医院医学影像科(厦门大学医学院医学影像研究所)提供。该影像科于2019年3至11月招募年龄19-78岁的志愿者110例。排除标准：①体内有巨大金属移植物的患者；②有幽闭恐惧症的人群；③腰椎发育异常、压缩骨折、脊柱畸形、脊柱肿瘤等影响MRI观察的患者；④有严重动脉硬化、糖尿病的患者；⑤有免疫性疾病(如强直性脊柱炎、骨性关节炎)的患者。

行腰椎常规 MRI 及 mDIXON-Quant 检查, 剔除运动伪影较大的 12 例, 纳入标准且图像质量满意者共 98 例。其中, 退变椎间盘相邻上、下位椎体的 MRI 脂肪分数 (FF) 和 T2\* 值皆得到检测的共 78 例, 每例包含 L<sub>1</sub>-S<sub>1</sub> 5 节腰椎间盘的检测数据, 因此得到 390 个腰椎间盘的检测值。椎体脂肪分数代表每个椎体脂肪的百分比含量, T2\* 值指组织在磁场不均匀的状态下的横向弛豫时间, 单位为毫秒 (ms), 它与腰椎间盘中的蛋白多糖和水含量等代谢物密切相关<sup>[24]</sup>。影像医师在后处理工作站 (Intellispace Portal workstation, Philips) 分别对椎体脂肪分数值和腰椎间盘 T2\* 值进行测量。L<sub>1</sub>-S<sub>1</sub> 5 个椎体的脂肪分数值测量兴趣区为正中矢状面椎体边缘皮质下区域 (图 1), 为提高数据准确性, 结果取 2 位影像医师测量的平均值。T2\* 值测量层面参照正中矢状面 T2WI 手动绘制, 测量兴趣区尽量包含椎间盘髓核及内层纤维环, 避开终板和血管等 (图 2)。

由 2 位磁共振诊断医师共同对每位志愿者腰椎间盘逐个行 Pfirrmann 分级, 完成后请一位具有丰富影像诊断经验的主任医师进行审核, 得到 L<sub>1</sub>-S<sub>1</sub> 腰椎间盘 Pfirrmann 等级分布, 如表 1 所示。

表 1 | 各椎间盘的 Pfirrmann 等级分布  
Table 1 | Pfirrmann grade distribution of each disc

Pfirrmann 等级	L <sub>1,2</sub>	L <sub>2,3</sub>	L <sub>3,4</sub>	L <sub>4,5</sub>	L <sub>5-S<sub>1</sub></sub>
I	36	38	29	14	14
II	29	19	21	24	31
III	10	14	11	13	18
IV	3	7	14	23	13
V	0	0	3	4	2

建立 MRI 腰椎间盘数据集  $D = \{(x_1, y_1), (x_2, y_2), \dots, (x_{390}, y_{390})\}$ 。对于样本  $i = 1, 2, \dots, 390$ ,  $x_i = (x_{i1}, x_{i2}, x_{i3})$ , 其中  $(x_{i1}, x_{i2}, x_{i3})$  分别表示第  $i$  个椎间盘样本的 T2\* 值和相邻上、下位椎体的 MRI 脂肪分数等生化代谢值;  $y_i$  是影像医师根据 Pfirrmann 分级方法判别的第  $i$  个椎间盘样本的退变等级,  $y_i \in \{1, 2, 3, 4, 5\}$ 。

(2) 训练样本集与测试样本集: 为了从数据集产生腰椎间盘退变等级分类器并对分类器的性能进行评估, 需要对数据集划分出训练样本集和测试样本集, 并使二者的样本类型具有一致性。应用基于矢量均值相似性的样本分割法 (MSSS)<sup>[26]</sup>, 从 390 个样本中取 1/3 作为测试样本集, 其余为训练样本集, 方法如下:

- 将每个样本  $x_i = (x_{i1}, x_{i2}, x_{i3})$  看成一个三维向量, 求出所有样本数据  $\{x_i : i = 1, 2, \dots, 390\}$  的均值向量  $\mu = (\mu_1, \mu_2, \mu_3)$ ;
- 计算每个样本  $x_i$  与均值向量  $\mu$  的夹角余弦  $\cos \theta_i = \mu^T x_i / (\sqrt{\mu^T \mu} \sqrt{x_i^T x_i})$  作为样本数据  $x_i$  与均值向量  $\mu$  的相似度,  $\cos \theta_i$  越接近于 1, 表明样本数据  $x_i$  与均值向量  $\mu$  越相似;
- 把样本数据集  $\{x_i : i = 1, 2, \dots, 390\}$  先按腰椎间盘退变等级排序, 在每个退变等级里再将样本按相似度从小到大进行排序;
- 把排序后的样本数据集分割成 130 段, 每段有 3 个样本, 取中间位置的作为测试样本, 其他样本为训练样本。最终得到 260 个训练样本和 130 个测试样本。

(3) 归一化: 为了加快训练速度, 消除不同生化代谢指标之间量纲的差异, 分别对训练样本集和测试样本集按照如下公式进行归一化处理:

$$x'_{ij} = \frac{x_{ij} - \min_i x_{ij}}{\max_i x_{ij} - \min_i x_{ij}}, \quad i = 1, 2, \dots, t, j = 1, 2, 3 \quad (1)$$

使得每个数据在 [0, 1] 之间取值, 其中,  $t$  为训练集 (或测试集) 样本数。

1.3.2 相关性分析 应用 SPSS 软件对所有受检者的数据  $\{(x_i, y_i) : i = 1, 2, \dots, 390\}$  做 Spearman 相关性分析, 检验腰椎间盘的各生化代谢指标  $x_j$  ( $j = 1, 2, 3$ ) 与 Pfirrmann 等级  $y$  的相关性, 文章使用的生化代谢指标包括椎间盘相邻上、下位椎体的 MRI 脂肪分数和 T2\* 值。如果某一生化代谢指标值  $x_j$  与 Pfirrmann 等级呈单调递增 (或递减) 关系, 则称两者为正 (或负) 相关。为度量第  $j$  个生化代谢指标  $x_j$  与 Pfirrmann 等级的相关性, 对于  $i = 1, 2, \dots, 390$ , 将  $x_{ij}$  和  $y_i$  从小到大排序, 得到 2 个序列, 每对  $x_{ij}$  和  $y_i$  在各自序列的排序位置称为秩, 分别记为  $R(x_{ij})$  和  $R(y_i)$ 。当序列中有相同的值时, 它们的秩取排序位置的平均值。

