

DOI: 10.16781/j.CN31-2187/R.20230373

· 论 著 ·

基于集成神经网络的类风湿关节炎中医证候分类器研究

杨晶东^{1*}, 江彪¹, 李熠伟¹, 姜泉², 韩曼², 宋梦歌²

1. 上海理工大学光电信息与计算机工程学院自主机器人实验室, 上海 200093

2. 中国中医科学院广安门医院风湿病科, 北京 100053

[摘要] **目的** 构建一种集成神经网络模型实现类风湿关节炎(RA)中医证候分类,并探究其中的特征重要性和风险因素。**方法** 针对基于人工智能技术的RA中医证候多标签分类中存在的标签关联性差、泛化性能低等问题,提出一种集成神经网络模型——集成神经网络链(FEN)。FEN模型采用一种基于深度神经网络的特征提取基分类器提取临床RA多标签样本的深层特征,增强RA特征区分度;根据协方差理论衡量标签相关性,调节分类器链的输入空间,减少RA错误信息传播和冗余度;并采用集成学习方法减小分类器链中不合理标签序列对RA特征分类的影响。此外,分析了RA中医证候主证和兼证的特征贡献度,挖掘其潜在的风险因素。**结果** FEN模型的10折交叉验证性能参数汉明损失、1-错误率、准确度和F1值分别为0.003 6、0.024 8、97.52%、99.18%。与7种典型多标签分类器(分类器链、标签幂集、二进制关联、随机k-标签集、多标签K最近邻、集成分类器链和集成二进制关联)相比,FEN模型具有较好的分类性能。特征贡献度分析提示,主证和次证特征均可作为RA中医证候分类的重要指标,是影响主证和兼证分类的主要因素。**结论** 基于集成神经网络模型的RA中医证候分类器具有较高的分类精度和效率,对于RA的临床诊断和治疗具有重要参考价值。

[关键词] 类风湿关节炎;多标签学习;神经网络;分类器链;集成学习

[引用本文] 杨晶东,江彪,李熠伟,等.基于集成神经网络的类风湿关节炎中医证候分类器研究[J].海军军医大学学报,2024,45(3):305-319. DOI: 10.16781/j.CN31-2187/R.20230373.

Classifiers for traditional Chinese medicine syndromes of rheumatoid arthritis based on ensemble neural network

YANG Jingdong^{1*}, JIANG Biao¹, LI Yiwei¹, JIANG Quan², HAN Man², SONG Mengge²

1. Autonomous Robot Lab, School of Optical-Electrical and Computer Engineering, University of Shanghai for Science and Technology, Shanghai 200093, China

2. Department of Rheumatology, Guang'anmen Hospital, China Academy of Chinese Medical Sciences, Beijing 100053, China

[Abstract] **Objective** To construct an ensemble neural network model to achieve traditional Chinese medicine (TCM) syndrome classification for rheumatoid arthritis (RA) and explore the importance of the features and risk factors. **Methods** An ensemble neural network model FEN (feature extration network) was proposed to solve the issues such as poor label correlation and low generalization performance in multi-label classification of TCM syndromes of RA based on artificial intelligence technology. The FEN model utilized a feature extraction classifier based on deep neural network to extract deep features from clinical multi-label RA samples, enhancing the discriminative power of RA features. By measuring label correlation based on covariance theory, the input space of the classifier chain was adjusted to reduce the spread of RA error information and redundancy. An ensemble learning method was used to mitigate the impact of unreasonable label sequences in the classifier chain on RA feature classification. Additionally, the importance of main and accompanying TCM syndrome features of RA was analyzed and potential risk factors were explored. **Results** The FEN model had excellent performance in a 10-fold cross-validation, with Hamming loss, one-error, accuracy, and F1-score being 0.003 6, 0.024 8, 97.52%, and 99.18%, respectively. Compared with 7 typical multi-label classifiers (classifier chain, label powerset, binary relevance, random k-labelsets, multi-label K-nearest neighbor, ensemble classifier chain, and ensemble binary relevance), the FEN model had better classification

[收稿日期] 2023-07-02 [接受日期] 2023-12-13

[基金项目] 国家自然科学基金(81973749),中国中医科学院科技创新工程重大攻关项目(CI2021A01503)。Supported by National Natural Science Foundation of China (81973749) and Major Research Project of Scientific and Technological Innovation Program of China Academy of Chinese Medical Sciences (CI2021A01503).

[作者简介] 杨晶东,博士,副教授。

*通信作者(Corresponding author). Tel: 021-55271064, E-mail: eerfriend@163.com

capabilities. The analysis of feature contribution indicated that the features of main and secondary symptoms might be used as important indicators of classification of TCM syndromes of RA, and were the main factors affecting the classification of main and accompanying syndromes. **Conclusion** The RA TCM syndrome classifier based on ensemble neural network has high classification accuracy and efficiency, providing important reference for the clinical diagnosis and treatment of RA.

[**Key words**] rheumatoid arthritis; multi-label learning; neural network; classifier chain; ensemble learning

[**Citation**] YANG J, JIANG B, LI Y, et al. Classifiers for traditional Chinese medicine syndromes of rheumatoid arthritis based on ensemble neural network[J]. Acad J Naval Med Univ, 2024, 45(3): 305-319. DOI: 10.16781/j.CN31-2187/R.20230373.

类风湿关节炎 (rheumatoid arthritis, RA) 是一种常见的自身免疫性风湿疾病, 主要特征表现为慢性关节滑膜病变, 通常影响手、膝和腕关节^[1]。有数据显示我国 RA 患病率为 0.2%~0.4%^[2]。目前与 RA 相关的中医研究主要集中在 RA 病因、发病机制、辨证分型、证候与实验室指标的相关性、临床治疗和疗效评估等方面^[3]。研究 RA 中医证候规律能为中医辨证论治提供依据^[4]。“至少 1 个关节存在滑膜炎且不是由其他疾病引起”是 RA 的诊断标准之一, 这一条要求有利于 RA 的诊断分类和有效治疗^[5]。中医证候的诊断是通过提取中医四诊信息中证候判别的症状并进行分类的过程。临床 RA 中医证候特征表现为多标签属性, 设计一个高效且准确的多标签分类模型, 对 RA 证候诊断具有重要意义。

深度学习神经网络具有强大的特征提取能力, 可以有效提取医学临床数据深层特征。Hügler 等^[6]提出了一种基于递归神经网络的自适应网络 (AdaptiveNet) 模型, 用于 RA 的疾病进展预测, 该方法采用长短期记忆网络提取临床 RA 样本间的长期依赖关系, 取得了较好预测精度, 但 RA 样本仅包括历史信息与实验室指标。Law 和 Ghosh^[7]提出了一种级联神经网络模型, 组合堆叠式自动编码器和极限学习机实现多标签分类, 通过自动编码器的特征编码技术增强样本差异性和减少多标签输入表征, 采用极限学习机计算类别分数并实现软分类。Hassan 等^[8]针对心血管自主神经病变构建了一种多级融合预测模型, 该模型结合了特征选择和多模态特征融合技术, 利用多类 ROC 曲线 AUC 值选择重要特征, 并融合生成新的数据集, 该方法虽然具有较好的分类精度, 但不完全依赖原始数据特征, 无法体现 RA 中医证候多标签数据特点。Pleiter 等^[9]在量表数据中引入新的后门攻击策略, 将专为联合学习设计的后门攻击调整为集中式设