令  $R(x_j) \in \mathbb{R}^{390}$  和  $R(y) \in \mathbb{R}^{390}$  表示所有受检者数据的秩组成的向量, 令向量  $h_j = R(x_j) - R(y) \in \mathbb{R}^{390}$ 。那么, 第  $j$  个生化代谢指标和 Pfirrmann 分级的 Spearman 相关系数为<sup>[27]</sup>

$$r_j = 1 - \frac{6(h_j)^T (h_j)}{N(N^2 - 1)}, \quad j = 1, 2, 3 \quad (2)$$

其中  $N = 390$  为样本数。又因为序列中相同的秩较多, 因此该文使用校正的 Spearman 相关系数  $r'_j$ <sup>[28]</sup>,

$$r'_j = \frac{R(x_j)^T R(y) - R(x_j)^T e * R(y)^T e}{\sqrt{[R(x_j)]^T [R(x_j)] - (R(x_j)^T e)^2 / N} \sqrt{[R(y)]^T [R(y)] - (R(y)^T e)^2 / N}}, \quad j = 1, 2, 3 \quad (3)$$

其中, 向量  $e \in \mathbb{R}^{N \times 1}$ ,  $r'_j$  在区间 [-1, 1] 取值。当  $r'_j > 0$  ( $< 0$ ) 表明腰椎间盘的第  $j$  个生化代谢指标与腰椎间盘退变的 Pfirrmann 等级呈正 (负) 相关。如果  $|r'_j|$  越接近 1, 则显示第  $j$  个生化代谢指标与腰椎间盘退变的 Pfirrmann 等级相关性越强。若  $|r'_j| > 2.575 / \sqrt{N - 1}$ , 则认为在 1% 的显著性水平下该生化代谢指标与腰椎间盘退变的 Pfirrmann 等级显著相关。

1.3.3 基本原理 在对腰椎间盘退变与疾病的分类研究中, 有许多经典的机器学习方法被尝试使用, 包括朴素贝叶斯、k 近邻、决策树、支持向量机、神经网络、Logistic 回归等<sup>[17-22, 29-30]</sup>。实验尝试使用多种机器学习方法, 希望找到较稳定且分类性能较好的腰椎间盘退变分类器。朴素贝叶斯算法的前提假设是各个特征之间相互独立, 而现实中生化代谢物之间往往相互影响。k 近邻算法需计算测试样本与每个训练样本间的距离, 但参数  $k$  不好确定。决策树算法虽然便于直观判断, 但腰椎间盘退变各个等级间的代谢指标无明显边界, 训练出来的分类边界可能非常复杂。因此未使用上述机器学习方法。

(1) 腰椎间盘退变的计算机辅助诊断方法: 由于所获得的数据集中腰椎间盘退变等级为 V 的样本较少, 将腰椎间盘退变等级为 V 的样本合并到等级 IV 的样本中, 只对 I、II、

III、IV等级的腰椎间盘退变进行分类。先在训练集上训练 Softmax 回归、神经网络和支持向量机模型，最后将这些训练好的多类别分类器用于对测试集分类。机器学习训练过程在 Matlab R2018a 软件 (图 3) 完成。

为了择优选择分类器，下面分别给出 3 种分类器的建模方法。

(2) Softmax 回归: Softmax 回归是 logistic 回归在多分类问题上的扩展，可以通过 Softmax 函数将特征向量映射到一个概率分布，即建立一个模型就可以得出腰椎间盘样本属于每个退变等级的概率，而不需要建立多个 logistic 回归模型才能预测，减少了训练时间。而且对于不同的退变等级其生化代谢指标重要性或多或少会有差异，Softmax 回归可以得到每个类别对应的代谢指标的权重，从而区分不同退变等级。

对于数据集，令  $\mathbf{x} \in R^{3 \times 1}$  是某个腰椎间盘样本的代谢组学数据。在建立 Softmax 回归分类器时，令  $\mathbf{x} = (1, \mathbf{x}^T)^T$ 。那么，Softmax 回归分类器可表示为：

$$f_L(\mathbf{x}) = \frac{1}{\sum_{j=1}^c e^{\theta_j^T \mathbf{x}}} \begin{bmatrix} e^{\theta_1^T \mathbf{x}} \\ \vdots \\ e^{\theta_c^T \mathbf{x}} \end{bmatrix} \quad (4)$$

其中， $c$  为类别数， $\theta_j \in R^{4 \times 1}$  是模型参数， $j=1, 2, \dots, c$ 。  $\sum_{j=1}^c e^{\theta_j^T \mathbf{x}}$  是分类器的归一化参数，使  $f_L(\mathbf{x}) \in R^{c \times 1}$  输出样本  $\mathbf{x}$  属于每个类别的概率，且概率之和等于 1，选择最大概率对应的退变等级作为样本  $\mathbf{x}$  的预测等级。

由极大似然法可得到求解分类器参数的优化模型：

$$\min_{\theta_1, \dots, \theta_c} -\frac{1}{n} \left( \sum_{i=1}^n \sum_{j=1}^c 1\{y_i = j\} \log \frac{e^{\theta_j^T \mathbf{x}_i}}{\sum_{j=1}^c e^{\theta_j^T \mathbf{x}_i}} \right) + \frac{\lambda}{2} \sum_{j=1}^c \|\theta_j\|_2^2 \quad (5)$$

其中， $n$  为训练样本数， $1\{y_i = j\}$  与示性函数。当第  $i$  个训练样本属于第  $j$  类时，该示性函数等于 1，否则等于 0。优化模型的第二项为正则化项，它使得优化函数为严格凸函数。用梯度下降法可以求得该模型的最优解，得到 Softmax 回归多类别分类器的参数  $\theta_j$ ， $j=1, 2, \dots, c$ 。

(3) BP 神经网络: 神经网络已被广泛应用于代谢组学分类问题，在对肿瘤、肺癌、慢性肝病等疾病的预测上都有应用。BP 神经网络是机器学习中的一种非线性分类器，通过误差逆传播 (Error Back Propagation) 算法训练得到一种多层前馈神经网络。神经网络隐层层数和神经元个数的确定是非常复杂的问题，不同的网络结构会有不同的分类结果，引用 Kolmogorov 定理来确定神经网络的结构 [31]。