置, 还开发了 2 种隐形攻击变体, 为保护深度神经网络 (deep neural network, DNN) 模型免受量表数据后门攻击提供有效支持。

多标签分类任务的机器学习方法主要包括算法自适应法和问题转换法^[10]。算法自适应法扩展了现有的单标签分类方法, 以适应多标签数据^[11]。多标签 K 最近邻 (multi-label K-nearest neighbor, ML-KNN) 通过改进 K 最近邻算法^[12]将一级和二级邻域信息结合解决多标签问题。排序支持向量机使用成对排序, 根据结果与特定查询的关联度自适应排序结果^[13]。随机 k-标签集 (random k-labelsets, RakEL) 将初始标签集分解为若干随机子集, 并利用标签幂集 (label powerset, LP) 训练相应的分类器^[14]。该方法通过改进单标签方法实现多标签分类, 却完全忽略了标签间的相关性。问题转换是将多标签分类问题转化为多个单标签二分类问题, 如二进制关联 (binary relevance, BR)^[15]和分类器链 (classifier chain)^[16]。BR 方法将每个标签作为一个独立的二进制分类问题, 但未考虑标签间的相关性。分类器链则利用标签集的输入顺序与局部标签的关联信息构建一个链状结构, 但不同的标签集顺序和错误传播会影响分类器链的预测性能。Eskandari 和 Ghassabi^[17]提出了一种基于信息论滤波器的多标签特征选择方法, 通过协同算法和问题转换策略对特征进行排序, 同时考虑个体标签的影响及抽象标签空间中的判别潜力。Lou 等^[18]通过改进模糊系统结构并利用模糊规则对特征和标签之间的关系进行建模, 提出了一种新的多标签分类方法, 其能够有效地学习多标签分类中标签和特征之间的隐藏关系, 提高了分类性能。分类器链的分类能力依赖于标签相关性。Zhang 等^[19]在多标签文本分类中用成对标签共现预测 (pairwise label co-occurrence prediction) 和条件标签共现预测 (conditional label co-occurrence prediction) 实现

共现预测任务,增强了标签相关性,减少了标签过拟合和误差传播,实验结果表明该方法具有较好的学习标签相关性。Li等^[20]研究了有噪声的多标签转移矩阵的估计问题,提出了一种新的转移矩阵估计器,充分利用标签相关性,实现了多标签数据集转移矩阵准确估计。Mencía等^[21]提出了一种基于随机决策树的预测方法,可以动态选择标签序列来增强标签之间的依赖关系。Han等^[22]通过挖掘标签之间的统计关系构建了一个邻接矩阵来对标签相关性进行建模,其使用了图卷积网络学习相邻节点间信息传播规律,实验结果表明模型具有较好的效果。Bakhshi和Can^[23]提出一种多标签数据流分类方法,引入基于单标签分类器预测的加权机制,并考虑二进制标签之间的相关性,使模型的有效性和效率达到均衡。

本文针对临床RA中医证候多标签样本,提出了一种基于特征提取的集成神经网络模型——集成神经网络链(feature extraction network, FEN)。该模型使用DNN构建一个基分类器特征提取分类器(feature extraction classifier, FEC),能够提取RA数据中的特征层次结构,学习输入和输出之间的复杂关系;根据协方差理论衡量标签间的相关性,确定特征提取分类器链(feature extraction classifier chain, FECC)的标签集输入,减少了弱相关或负相关标签之间的干扰;还采用集成学习方法构建多条关联分类器链,减少了不合理的标签序列对分类的影响,提高了分类性能。

1 基于集成学习的多标签分类器FEN模型的构建

1.1 基于DNN的特征提取基分类器FEC

分类器链是解决多标签分类问题的一种常用方法,一般采用基分类器学习数据特征,而基分类器由基于序列最小最优化算法的支持向量机构成。在传统神经网络模型中,循环神经网络主要用于处理序列数据,在自然语言处理领域应用广泛,卷积神经网络在数字图像处理领域取得了巨大的成功,DNN则更适合处理临床医学表格数据。在处理多标签数据时,需要考虑数据样本的复杂性和关联性,DNN具有强大的数据特征学习能力,能学习到标签数据间的深层特征。利用DNN构建一个基分类器,可以学习RA中医证候样本的深层特征。DNN内部的神经网络层可以分为3类,分别为输入层、输出层和隐藏层(图1A)。神经网络中的隐藏层能够提取数据的深层特征,并且随着隐藏层层数的增加其特征提取能力增强。然而,隐藏层层数过多会导致网络模型过拟合等问题。本研究设计了一个4层的网络模型FEC,如图1B所示。网络的输入层为第1层,中间2层为隐藏层,最后1层为输出层,同时修改了输出层的节点个数,将该网络作为FEN模型中的基分类器FEC。其中, x_a 表示第 a 个输入单元, h_n 和 h_m 分别表示第 n 个和第 m 个隐藏层单元, y_b 和 y_k 分别为第 b 个和第 k 个输出单元, k 对应第 k 个标签。隐藏层由线性层(linear)和ReLU激活函数构成,输出层的激活函数为Sigmoid。

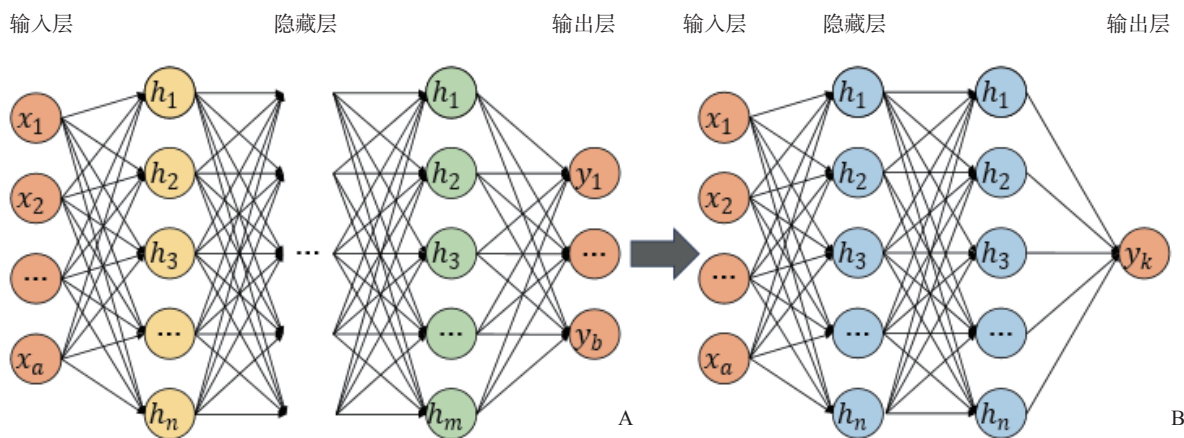


图1 DNN (A) 和 FEC (B) 的网络结构

Fig 1 Network structure of DNN (A) and FEC (B)

DNN: 深度神经网络; FEC: 特征提取分类器. x_a : 第 a 个输入单元; h_n : 第 n 个隐藏层单元; h_m : 第 m 个隐藏层单元; y_b : 第 b 个输出单元; y_k : 第 k 个输出单元; k : 第 k 个标签.