**Kolmogorov 定理:** 给定任一连续函数  $f: U^d \rightarrow R^m$ ，即  $f(\mathbf{x}) = \mathbf{y}$ ，其中  $U$  为闭区间  $[0, 1]$ ，那么  $f$  可以精确地用一个 3 层前向网络实现。此网络的输入层有  $d$  个神经元，隐层有  $2d+1$  个神经元，输出层有  $m$  个神经元。

采用该 3 层前馈神经网络作为分类器，其中输入层有 3 个神经元，对应 3 个腰椎间盘退变代谢指标，隐层神经元数为 7，输出层神经元数为 4。此外，神经网络的隐层和输出层神经元的激活函数均选择为 sigmoid 函数  $g(z) = (1 + e^{-z})^{-1}$ 。因此该神经网络的数学形式可以表示为：

$$f_{BP}(\mathbf{x}) = \left( 1 + \exp \left( -\boldsymbol{\omega}^T \left( 1 + \exp(-\mathbf{v}^T \mathbf{x} + \gamma) \right)^{-1} + \theta \right) \right)^{-1} \quad (6)$$

其中， $\mathbf{v} \in R^{21}$  为输入层和隐层之间的连接权向量， $\boldsymbol{\omega} \in R^{28}$  为隐层和输出层之间的连接权向量， $\gamma \in R^7$  和  $\theta \in R^4$  分别为隐层和输出层神经元的阈值向量。分类函数  $f_{BP}(\mathbf{x}) \in R^4$  输出为一个向量，样本  $\mathbf{x}$  的退变等级预测为向量中最大值位置所对应的等级。

为了优化求解分类器参数，取优化目标为训练集累积误差的最小化，即：

$$\min_{\mathbf{v}, \boldsymbol{\omega}, \gamma, \theta} E = \frac{1}{n} \sum_{i=1}^n E_i = \frac{1}{2n} \sum_{i=1}^n (f_{BP}(\mathbf{x}_i) - y_i)^2 \quad (7)$$

其中， $\mathbf{x}_i$  和  $n$  为 BP 神经网络分类器的训练样本和训练样本数， $y_i$  是第  $i$  个训练样本的输出，训练前将腰椎间盘退变等级为 I 的样本标签设置为 (1, 0, 0, 0)，退变等级为 II 的样本标签设置为 (0, 1, 0, 0)，退变等级为 III 的样本标签设置为 (0, 0, 1, 0)，退变等级为 IV 的样本标签设置为 (0, 0, 0, 1)。

使用梯度下降法求解优化问题 (7)，不断修正 BP 神经网络的多个连接权和阈值，直到累积误差小于限定范围或迭代次数达到预设值后停止修正。连接权和阈值的修正模型统一选择为  $w \leftarrow w + \alpha \Delta w$ ，其中学习率  $\alpha = 0.02$ 。

(4) 支持向量机: 支持向量机是一种可处理过拟合现象的分类器 [32]。文献 [20] 基于脊椎强度和椎间盘形状特征等数据，采用支持向量机辅助诊断腰椎间盘突出。但该文献没有考虑腰椎间盘退变等级的诊断问题。

实验获得的每个样本数据包含了用于预测腰椎间盘退变等级的 3 个生化代谢指标。建立由该 3 个生化代谢指标组成的三维坐标系，并把获得的 390 个样本数据展示在该坐标系上可以发现，属于不同退变等级的样本之间很难线性划分，如图 4 所示。而采用支持向量机方法，可以通过引入非线性映射  $\phi(\mathbf{x})$ ，将原来的样本数据映射到另一个线性可分的高维空间，再寻找一个最优超平面进行分类。

由于要划分 4 个类别，先用 one-against-one 方法训练支持向量机分类器，即每次选取不同类别进行两两组合，构建  $c(c-1)/2 = 6$  个支持向量机二分类器。然而，如果直接用这些分类器对测试集进行投票预测，可能会出现分类重叠的现象，即多个类别得票相同，无法预测样本类别。因此，在得到 6 个支持向量机二分类器后，用 Platt 提出的两类分类概率建模方法将每个支持向量机二分类器输出转换为后验概率，最后再用投票法得到样本属于每个类别的后验概率 [33]。

**支持向量机二分类器:** 在训练支持向量机分类器前，将正类样本的类别记为 +1，负类样本的类别记为 -1，即  $y \in \{-1, +1\}$ 。支持向量机的数学表达式为

$$f_{SVM}(\mathbf{x}) = \sum_{i=1}^l \alpha_i y_i \exp(-\gamma \|\mathbf{x} - \mathbf{x}_i\|^2) + b \quad (8)$$

其中， $l$  是支持向量的个数， $\alpha_i$  为拉格朗日乘子， $\mathbf{x}_i$  为支持向量， $y_i$  为支持向量对应的类别标记， $b$  为模型参数。显然，

只有  $\alpha_i > 0$  对应的样本对分类器才有影响。当样本  $\mathbf{x}$  的分类函数值  $f_{SVM}(\mathbf{x})$  不小于 0 时, 则该样本为正类, 否则为负类。

为了最大化 2 个类别之间的分类间隔, 由拉格朗日乘子法可以得到以下优化问题:

$$\max_{\alpha} \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j y_i y_j \kappa(\mathbf{x}_i, \mathbf{x}_j), \quad (9)$$

$$s.t. \sum_{i=1}^n \alpha_i y_i = 0, \quad 0 \leq \alpha_i \leq C, \quad i = 1, \dots, n$$

其中,  $C > 0$  为常数,  $n$  为训练样本数,  $\kappa(\mathbf{x}_i, \mathbf{x}_j) = \phi(\mathbf{x}_i)^T \phi(\mathbf{x}_j)$  称为核函数。该文选用常用的径向基核函数  $\kappa(\mathbf{x}, \mathbf{x}_j) = \exp(-\gamma \|\mathbf{x}_i - \mathbf{x}_j\|^2)$ 。

采用 SMO 算法求解优化问题 (9) 中的参数  $\alpha_i (i = 1, \dots, n)$  和  $b$ 。借助文献 [34] 开发的一种 SMO 软件包 LIBSVM 来完成该优化问题的求解, 其中的参数  $C$  和  $\gamma$  由软件包中的 SVMcgForClass 函数进行网格搜索来确定。

**多分类后验概率支持向量机:** 在训练集构建 6 个支持向量机二分类器后, 用 sigmoid 函数将支持向量机二分类器输出映射到  $[0, 1]$ , 形式如下:

$$P_{kl}(k|x) = \frac{1}{1 + \exp(A^{kl} f_{SVM}^{kl} + B^{kl})}, \quad k, l = 1, \dots, 4, \quad k \neq l \quad (10)$$

$f_{SVM}^{kl}$  是类别  $k$  和类别  $l$  构成的支持向量机二分类器在对应训练集的输出值,  $P_{kl}(k|x)$  只表示支持向量机二分类器中样本  $\mathbf{x}$  属于类别  $k$  的后验概率, 简称为  $P_{kl}$ 。 $A^{kl}$  和  $B^{kl}$  是需要优化的参数, 可以通过最小化对应训练集的负对数似然函数得到:

$$\min_{A^{kl}, B^{kl}} - \sum_{i=1}^{n_{kl}} (t_i \ln(P_{kl}) + (1-t_i) \ln(1-P_{kl})), \quad (11)$$

$$t_i = \begin{cases} (N_+ + 1)/(N_+ + 2), & y_i = +1 \\ 1/(N_- + 1), & y_i = -1 \end{cases}, \quad i = 1, \dots, n_{kl}$$

$n_{kl}$  是类别  $k$  和  $l$  组成的训练样本数。该文使用梯度下降法求解上述优化问题, 可以得到 6 对参数  $A^{kl}$  和  $B^{kl}$ 。最后将 6 个支持向量机二分类器应用于测试集, 每个测试样本都得到 6 个支持向量机输出值  $f_{SVM}^{kl}$ , 再将其代入对应的 sigmoid 函数中计算得到 6 个  $P_{kl}$  值。使用投票法计算测试样本  $\mathbf{x}$  属于类别  $k$  的后验概率为:

$$P(k|x) = \frac{\sum_{l=1, l \neq k}^c P_{kl}}{\sum_{k=1, k \neq l}^c \sum_{l=1, l \neq k}^c P_{kl}}, \quad k = 1, \dots, c \quad (12)$$

将最大后验概率对应的类别作为样本  $\mathbf{x}$  的预测等级。

## 2 结果与分析 Results and analysis

**2.1 统计分析结果** 对收集到的 390 个受检者样本数据进行 Spearman 相关性计算, 得到腰椎间盘退变的 Pfirrmann 等级与该椎间盘相邻上下位椎体脂肪分数值和 T2\* 值的 Spearman 相关系数  $r_j'$  分别为 0.585, 0.560 和 -0.445, 因此, 椎间盘相邻上、下位椎体脂肪分数与退变等级呈正相关, 生化代谢指标 T2\* 值与退变等级呈负相关, 其中上位椎体脂肪分数与退变等级相关性最强。由于 3 种代谢指标的 Spearman 相关系数的绝对值  $|r_j'|$  都大于  $2.575/\sqrt{N-1} = 0.1515$ , 因此 3 种生化代

谢指标均与退变等级显著相关, 可以作为预测腰椎间盘退变 Pfirrmann 等级的特征。

**2.2 分类器预测过程** 为了展示分类过程, 从测试集随机挑选一个样本  $\mathbf{x} = (0.048 \ 9, 0.122 \ 3, 0.499 \ 5)^T$ , 分别用训练好的 softmax 回归、BP 神经网络和支持向量机分类器对该样本进行分类。

Softmax 回归分类器中的参数训练结果为:  $\theta_1 = (0.869 \ 3, -1.364 \ 8, -1.025 \ 1, 1.062 \ 1)^T$ ,  $\theta_2 = (0.046 \ 2, 0.044 \ 8, -0.002 \ 5, 0.624 \ 8)^T$ ,  $\theta_3 = (-0.361 \ 7, 0.239 \ 2, -0.445 \ 2)^T$ ,  $\theta_4 = (-0.552 \ 1, 0.935 \ 6, 0.804 \ 1, -1.248 \ 0)^T$ 。把  $\mathbf{x} = (1, \mathbf{x}^T)^T$  代入 (4) 式可得  $f_L(\mathbf{x}) = (0.584 \ 8, 0.250 \ 6, 0.102 \ 3, 0.062 \ 3)^T$ , 可见样本退变等级为 I 的概率最大, 因此, Softmax 回归将该样本预测为 I 级退变。

BP 神经网络分类器中的参数训练结果为:

$$v = \begin{bmatrix} 16.852 \ 3, & -20.390 \ 9, & -15.608 \ 4, & 24.941 \ 9, & 6.979 \ 6, & -11.289 \ 4, & 1.856 \ 5 \\ -5.368 \ 5, & -9.539 \ 3, & 9.668 \ 5, & -4.932 \ 6, & 22.137 \ 8, & -12.080 \ 3, & -1.826 \ 1 \\ -13.229, & -33.481 \ 2, & 13.141 \ 6, & 27.042 \ 8, & -13.714 \ 1, & 24.054 \ 7, & 8.393 \ 9 \end{bmatrix}$$

$$\gamma = (1.701, -26.173 \ 2, -0.485 \ 7, 18.858 \ 7, 10.762 \ 3, 7.325 \ 2, 0.731 \ 4)^T,$$

$$\omega^T = \begin{bmatrix} -5.418 \ 8, & -7.092, & 10.440 \ 5, & -8.413, & -1.481 \ 7, & 2.347 \ 1, & -0.686 \ 2 \\ -4.283 \ 5, & 6.670 \ 6, & 2.694 \ 5, & 8.611 \ 6, & 2.792 \ 5, & -1.601 \ 2, & -5.865 \ 5 \\ 17.940 \ 4, & -14.653 \ 1, & 14.415 \ 1, & -15.345 \ 1, & -9.897 \ 5, & -1.807 \ 2, & 2.990 \ 2 \\ -1.002, & 9.404 \ 7, & -2.912 \ 7, & 8.409 \ 2, & 4.447 \ 5, & -6.836 \ 5, & -6.574 \ 6 \end{bmatrix}$$

$\theta = (1.765 \ 1, 5.475 \ 4, 4.331 \ 4, 4.057)^T$ 。将样本  $\mathbf{x}$  代入 (6) 式可得,  $f_{BP}(\mathbf{x}) = (0.954 \ 3, 0.039 \ 1, 0.029 \ 4, 0)^T$ 。因此 BP 神经网络分类器将该样本预测为 I 级退变。