1.2 基于标签相关性的 FECC 传统的分类器链结构由输入的标签集顺序确定标签链,链中每个基分类器的输入空间均接受先前的基分类器输出结果,并训练一个标签。分类器链可以让每个基分类器学习标签间的局部关联特征,使分类器具有更好的分类性能。假设 2 个强正相关的标签可以相互学习更多的特征,而 2 个不相关或负相关的标签可能会增加分类器训练时的数据冗余度,甚至发生误差传播影响分类精度。因此,解决不相关或负相关的标签问题可以有效提高分类器链的分类性能。

单个 FEC 可以学习多标签数据的深层特征,却忽略了标签间的关联信息。使用链式结构将 FEC 连接起来构建一条 FECC 能够有效学习不同 RA 中医证候之间的关联信息。本研究在标签集 $y = [y_1, \dots, y_{i_s}, \dots, y_j, \dots, y_k]$ 输送到分类器链前,使用一种相关系数衡量标签间的相关性。首先利用协方差公式度量 2 个标签间的相关性,协方差计算公式为

$$\text{cov}(y_i, y_j) = \frac{\sum_{n=1}^N (y_{ni} - \bar{y}_{ni})(y_{nj} - \bar{y}_{nj})}{N-1} \quad (1)$$

其中 N 为实例个数, y_i 和 y_j 分别为 2 个标签,标签 y_i 和 y_j 的相关性强弱取决于 $\text{cov}(y_i, y_j)$ 值;协方差受标签尺度影响,从协方差中引出相关系数 ρ 来衡量标签相关性, ρ 计算公式为

$$\rho = \frac{\text{cov}(y_i, y_j)}{\sigma_{y_i} \sigma_{y_j}} \quad (2)$$

其中 σ_{y_i} 和 σ_{y_j} 分别为 2 个标签的方差。方差计算公式为

$$\sigma_{(y_i, y_j)} = \frac{\sum_{n=1}^N (y_{(n, n_i)} - \bar{y}_{(n, n_i)})(y_{(n, n_j)} - \bar{y}_{(n, n_j)})}{N-1} \quad (3)$$

在分类器链中设置标签阈值,即

$$\text{csign}(\rho) = \begin{cases} 1, \rho > 0.5 \\ 0, \rho \leq 0.5 \end{cases} \quad (4)$$

若 ρ 大于阈值 0.5, 标签 y_i 和标签 y_j 的正相关性较强,即为强正相关标签,将基分类器 FEC_i 的输出结果输入基分类器 FEC_j 的输入空间;否则标签 y_i 和标签 y_j 不相关或弱正相关甚至负相关,基分类器 FEC_i 的输出结果不输入基分类器 FEC_j 的输入空间。

FECC 增强了标签的局部相关性,也减少了分类器冗余,具有更好的分类性能。对于标签间存在相关性的数据集,通过考虑标签间相关性强弱来调节基分类器输入,使其合理地利用标签间的关联信息,减少分类器链的冗余度和误差传播,增强模型的分类能力。图 2 所示为 FECC,该分类器链中共有 8 个 FEC。通过衡量标签间的相关性重新确定 FECC 中的标签集顺序 $y = [y'_1, y'_2, \dots, y'_8]$, 每个基分类器 FEC 的输入为 $X_s = [X, (y'_1, \dots, y'_{s-1})]$, 其中 $1 \leq s \leq 8$ 。例如,若标签 y'_3 与标签 y'_4 的相关性较弱,基分类器 FEC_3 的输出 y'_3 不会作为输入结果送入基分类器 FEC_4 的输入空间,基分类器 FEC_4 的输入空间为 $X_4 = [X, (y'_1, y'_2)]$;若标签 y'_2 和标签 y'_7 的相关性较弱,则基分类器 FEC_2 的输出 y'_2 不能作为输入结果送入基分类器 FEC_7 的输入空间,基分类器 FEC_7 的输入空间为 $X_7 = [X, (y'_1, y'_3, y'_4, y'_5, y'_6)]$ 。

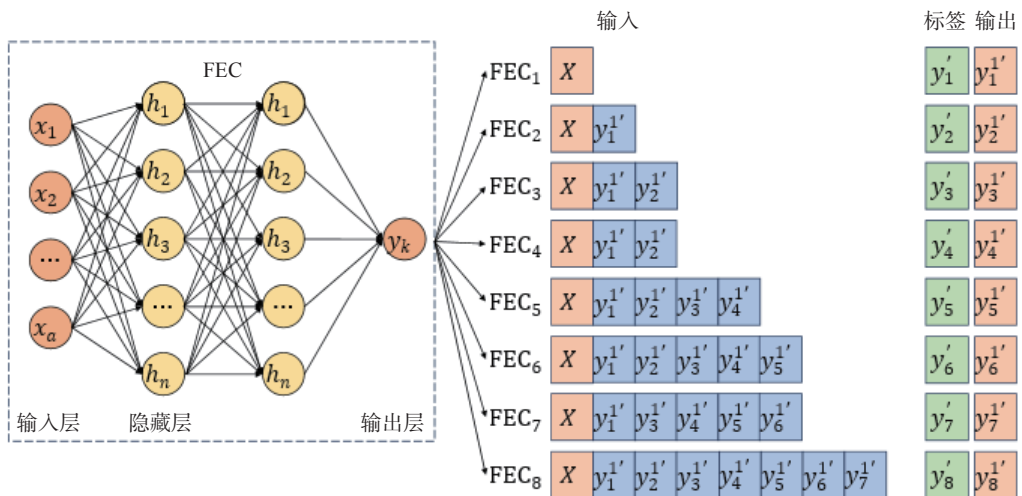


图 2 FECC 结构图

Fig 2 Structure diagram of FECC

FECC: 特征提取分类器链; FEC: 特征提取分类器。 x_a : 第 a 个输入单元; h_n : 第 n 个隐藏层单元; y_k : 第 k 个输出单元; k : 第 k 个标签; y' : 输入到 FEC 中的乱序标签集。

1.3 集成学习融合策略 分类器链能够学习标签间的局部关联信息,对于多标签数据具有较好的分类效果,但不同的标签集顺序也会影响分类器链的分类性能。与单一分类模型相比,集成学习模型结合了多个弱模型预测结果,具有更高的分类能力。典型的集成学习方法包括袋装法 (bagging) 和提升法 (boosting)。袋装法是指对原始数据集有放回的随机抽样,采用投票机制,将每个基学习器分类的结果进行统计,然后根据投票结果确定最终分类标签。提升法为各模型的预测结果赋予不同的权重,加权求和结果作为最终结果。本研究采用集

成学习方法,对 FECC 进行扩展,随机打乱输入到分类器链的标签集顺序,每个标签的分类结果由每条链对应的子标签分类结果投票决定。

图 3 所示为 FEN 模型结构,基本思想是构建 M 条 FECC,每条由 K 个基分类器 FEC 连接构成。其中,第 m 条链的第 k 个基分类器可表示为 FEC_k^m 。由于 RA 中医证候的标签数目为 8,则每条分类器链中有 8 个基分类器 FEC。确定相关性强的标签后,将新的标签集 $Y = [y_{\alpha_1}, y_{\alpha_2}, \dots, y_{\alpha_8}]$ 随机排序后输入 FECC 中,最终每个标签的分类结果由每条链中相应子标签分类结果投票决定。