用 one-against-one 方法训练的 6 个支持向量机分类器  $f_{SVM}^{12}$ 、 $f_{SVM}^{13}$ 、 $f_{SVM}^{14}$ 、 $f_{SVM}^{23}$ 、 $f_{SVM}^{24}$ 、 $f_{SVM}^{34}$  对样本  $\mathbf{x}$  的输出结果分别为 1.077 6, 1.634 7, 1.116 3, -0.166 3, 0.951 3, 1.525 3。

将支持向量机输出转换为后验概率的 6 个 sigmoid 函数的参数值 A 和 B, 见表 2。

表 2 | 各 sigmoid 函数的参数  
Table 2 | Parameters of each sigmoid function

参数值	后验概率					
	$P_{12}$	$P_{13}$	$P_{14}$	$P_{23}$	$P_{24}$	$P_{34}$
A	1.853 5	2.032 4	1.942 8	1.822 2	1.760 2	1.316 5
B	-0.263 1	-0.142 4	0.237 8	-0.091 9	0.043 1	0.032 1

将样本的支持向量机输出  $f_{SVM}^{kl}$  和表 2 的参数代入 (10) 式可得  $P_{12} = 0.85$ ,  $P_{13} = 0.960 \ 1$ ,  $P_{14} = 0.917 \ 3$ ,  $P_{23} = 0.402 \ 5$ ,  $P_{24} = 0.847 \ 8$ ,  $P_{34} = 0.884 \ 9$ 。最后根据式 (12) 计算出样本属于每个类别的后验概率分别为:  $P=(1|\mathbf{x}) = 0.455$ ,  $P=(2|\mathbf{x}) = 0.233$ ,  $P=(3|\mathbf{x}) = 0.254$ ,  $P=(4|\mathbf{x}) = 0.058$ 。可见样本属于第一类的后验概率最大, 因此支持向量机分类器预测该腰椎间盘样本属于 I 级退变。

**2.3 分类器评估** Softmax 回归、BP 神经网络和支持向量机分类器分别对 130 个测试样本进行分类后得到分类结果, 见表 3-5 所示。分别计算每个分类器的准确率、kappa 系数、“宏查全率”(macro-R)、“宏查准率”(macro-P)和“宏  $F_1$ ”(macro- $F_1$ )。准确率即总共预测正确的样本数比测试集总样本数的值, kappa 系数是用来衡量分类器预测结果和实际分类的一致性的指标, 计算公式如下:

$$\kappa = \frac{P_0 - P_e}{1 - P_e} \quad (13)$$

其中  $P_0$  为正确率,  $P_e$  称为期望一致率, 假设各类别的真实样本数分别  $a_1, \dots, a_c$ , 各类别的预测样本数分别  $b_1, \dots, b_c$ , 总样本数为  $n$ , 则  $p_e = \sum_{i=1}^c a_i * b_i / n^2$ 。Kappa 系数取值范围在  $[-1, 1]$ , 值越大代表分类准确度越高。此外, 将各分类器的输出结果拆分成 4 个的混淆矩阵 (表 6), 分别

表 3 | Softmax 回归分类器的输出结果

Table 3 | Output of Softmax classifier

真实等级	预测等级				合计
	I 级	II 级	III 级	IV 级	
I 级	32	12	0	0	44
II 级	18	21	0	2	41
III 级	8	11	0	3	22
IV 级	1	13	0	9	23
合计	59	57	0	14	130

表 4 | 神经网络分类器的输出结果

Table 4 | Output of neural network classifier

真实等级	预测等级				合计
	I 级	II 级	III 级	IV 级	
I 级	29	14	1	0	44
II 级	14	22	2	3	41
III 级	8	8	5	1	22
IV 级	3	4	5	11	23
合计	54	48	13	15	130

表 5 | 支持向量机分类器的输出结果

Table 5 | Output of support vector machine classifier

真实等级	预测等级				合计
	I 级	II 级	III 级	IV 级	
I 级	31	11	2	0	44
II 级	16	19	2	4	41
III 级	8	5	6	3	22
IV 级	1	5	5	12	23
合计	56	40	15	19	130

表 6 | 不同分类器各个等级的混淆矩阵

Table 6 | Confusion matrices of various levels in different classifiers

分类器	真实类别	预测类别 (I 级为正常)		预测类别 (II 级为正常)		预测类别 (III 级为正常)		预测类别 (IV 级为正常)	
		正类	负类	正类	负类	正类	负类	正类	负类
		正类	负类	正类	负类	正类	负类	正类	负类
Softmax 回归	正类	32	12	21	20	0	22	9	14
	负类	27	59	36	53	0	108	5	102
	合计	59	71	57	73	0	130	14	116
神经网络	正类	29	15	22	19	5	17	11	12
	负类	25	61	26	63	8	100	4	103
	合计	54	76	48	82	13	117	15	115
支持向量机	正类	31	13	19	22	6	16	12	11
	负类	25	61	21	68	9	99	7	100
	合计	56	74	40	90	15	115	19	111

以 I - IV 级为正常, 其他为负类, 从而计算 4 个退变等级的查全率  $R_1-R_4$ , 以及查准率  $P_1-P_4$ 。查全率是正类的样本被成功预测的概率, 查准率是预测为正类的样本中确实为正类的概率。宏查全率即为 4 个混淆矩阵查全率的均值, 宏查准率为 4 个查准率的均值, 再得到 macro- $F_1$  为:

$$\text{macro-}F_1 = \frac{2 \times \text{macro-}P \times \text{macro-}R}{\text{macro-}P + \text{macro-}R} \quad (14)$$

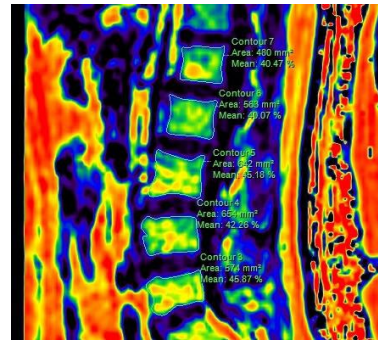


图 1 | 椎体脂肪分数测量  
Figure 1 | Measurement of vertebral body fat fraction

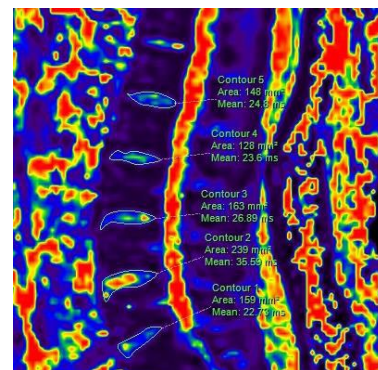


图 2 | 腰椎间盘 T2\* 值测量  
Figure 2 | Measurement of T2\* value of lumbar intervertebral disc

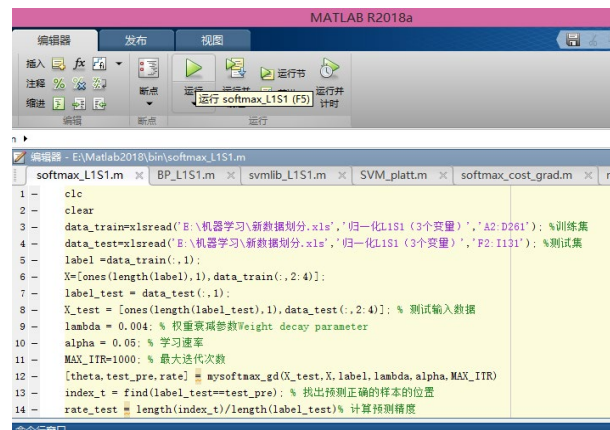


图 3 | 机器学习界面  
Figure 3 | Interface of machine learning

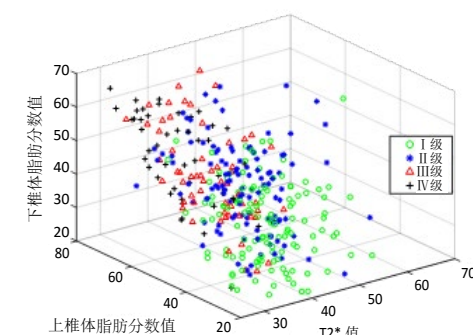


图 4 | 不同 Pfirrmann 等级的样本点  
Figure 4 | Sample points of different Pfirrmann grades

**2.3.1 Softmax 回归分类器结果** 测试集中 I - IV 级退变样本分别有 44, 41, 22, 23 个。由表 3 可知 Softmax 回归分类器对 I - IV 退变等级分类正确的样本数分别有 32, 21, 0, 9 个, 分类正确率为 0.477,  $P_e=(44 \times 59+41 \times 57+22 \times 0+23 \times 14)/130^2=0.311$ , kappa 系数为 0.311。由表 6 可计算退变等级 I - IV 的混淆矩阵的查全率分别为 0.727(32/44), 0.512(21/41), 0(0/22), 0.391(9/23), 查准率分别为 0.542(32/59), 0.368(21/57), 0(0/0), 0.643(9/14)。取平均得到 macro-R 为 0.408, macro-P 为 0.388, macro-F<sub>1</sub> 为 0.398。从 I - IV 级退变的混淆矩阵的查全率来看, 分类器对退变 I 级的样本有较高的识别率, 其次是对 II 级样本的识别, III 级和 IV 级样本识别效果较差, 甚至没有识别出 III 级退变样本。

**2.3.2 神经网络分类器结果** 由表 4 可知, 神经网络分类器对 I - IV 退变等级分类正确的样本数分别有 29, 22, 5, 11 个, 分类正确率  $P_0$  为 0.515,  $P_e=(44 \times 54+41 \times 48+22 \times 13+23 \times 15)/130^2=0.294$ , kappa 系数为 0.3。退变等级 I - IV 的混淆矩阵的查全率分别为 0.659(29/44), 0.537(22/41), 0.227(5/22), 0.478(11/23), 查准率分别为 0.537(29/54), 0.458(22/48), 0.385(5/13), 0.733(11/15)。取平均得到 macro-R 为 0.475, macro-P 为 0.528, macro-F<sub>1</sub> 为 0.5。神经网络分类器对各级退变样本的识别率从高到低分别为 I 级、II 级、IV 级和 III 级。

**2.3.3 支持向量机分类器结果** 由表 5 可知, 神经网络分类器对 I - IV 退变等级分类正确的样本数分别有 31, 19, 6, 12 个, 分类正确率  $P_0$  为 0.523,  $P_e=(44 \times 56+41 \times 40+22 \times 15+23 \times 19)/130^2=0.288$ , kappa 系数为 0.33。退变等级 I - IV 的混淆矩阵的查全率分别为 0.705(31/44), 0.463(19/41), 0.273(6/22), 0.522(12/23), 查准率分别为 0.554(31/56), 0.475(19/40), 0.4(6/15), 0.632(12/19)。取平均得到 macro-R 为 0.491, macro-P 为 0.515, macro-F<sub>1</sub> 为 0.503。支持向量机分类器对各级退变样本的识别率从高到低分别为 I 级、IV 级、II 级和 III 级。

### 3 讨论 Discussion

回归算法、支持向量机、神经网络等在疾病预测和代谢组学分析方面都有着广泛的应用且取得显著成果。在代谢组学中, 郗群等<sup>[9]</sup>就应用逻辑回归、神经网络和支持向量机 3 种方法基于多种血清标志物建立了结直肠癌早期诊断模型, 逻辑回归和支持向量机模型都获得可观的分类效果, 但认为神经网络受隐层节点数等超参数影响缺乏一定的稳定性。张岩等<sup>[8]</sup>也使用逻辑回归对结直肠癌的 4 种血清标志物进行分析, 结果发现逻辑回归能同时提高诊断的敏感度和特异度。可见逻辑回归作为一种简单有效的方法在疾病诊断和代谢组学分析非常受欢迎。实验使用的 Softmax 回归是 logistic 回归在多分类任务中的扩展, 它对退变 I 级的识别率是 3 种模型中最高的, 但对 III 级退变样本识别上有较大缺陷, 也许是因为 Softmax 回归模型复杂度较低, 无法很好拟合所有数据以致测试结果比较一般。神经网络是适合处理非线性问题的机