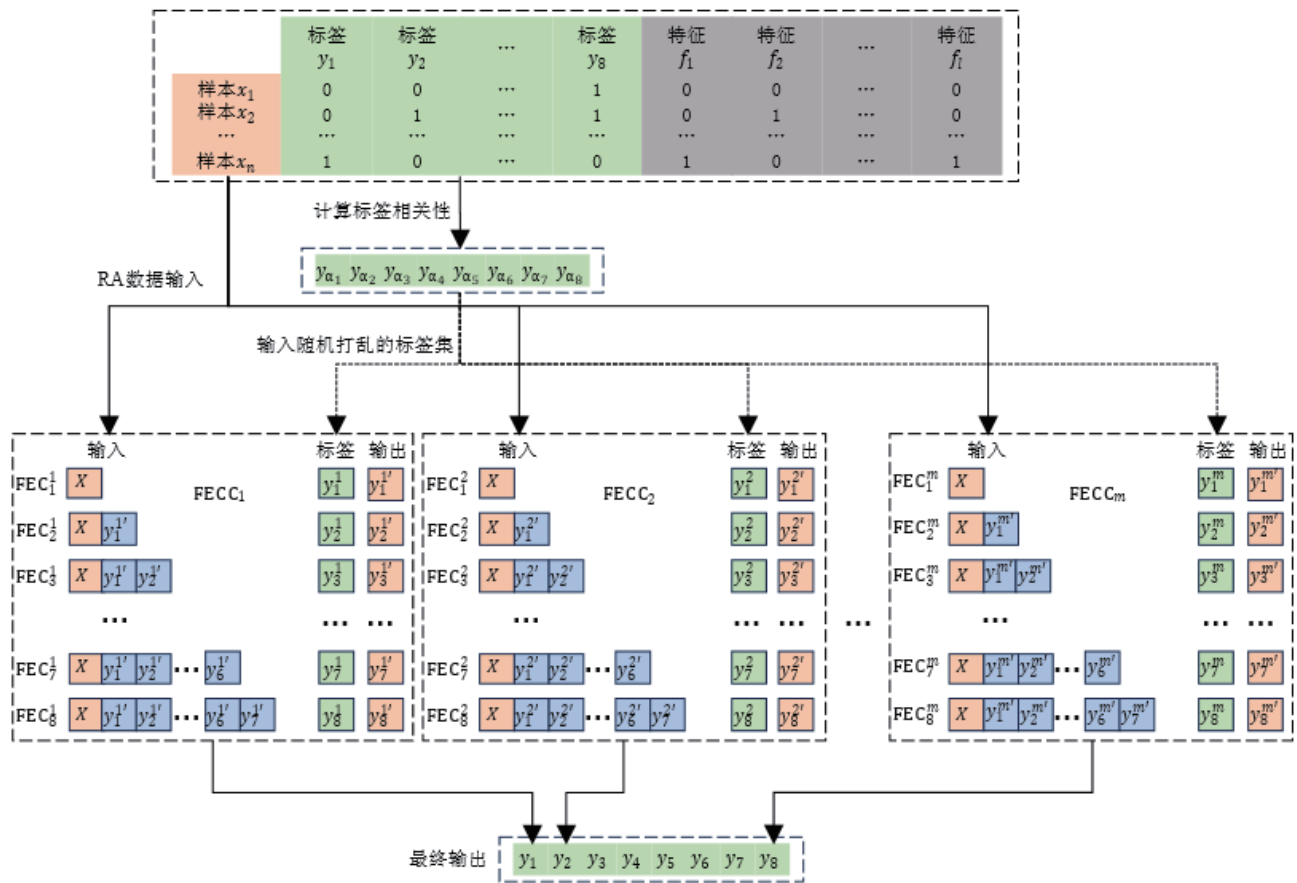


图 3 FEN 模型结构图

Fig 3 Structure diagram of FEN model

FEN: 集成神经网络链; RA: 类风湿关节炎; FEC: 特征提取分类器; FECC: 特征提取分类器链. x_n : 第 n 个样本; f_l : 第 l 个特征; $FECC_m$: 第 m 个 FECC; $FEC_1^m \sim FEC_8^m$: 第 m 个 FECC 中第 1~8 个 FEC; $y_1^m \sim y_8^m$: 第 m 个 FECC 中第 1~8 个 FEC 的对应标签; $y_1^m \sim y_8^m$: 第 m 个 FECC 中第 1~8 个 FEC 的输出标签。

在 FEN 模型中,扩展了 M 条 FECC,每条有 K 个基分类器 FEC。在模型训练中首先通过协方差理论衡量标签间的相关性,确定新的标签集 $Y = [y_{\alpha_1}, y_{\alpha_2}, \dots, y_{\alpha_8}]$ 。将新的标签集随机打乱顺序分别输入到集成分类器链 (ensemble classifier chain,

Ecc) 中。假设第 m 条 FECC 的随机标签集序列为 $Y^m = [y_{\alpha_1}^m, y_{\alpha_2}^m, \dots, y_{\alpha_8}^m]$ 。对于第 m 条 FECC 的第 i 个基分类器的输入为 $X_i^m = [X, (y_{\alpha_1}^m, \dots, y_{\alpha_{i-1}}^m)]$, 对应的函数为 $f_i^m(X^m)$, 输出结果为 $O_{\alpha_i}^m$, 输出结果乘以前一个的输出为

$$P_{\alpha_i}^m = P_{\alpha_{i-1}}^m \cdot O_{\alpha_i}^m \quad (5)$$

预测结果为

$$y_{\alpha_i}^m = \text{csign}(P_{\alpha_i}^m) \quad (6)$$

在模型测试中,对于第 α_k 个标签,其最终输出为 M 个

基分类器输出的加权求和,即

$$y_i = \frac{1}{M} \sum_{i=1}^M y_{\alpha_i}^m \quad (7)$$

最终得到的预测结果为 $Y = [y_1, y_2, \dots, y_k]$ 。

FEN 的详细算法流程如图 4 所示。

模型的训练过程	模型的测试过程
Input: X, Y ; output: FEN for $i=1 \rightarrow K$: train FEC_k with Y for $j=i+1 \rightarrow K$: calculate label correlation: $y_{\alpha_k} = \text{csign}(\rho)$ end for end for determine the label: $Y = [y_{\alpha_1}, y_{\alpha_2}, \dots, y_{\alpha_k}]$ for $m=1 \rightarrow M$ shuffle the label: $Y_{\alpha}^m = [y_{\alpha_1}^m, y_{\alpha_2}^m, \dots, y_{\alpha_k}^m]$ for $i=1 \rightarrow K$: $X_i^m = [X, (y_{\alpha_1}^m, \dots, y_{\alpha_{i-1}}^m)']$ train FEC_i^m with (X_i^m, Y_{α}^m) $O_{\alpha_i}^m = f_i^m(X_i^m)$ $P_{\alpha_i}^m = P_{\alpha_{i-1}}^m \cdot O_{\alpha_i}^m$ $y_{\alpha_i}^m = \text{csign}(P_{\alpha_i}^m)$ end for end for end for	Input: X, FEN ; output: Y for $i=1 \rightarrow K$: train FEC_k with Y for $j=i+1 \rightarrow K$: calculate label correlation: $y_{\alpha_k} = \text{csign}(\rho)$ end for end for determine the label: $Y = [y_{\alpha_1}, y_{\alpha_2}, \dots, y_{\alpha_k}]$ for $m=1 \rightarrow M$ shuffle the label: $Y^m = [y_{\alpha_1}^m, y_{\alpha_2}^m, \dots, y_{\alpha_k}^m]$ for $i=1 \rightarrow K$: $X_i^m = [X, (y_{\alpha_1}^m, \dots, y_{\alpha_{i-1}}^m)']$ $O_{\alpha_i}^m = f_i^m(X_i^m)$ $P_{\alpha_i}^m = P_{\alpha_{i-1}}^m \cdot O_{\alpha_i}^m$ $y_{\alpha_i}^m = \text{csign}(P_{\alpha_i}^m)$ end for end for for $i=1 \rightarrow K$: $y_i = \frac{1}{M} \sum_{i=1}^M y_{\alpha_i}^m$ end for $Y = [y_1, y_2, \dots, y_k]$

图 4 FEN 的算法流程

Fig 4 Algorithm flow of FEN

FEN:集成神经网络链;FEC:特征提取分类器.

2 实验结果与分析

2.1 RA 中医证候数据集及预处理 所采用的实验数据来自中国中医科学院广安门医院 1 683 例 RA 患者的临床数据集。删除存在缺失值和错误值的病例,去除部分冗余信息如患者社会学资料、实验室指标等,保留了 11 种 RA 中医证候特征,分别为主证、兼证、主证-主症、主证-次症、兼证-主症、兼证-次症、脉象、舌色、舌形、苔质、苔色。主证是 RA 的主要病症,兼证是与主证在同一时间出现的兼杂病症。主证和兼证均含有相同的 8 种 RA 中医证候,分别为风湿痹阻证、寒湿痹阻证、肝肾不足证、气血两虚证、气阴两虚证、湿热痹阻证、痰瘀痹阻证、瘀血阻络证。对于每例 RA

患者既有主证,也可能具有兼证,即每例 RA 患者具有 1~2 个 RA 中医证候,因此 RA 中医证候分类属于多标签分类。图 5 所示为 1 683 例患者不同主证的数量及不同主证下兼证所占比例,其中湿热痹阻证样本最多(511 例),气阴两虚证最少(69 例);主证为瘀血阻络证的样本中无其他兼证者占 46% (38/82),既有瘀血阻络证又有兼证风湿痹阻证者占 27% (22/82)。

另外 9 种(主证-主症、主证-次症、兼证-主症、兼证-次症、脉象、舌色、舌形、苔质、苔色) RA 中医证候特征则是 RA 中医证候的常见症状,本研究将这 9 种特征作为 RA 中医证候分类特征,如表 1 所示。

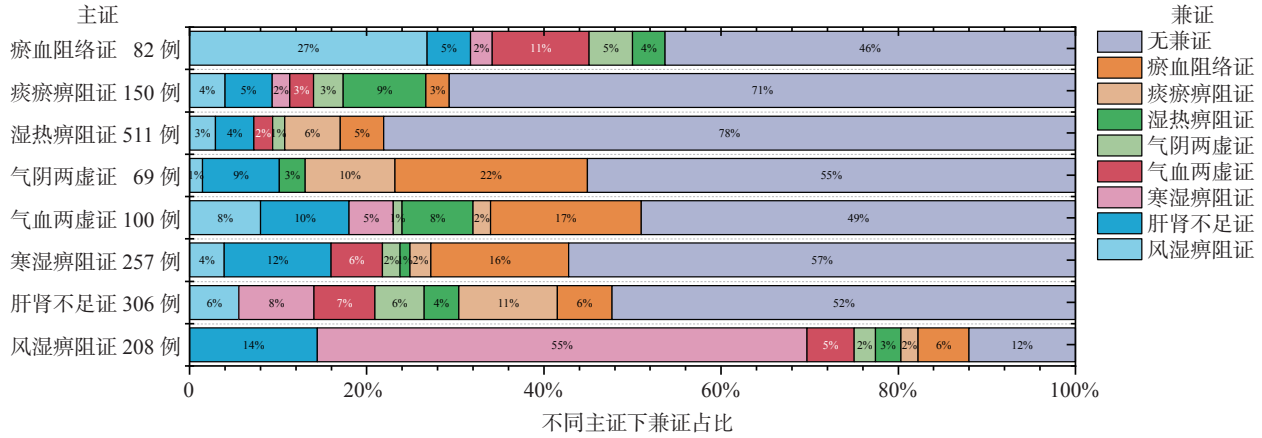


图5 1683例类风湿关节炎(RA)患者不同主证、兼证分布

Fig 5 Distribution of different main and accompanying syndromes in 1 683 rheumatoid arthritis (RA) patients

表1 9种类风湿关节炎(RA)中医证候特征描述

Tab 1 Feature description of 9 kinds of traditional Chinese medicine syndromes of rheumatoid arthritis (RA)

特征	特征条目	特征条目数
主证-主症	关节疼痛、肿大或僵硬变形、腰膝酸软或腰背酸痛、关节肿痛久不消、疼痛夜甚等	27
主证-次症	关节僵硬变形、关节局部皮色发红、口渴或渴不欲饮、关节肌肉刺痛、发热等	114
兼证-主症	关节疼痛、肿大或僵硬变形、腰膝酸软或腰背酸痛、关节肿大伴气短乏力等	25
兼证-次症	关节僵硬变形、关节局部皮色发红、口渴或渴不欲饮、关节局部色暗等	112
脉象	浮、滑、细、缓、涩、弦、数、结等	86
舌色	暗、淡白、淡红、淡白、绛、紫、淡红等	30
舌形	娇嫩、瘦薄、胖大、芒刺、裂纹、齿痕、正常等	44
苔质	薄、腻、燥、腐、滑、润、厚、剥等	49
苔色	黑、白、黄、灰	4

从表1中可以看出每个特征包括不同数量的特征条目,且这些特征条目均为离散型。为了保证结果的可靠性,对于轻微缺失值进行补零处理,并删除具有较多缺失值的样本。本研究采用 one-hot 编码方式将 RA 中医证候分类特征转换为数字特征,共 491 个数据特征。因此,经过处理后的数据样本包含 8 个标签和 491 个特征。使用 one-hot 编码将每个病例的标签和特征使用 0、1 的方式表示,0 表示该病例没有该证候或特征,1 表示该病例有该证候或特征。

2.2 评价指标 采用 3 种损失指标和 3 种性能指

标评价模型的分类效果。3 种损失指标分别为汉明损失 (Hamming loss)、1- 错误率 (one-error) 和排名损失 (ranking loss), 损失指标的值越小表明模型的分类效果越好; 3 种性能指标分别为准确度 (accuracy)、召回率 (recall) 和 F1 值 (F1-score), 性能指标的值越大表明模型的分类效果越好。此外,还使用平均精度 (average precision) 作为多标签分类结果的评价指标,其值越大表明模型性能越好; 使用几何平均值 (geometric mean, G-mean) 衡量模型的综合表现。各指标定义为

$$\text{汉明损失} = \frac{1}{n} \sum_{i=1}^n \frac{1}{m} |Y_i \Delta Z_i| \tag{8}$$

$$\text{1- 错误率} = \frac{1}{n} \sum_{i=1}^n \delta(\text{argmin} \gamma_i(\lambda)) \tag{9}$$

$$\text{排名损失} = \frac{1}{n} \sum_{i=1}^n \frac{1}{|Y_i|} |\{(\lambda_a, \lambda_b) : \gamma_i(\lambda_a) > \gamma_i(\lambda_b), (\lambda_a, \lambda_b) \in Y_i \times \bar{Y}_i\}| \tag{10}$$