器学习方法, 善于建立输入特征和输出类别的非线性关系。MANI-VARNOSFADERANI 等<sup>[12]</sup>基于髂嵴骨组织的矿物质建立预测鸦片成瘾的神经网络, 发现预测效果比线性模型的要好。常炳国等<sup>[13]</sup>也基于代谢数据建立神经网络, 对慢性肝病预测有一定精度。实验建立的神经网络比 Softmax 回归有更高的准确率、宏查全率、宏查准率和宏 F1 值, 且与支持向量机分类效果不相上下, 但神经网络只建立一个分类器就可以实现多分类, 而此次实验的支持向量机方法需要建立多个二分类器, 从减少训练时间和降低模型复杂度的角度来看, 神经网络比支持向量机更有优势。另外, 代谢组学数据通常具有小样本、高噪声等特征, 而支持向量机是一种能够有效处理样本小和非线性等实际问题的方法。SONG 等<sup>[10]</sup>使用糖尿病认知功能障碍的生物代谢指标建立了准确性较高的支持向量机分类器, 为糖尿病认知功能障碍早期诊断提供新的思路。WANG 等<sup>[11]</sup>应用支持向量机对肝癌分类, 发现整合蛋白组学和代谢组学等数据能改善诊断效果。此次实验使用的 3 种分类器中支持向量机的分类准确率、macro-R 和 macro-F<sub>1</sub> 最高, 因此认为将支持向量机应用于腰椎间盘的代谢组学数据中是值得考虑的。

虽然目前已有部分关于腰椎间盘退变等级计算机辅助诊断的研究, 但与此次实验使用生化代谢指标建立腰椎间盘退变诊断模型不同, 这些研究大多数是基于腰椎间盘的图像特征进行, 例如 CASTRO-MATEOS 等<sup>[19]</sup>用 one-against-all 策略将 Adaboost 算法应用于腰椎间盘图像特征, 自动识别 Pfirrmann I & II 级、III 级、IV 级和 V 级样本, 得到了不错的分类效果。但因为采用了 one-against-all 策略, 因此在训练模型时需要建立 4 个分类器, 在节省训练分类器时间的层面看, 此次实验使用的 Softmax 回归和 BP 神经网络更有利。另外, 由于识别 Pfirrmann 等级中 I 级退变对疾病预防和及早治疗有重大意义, 此次实验对 I 级退变样本也做了分类, 测试结果显示, 基于生化代谢指标建立的 3 种诊断模型都对腰椎间盘 I 级退变样本有较高识别率, 但是对其他退变等级样本的识别效果一般。HE 等<sup>[20]</sup>则引入了有监督距离学习方法, 基于腰椎间盘图像特征自动将腰椎间盘退变等级分为 4 级, 未说明是对腰椎间盘退变的 Pfirrmann 等级分类。此次实验使用世界范围内得到广泛认可的 Pfirrmann 标准, 在判别腰椎间盘退变等级时更有可靠性。多研究表明, 腰椎间盘 T2\* 值与 Pfirrmann 等级呈负相关<sup>[23]</sup>, 这在此次实验中也得到证实。李勉文等<sup>[25]</sup>利用 mDIXON-Quant 技术测量椎体终板下骨髓脂肪分数 FF 值, 检验发现上位、下位椎体脂肪分数值均在不同 Pfirrmann 等级间有显著性差异, 实验经过 Spearman 相关分析得到上、下位椎体脂肪分数值与 Pfirrmann 等级是显著相关的。经过统计相关性分析, 作者认为采用椎间盘相邻上下位椎体脂肪分数值和 T2\* 值建立腰椎间盘退变计算机辅助诊断模型是合理的。

实验研究基于代谢组学的腰椎间盘退变计算机辅助诊断方法, 通过 390 个腰椎间盘代谢指标检测值, 验证了椎间

盘相邻上下位椎体脂肪分数值和 T2\* 值等 3 种生化代谢指标与腰椎间盘突出退变显著相关, 并应用 MSSS 方法把 390 个检测值划分出训练样本集和测试样本集, 保证了训练样本集和测试样本集的一致性; 基于 260 个训练样本建立了 Softmax 回归、神经网络和支持向量机等 3 种分类器, 并在测试样本集上进行分类器性能的比较。结果表明, 采用 MRI 代谢指标建立腰椎间盘突出退变诊断分类器是可行性的。此次实验提供了一种新的腰椎间盘突出退变诊断计算机辅助方法, 值得后续进一步研究。

**致谢:** 在此特别感谢第九〇九医院医学影像科影像医师在此次研究数据收集整理上提供的帮助。

**作者贡献:** 吴晓锋和欧阳林进行实验设计, 实验实施为江丽红, 实验评估为江丽红, 资料收集为罗爱芳和黄丽, 江丽红成文, 吴晓锋和欧阳林审校。

**经费支持:** 该文章接受了“福建省科技计划项目 (2019Y31010067) 及第九〇九医院青年苗圃基金 (18Y021)”的资助。所有作者声明, 经费支持没有影响文章观点和对研究数据客观结果的统计分析及其报道。

**利益冲突:** 文章的全部作者声明, 在课题研究和文章撰写过程中不存在利益冲突。

**写作指南:** 该研究遵守国际医学期刊编辑委员会《学术研究实验与报告和医学期刊编辑与发表的推荐规范》。

**文章查重:** 文章出版前已经过专业反剽窃文献检测系统进行 3 次查重。

**文章外审:** 文章经小同行外审专家双盲外审, 同行评议认为文章符合期刊发稿宗旨。

**文章版权:** 文章出版前杂志已与全体作者授权人签署了版权相关协议。

**开放获取声明:** 这是一篇开放获取文章, 根据《知识共享许可协议》“署名-非商业性使用-相同方式共享 4.0”条款, 在合理引用的情况下, 允许他人以非商业性目的基于原文内容编辑、调整和扩展, 同时允许任何用户阅读、下载、拷贝、传递、打印、检索、超级链接该文献, 并为之建立索引, 用作软件的输入数据或其它任何合法用途。