$$\text{准确度} = \frac{\text{真正例数} + \text{真负例数}}{\text{真正例数} + \text{真负例数} + \text{假正例数} + \text{假负例数}} \times 100\% \tag{11}$$

$$\text{召回率} = \frac{\text{真正例数}}{\text{真正例数} + \text{假负例数}} \times 100\% \tag{12}$$

$$F1 \text{ 值} = 2 \times \frac{\text{精确率} \times \text{召回率}}{\text{精确率} + \text{召回率}} \times 100\% \quad (13)$$

$$\text{精确率} = \frac{\text{真正例数}}{\text{真正例数} + \text{假正例数}} \times 100\% \quad (14)$$

$$\text{平均精度} = \frac{1}{n} \sum_{i=1}^n \frac{1}{|Y_i|} \sum_{\lambda \in Y_i} \frac{|\lambda' \in Y_i: \gamma_i(\lambda') \leq \gamma_i(\lambda)|}{\gamma_i(\lambda)} \quad (15)$$

$$\text{灵敏度} = \frac{\text{真正例数}}{\text{真正例数} + \text{假负例数}} \quad (16)$$

$$G\text{-mean} = \sqrt{\text{灵敏度} \times \text{特异度}} \quad (17)$$

$$\text{特异度} = \frac{\text{真负例数}}{\text{真负例数} + \text{假正例数}} \quad (18)$$

其中 n 为样本数, 式 (8) 中 Y_i 是第 i 个样本的真实标签集合, ΔZ_i 表示第 i 个样本的预测值 Z 的偏移量, δ 表示指示函数, $\operatorname{argmin}_\lambda \gamma_i(\lambda)$ 表示在给定样本 i 的相关度评分 $\gamma_i(\lambda)$ 中取得最小值的 λ 的索引, 式 (10) 和 (15) 中的 Y_i 是第 i 个样本的相关样本集合, \bar{Y}_i 是第 i 个样本的不相关样本集合, $|Y_i|$ 和 $|\bar{Y}_i|$ 分别是它们的元素个数, λ_a 和 λ_b 表示第 i 个样本与其他样本的关系, a 、 b 代表样本在序列中的排列顺序, $\gamma_i(\lambda)$ 是样本 i 与 λ 的相关度评分。

2.3 消融实验

2.3.1 FEC 网络隐藏层层数对多标签分类的影响

当网络隐藏层的层数较少时, 模型无法学习到数据的深层特征, 导致模型各项评价指标均较差。如果网络隐藏层的层数过多则会导致模型过拟合, 因此需要选择合适的层数以提高模型的精度和泛化性能。图 6 所示为不同隐藏层层数对模型性能的影响结果。当隐藏层层数为 2 时, 汉明损失、1- 错误率和排名损失均最低, 准确度、召回率和 F1 值均最高, 故将该隐藏层的层数作为模型 FEC 网络参数。

在神经网络中, 隐藏层节点数对模型的性能影响很大, 不仅与输入、输出层的节点数有关, 更与需要解决问题的复杂程度和样本数据的特性有关。

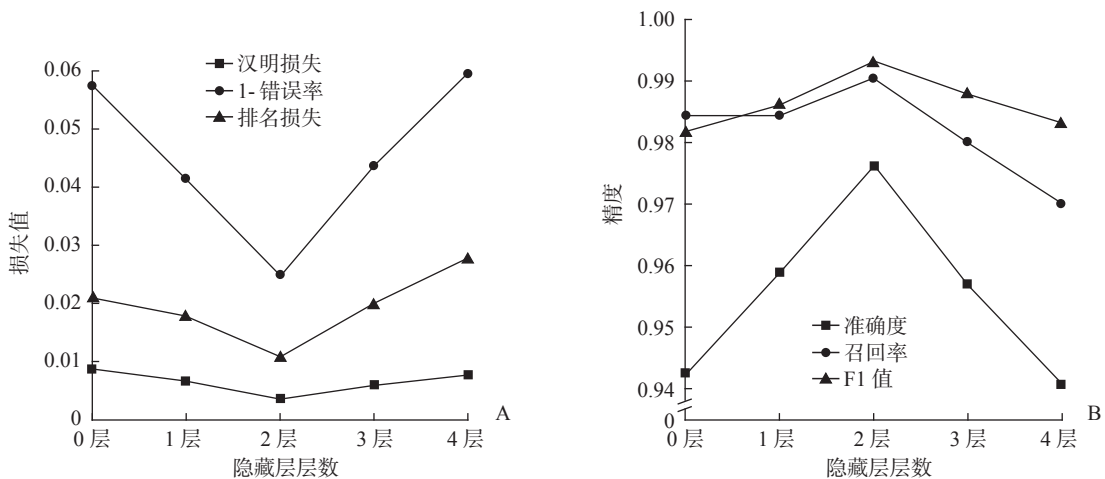


图 6 不同隐藏层层数的模型性能比较

Fig 6 Comparison of model performance with different hidden layers

A: 损失指标; B: 性能指标.

1 683 例 RA 样本按照 7 : 3 的比例划分为训练集 (1 178 例) 和测试集 (505 例), 为了使隐藏层节点数接近训练样本数, 将隐藏层节点数设置为 1 024, 与其他具有不同隐藏层节点数的模型的性能对比结果见图 7。当 2 层隐藏层节点数均设置为 1 024 时, 模型的损失指标和性能指标均最优, 可见当节点数接近训练样本数时模型的性能更好。

2.3.2 FECC 标签相关性对多标签分类的影响

对 RA 多标签样本进行相关性分析, 引入属性协方差描述 RA 中中医证候任意 2 个标签之间的相关性, 图 8 展示了 8 个标签协方差分数。矩阵中的分数为正值表明 2 个标签为正相关, 分数值越高说明相关性越强; 分数为负值表明 2 个标签为负相关, 分数值越小说明其反向程度越大。分析可知, 痰瘀痹阻证与肝肾不足证之间具有最强的正相关性, 而瘀血阻滞证与肝肾不足证之间具有最强的负相关性。

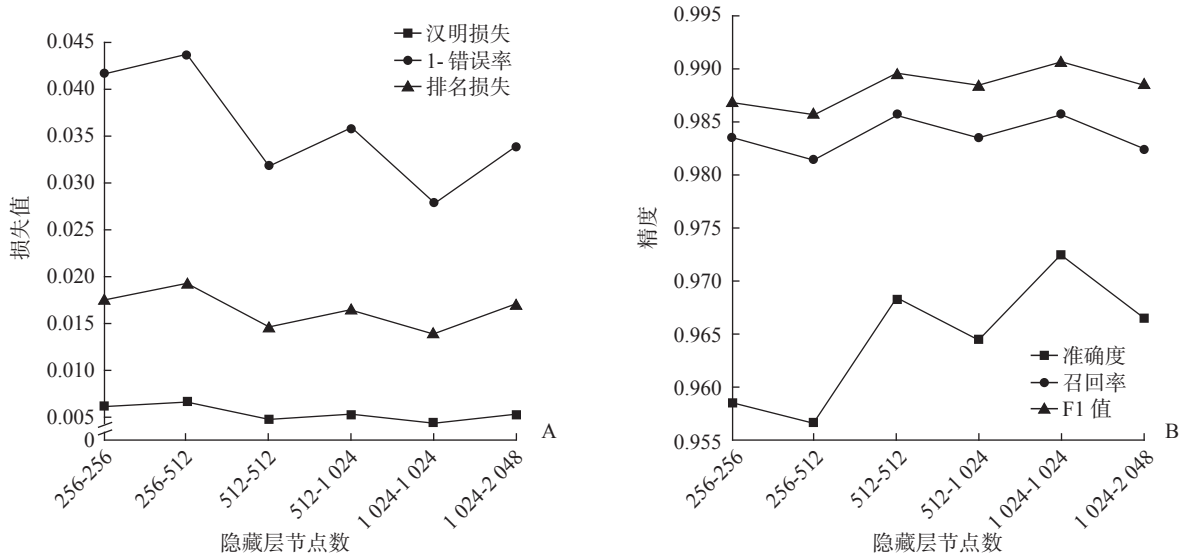


图7 不同隐藏层节点数的模型性能比较

Fig 7 Comparison of performance among models with different number of nodes in hidden layers

A: 损失指标; B: 性能指标.

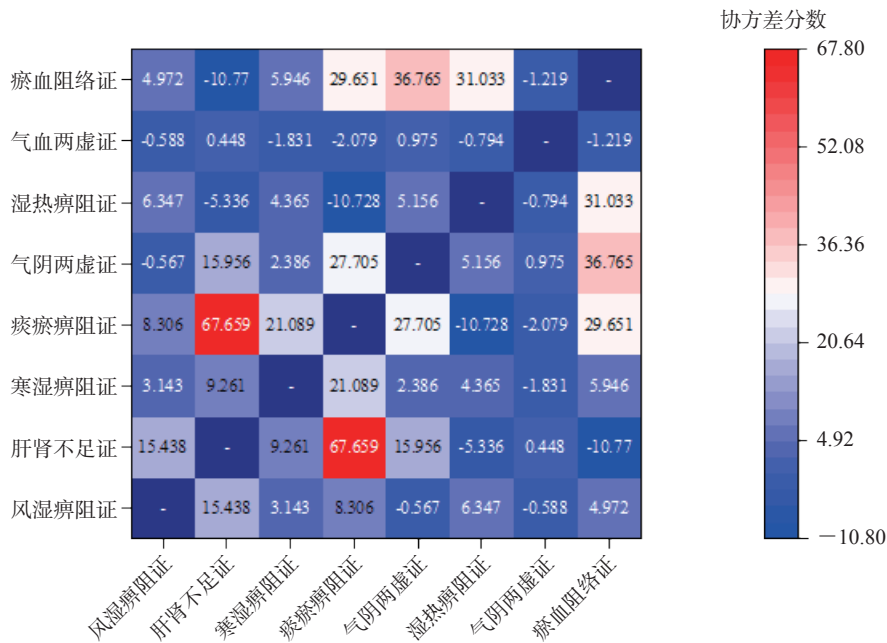


图8 8个标签协方差矩阵

Fig 8 Covariance matrix of 8 labels

RA 中医证候真实值与模型预测值对比结果如表 2 所示。病例 1 和病例 2 均患有痰瘀痹阻证和肝肾不足证，由于这 2 个标签之间具有最强的正相关性，对应的基分类器输入空间都能相互接收彼此的信息，FEN 和分类器链都能准确预测其真实值。病例 5 和病例 6 均患有瘀血阻络证和肝肾不足证，2 个标签之间具有最强的负相关性，对于这 2 个病例，FEN 在模型训练时，2 个标签对应的基分类器

输入空间不会接收彼此的信息，避免了噪声干扰和错误传播，可以正确预测其真实值；而分类器链由于负相关标签的干扰，不能同时正确预测出肝肾不足证和瘀血阻络证，对病例 5 只能预测出肝肾不足证，对病例 6 只能预测出瘀血阻络证。通过衡量标签相关性的方法可以避免噪声干扰和误差传播，有效地提高模型的预测效果。

表 2 RA 中医证候真实值与模型预测值对比结果

Tab 2 Comparison of actual and predicted values of traditional Chinese medicine syndromes of RA

病例	真实值			FEN 预测值			分类器链预测值		
	痰瘀痹阻证	肝肾不足证	瘀血阻络证	痰瘀痹阻证	肝肾不足证	瘀血阻络证	痰瘀痹阻证	肝肾不足证	瘀血阻络证
1	1	1	0	1	1	0	1	1	0
2	1	1	0	1	1	0	1	1	0
3	0	1	0	0	1	0	0	1	0
4	0	0	1	0	0	1	0	0	1
5	0	1	1	0	1	1	0	1	0
6	0	1	1	0	1	1	0	0	1

“0”表示病例没有该证候，“1”表示病例有该证候。RA：类风湿关节炎；FEN：集成神经网络。

2.3.3 FEN 中分类器链数量对多标签分类的影响 FEN 模型中分类器链的数目 M 选取值会对分类结果产生影响。若 M 值过小，会减弱集成效果，不能有效减少不合理标签序列对分类结果的干扰；若 M 值过大，则使得模型过大，导致参数量较多，训练时间较长。根据经验值，本研究将 M 值设置为 8，图 9A 展示了 M 值为 8 时每条分类器链 FECC 随着训练次数增加损失值的变化曲线，图 9B、9C 分别展示了 FEN 模型在不同阈值下针对 RA 中医证候样本

多标签分类的平均精度和 F1 值。平均精度是一种基于标签排序的评估指标，在计算分类精度时考虑了标签间的相关性，能更好地衡量多标签分类器的分类性能。其中，FEN 模型在标签阈值设置为 0.5 时平均精度达到最大值，为 0.984 5。F1 值作为精确率和召回率的调和平均数，可以精确地评价模型效果的优劣，当标签阈值设置为 0.5 时模型的效果达到最优，F1 值为 0.990 5。

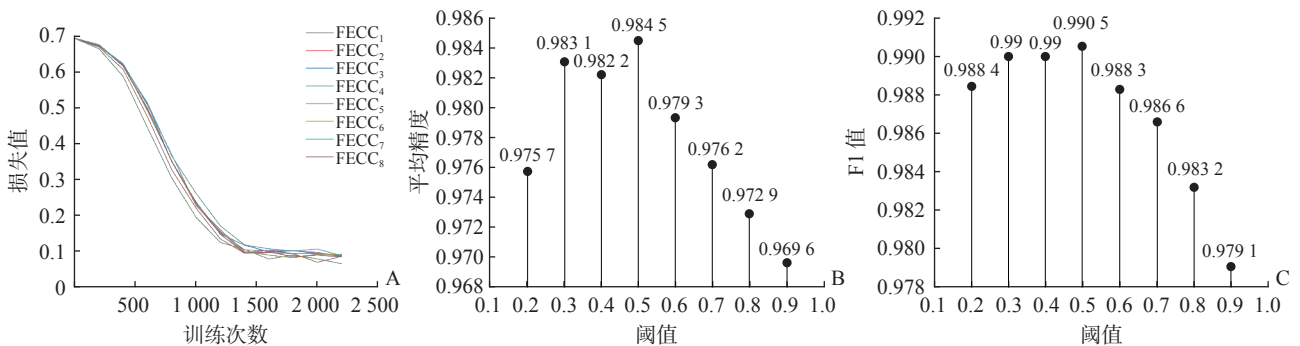


图 9 FEN 模型的损失曲线及不同阈值下的平均精度和 F1 值

Fig 9 Loss curves of FEN model and average precision and F1-scores under different thresholds

A:损失曲线;B:平均精度;C:F1 值. FEN:集成神经网络链;FECC:特征提取分类器链.