#### 4 参考文献 References

- [1] WALDENBERG C, HEBELKA H, BRISBY H, et al. MRI histogram analysis enables objective and continuous classification of intervertebral disc degeneration. *Eur Spine J.* 2018;27(5):1042-1048.
- [2] 杨立杰, 欧阳林, 陈鼎伟, 等. 下腰疼痛的生物力学分析 [J]. *中国组织工程研究*, 2020,24(33):5267-5271.
- [3] 李钦海, 陈光玉. DWI 和 DTI 对腰椎间盘突出早期退变的诊断价值 [J]. *中国实用医药*, 2016,11(19):74-75.
- [4] 凌人男, 龚静山, 马捷. 腰椎间盘突出退变的 MRI 早期诊断和定量评价的研究进展 [J]. *医学影像学杂志*, 2016,26(10):1918-1920.
- [5] 赵寅, 周盛源, 苑博, 等. X 射线九分法评估腰椎间盘退变 [J]. *中国组织工程研究*, 2018,22(3):450-455.
- [6] 龚静山, 凌人男, 周阳洪, 等. 采用 3.0T MRI T2 和 T2\* mapping 评价下腰痛患者腰椎间盘突出退变 [J]. *中国 CT 和 MRI 杂志*, 2016,14(10):113-116.
- [7] ZHAO H, HEIMBERGER AB, LU Z, et al. Metabolomics profiling in plasma samples from glioma patients correlates with tumor phenotypes. *Oncotarget.* 2016;7(15):20486-20495.
- [8] 张岩, 杨桃. logistic 回归和 ROC 曲线评价 4 种血清肿瘤标志物在结肠癌中的诊断价值 [J]. *世界最新医学信息文摘*, 2018,18(58):25-27.
- [9] 郝群, 毛文虹. 基于机器学习的结肠癌血清标志物筛选及早期诊断模型评估 [J]. *解放军预防医学杂志*, 2016,34(6):879-883.
- [10] SONG L, ZHUANG P, LIN M, et al. Urine Metabolomics Reveals Early Biomarkers in Diabetic Cognitive Dysfunction. *J Proteome Res.* 2017;16(9):3180-3189.

- [11] WANG M, YU G, RESSOM HW, et al. Integrative Analysis of Proteomic, Glycomic, and Metabolomic Data for Biomarker Discovery. *IEEE J Biomed Health Inform.* 2016;20(5):1225-1231.
- [12] MANI-VARNOSFADERANI A, JAMSHIDI M, YEGANEH A, et al. Concentration profiling of minerals in iliac crest bone tissue of opium addicted humans using inductively coupled plasma and discriminant analysis techniques. *J Pharm Biomed Anal.* 2016;120:92-99.
- [13] 常炳国, 李玉琴, 冯智超, 等. 基于主成分机器学习算法的慢性肝病的智能预测新方法 [J]. *计算机科学*, 2017,44(Z11):65-67,91.
- [14] 李祥蓉, 程琳, 席家宁, 等. 三种模式识别模型诊断腰椎间盘突出症受压神经根的准确率 [J]. *中国组织工程研究*, 2018,22(19):3005-3013.
- [15] 覃伟, 高敏, 沈莹, 等. 基于机器学习算法的 2 型糖尿病患者 3 个月血糖预测 [J]. *中华疾病控制杂志*, 2019,23(11):1313-1317.
- [16] DETERMAN JR CE. Optimal algorithm for metabolomics classification and feature selection varies by dataset. *Int J Biol.* 2015;7(1):100-105.
- [17] BOUNDS DG, LLOYD PJ, MATHEW BG, et al. A multilayer perceptron network for the diagnosis of low back pain. *IEEE 1988 International Conference on Neural Networks, IEEE, 1988:481-489.*
- [18] KARABULUT EM, IBRIKCI T. Effective Automated Prediction of Vertebral Column Pathologies Based on Logistic Model Tree with SMOTE Preprocessing. *J Med Syst.* 2014;38(5):1-9.
- [19] CASTRO-MATEOS I, POZO JM, LAZARY A, et al. 2D segmentation of intervertebral discs and its degree of degeneration from T2-weighted magnetic resonance images. *Proc SPIE 9035, Medical Imaging 2014: Computer-Aided Diagnosis*, 903517.
- [20] HE X, LANDIS M, LEUNG S, et al. Automated grading of lumbar disc degeneration via supervised distance metric learning. *Proc SPIE 10134, Medical Imaging*, 2017: Computer-Aided Diagnosis, 1013443.
- [21] EBRAHIMZADEH E, FAYAZ F, NIKRAVAN M, et al. Towards an Automatic Diagnosis System for Lumbar Disc Herniation: The Significance of Local Subset Feature Selection. *Biomed Eng Appl Basis Commun.* 2018;30(6):1850044.
- [22] ALOMARI RS, CORSO JJ, CHAUDHARY V, et al. Toward a clinical lumbar CAD: herniation diagnosis. *Int J Comput Assist Radiol Surg.* 2011;6(1):119-126.
- [23] 熊玉超, 曾旭文, 梁治平, 等. T2\*-mapping 在腰椎间盘突出退行性变中的研究进展 [J]. *生物医学工程与临床*, 2019,23(2):239-242.
- [24] 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.
- [25] 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]. *临床放射学杂志*, 2017,36(12):1838-1842.
- [26] 陈先来, 杨路明. 基于均矢量相似性的机器学习样本集划分 [J]. *中南大学学报 (自然科学版)*, 2009,40(6):1636-1641.
- [27] 龚凤乾. 等级相关系数公式的证明应用及统计显著性检验 [J]. *天津财经学院学报: 现代财经*, 1987(1): 33-36.
- [28] 何艳频, 孙爱峰. Spearman 等级相关系数计算公式及其相互关系的探讨 [J]. *中国现代药物应用*, 2007,1(7): 72-73.
- [29] HANGAI M, KANEOKA K, KUNO S, et al. Factors associated with lumbar intervertebral disc degeneration in the elderly. *Spine J.* 2008;8(5):732-740.
- [30] ABUQASMIH I. Multi-class Multi-label Classification and Detection of Lumbar Intervertebral Disc Degeneration MR Images using Decision Tree Classifiers. *Comput Eng Intell Syst.* 2013;4(9):33-38.
- [31] 赵振宇, 徐用懋. 模糊理论和神经网络的基础与应用 [M]. 南宁: 广西科技出版社, 1996:85-86.
- [32] CORTES C, VAPNIK VN. Support Vector Networks. *Mach Learn.* 1995;20(3):273-297.
- [33] 肖小玲, 李腊元, 张翔, 等. 一种多类支持向量机概率建模新方法 [J]. *计算机工程*, 2006,32(20):28-29.
- [34] CHANG CC, LIN CJ. LIBSVM: A library for support vector machines. *ACM Trans Intell Syst Technol.* 2011;2(3):1-27.

(责任编辑: GW, ZN, SX)