2.4 交叉验证实验结果 表 3 所示为 FEN 模型 RA 中医证候多标签分类交叉验证结果。采用 10 折交叉验证法，将 1 683 例样本数据分成 10 份，然后进行 10 次实验，每次实验都从这 10 份数据中选取其中 9 份作为训练集来训练模型，其余 1 份作为测试集（168 例）用于评估模型的性能。分析可知，各项评价指标的标准差均小于 0.003，实验过程中数据波动较小，每轮实验所用的数据均服从同一分布。汉明损失、1-错误率和排名损失分别为 0.003 6、0.024 8 和 0.010 9，准确度、召回率和 F1 值分别为 97.52%、98.92% 和 99.18%，说明 FEN 对 RA 中医

证候样本的多标签分类具有较高的分类精度和较好的泛化性能。

2.5 各种多标签分类器分类效果比较 FEN 与其他 7 种典型多标签分类器 [分类器链、LP、BR、RakEL、ML-KNN、Ecc 和集成二进制关联 (ensemble binary relevance, Ebr)] 对 168 例测试集样本 RA 中医证候的分类效果对比结果见表 4。FEN 的 3 种损失指标汉明损失、1-错误率和排名损失分别为 0.003 6、0.024 8 和 0.010 9，均低于其他 7 种分类模型；FEN 的 3 种性能指标准确度、召回率和 F1 值分别为 97.52%、98.92% 和 99.18%，均高于其

他7种分类模型,可见FEN与其他多标签分类器相比具有较高的分类精度和泛化性能。这是由于基于问题转换的多标签分类模型(BR、分类器链)将多标签分类转换为多个单标签分类,未能合理利用RA中医证候标签之间的关联性,降低了RA样本的分类精度;基于算法适应的多标签分类模型(ML-KNN、RakEL、LP)通过扩展单标签方法实现多标签分类,无法有效提取具有强关联特征的RA中医证候样本深层特征,使其分类性能较低;而集成学习方法(Ecc、Ebr)则通过集成多个单分

类器仅学习部分标签之间的相关性,使分类精度降低。本研究FEN模型采用DNN结构构建基分类器FEC提取RA中医证候特征,通过计算不同标签的协方差分数划分标签的强弱关系,并采用链式结构FECC将标签输入到与之具有强相关性的标签基分类器中,充分利用了标签间的关联性,减少了数据冗余度和误差传播,而且基于集成学习方法分类器链FECC结构,减少了不合理标签序列的干扰,提高了模型的泛化性能。

表3 1 683例RA中医证候分类10折交叉验证结果

Tab 3 10-fold cross-validation for classification of traditional Chinese medicine syndromes of 1 683 RA patients

交叉验证	汉明损失	1-错误率	排名损失	准确度/%	召回率/%	F1值/%
1折	0.004 2	0.029 7	0.013 9	97.03	98.54	99.04
2折	0.003 7	0.023 8	0.009 7	97.60	99.11	99.16
3折	0.003 0	0.021 8	0.009 7	97.82	98.99	99.32
4折	0.003 2	0.019 8	0.009 1	98.02	99.11	99.27
5折	0.003 2	0.023 8	0.011 4	97.62	98.78	99.27
6折	0.004 0	0.025 7	0.012 0	97.43	98.77	99.15
7折	0.004 5	0.027 7	0.011 7	97.23	98.88	98.99
8折	0.003 2	0.023 8	0.010 7	97.62	98.89	99.27
9折	0.004 0	0.027 7	0.010 9	97.23	99.00	99.11
10折	0.003 5	0.023 8	0.009 6	97.62	99.10	99.22
平均值, $\bar{x} \pm s$	0.003 6 ± 0.000 5	0.024 8 ± 0.002 8	0.010 9 ± 0.001 4	97.52 ± 0.28	98.92 ± 0.17	99.18 ± 0.11

RA:类风湿关节炎.

表4 各多标签分类器对RA中医证候的分类结果

Tab 4 Classification results of traditional Chinese medicine syndromes of RA by each multi-label classifier

模型	汉明损失	1-错误率	排名损失	准确度/%	召回率/%	<i>n</i> = 168, $\bar{x} \pm s$
						F1值/%
CC	0.006 9 ± 0.000 7	0.048 1 ± 0.004 2	0.024 3 ± 0.002 3	95.19 ± 0.42	97.34 ± 0.27	98.42 ± 0.16
LP	0.018 3 ± 0.003 9	0.073 9 ± 0.014 2	0.043 8 ± 0.008 8	92.61 ± 1.42	95.66 ± 0.87	95.87 ± 0.90
BR	0.007 7 ± 0.000 5	0.057 2 ± 0.004 3	0.029 0 ± 0.002 2	94.28 ± 0.43	96.79 ± 0.25	98.25 ± 0.12
RakEL	0.009 9 ± 0.001 1	0.040 8 ± 0.004 5	0.023 0 ± 0.002 7	95.92 ± 0.45	97.73 ± 0.28	97.75 ± 0.25
ML-KNN	0.032 5 ± 0.002 0	0.165 3 ± 0.010 6	0.093 6 ± 0.004 3	83.47 ± 1.06	90.03 ± 0.53	92.38 ± 0.48
Ecc	0.006 3 ± 0.000 6	0.042 2 ± 0.003 5	0.021 9 ± 0.002 0	95.77 ± 0.35	97.58 ± 0.22	98.54 ± 0.16
Ebr	0.006 8 ± 0.000 6	0.052 8 ± 0.004 7	0.026 5 ± 0.002 6	94.71 ± 0.47	97.04 ± 0.30	98.42 ± 0.15
FEN	0.003 6 ± 0.000 5	0.024 8 ± 0.002 8	0.010 9 ± 0.001 4	97.52 ± 0.28	98.92 ± 0.17	99.18 ± 0.11

RA:类风湿关节炎;CC:分类器链;LP:标签幂集;BR:二进制关联;RakEL:随机k-标签集;ML-KNN:多标签K最近邻;Ecc:集成分类器链;Ebr:集成二进制关联;FEN:集成神经网络链.

图10为505例测试集样本中,各分类器对各子标签在灵敏度和G-mean上的分类比较结果。灵敏度衡量了模型对真正例的识别能力,本研究FEN模型判断寒湿痹阻证、湿热痹阻证、痰瘀痹阻证、淤血阻络证和风湿痹阻证的灵敏度均达到1.0,

FEN能够有效地捕获并识别这些真正例。G-mean综合考虑了分类器在正负2个类别上的性能,对每个子标签的G-mean进行统计发现FEN对8个子标签的G-mean均超过0.99,寒湿痹阻证、湿热痹阻证、痰瘀痹阻证、瘀血阻络证和风湿痹阻证的

G-mean 均达到 1.00。FEN 在标签输入时避免了弱相关和负相关标签的干扰，充分利用了标签间的关联性，减少数据冗余度和误差传播，对于 2 种评估指标，FEN 均有较好的表现。而其他多标签分类器如 BR 和 Ebr 在处理高度关联标签时可能表现较差，在预测时无法正确捕捉到标签之间的关联性，从而影响灵敏度和 G-mean；分类器链会按照标签顺序逐个预测每个标签，如果标签的顺序选择不当，可能无法准确地捕捉到标签之间的依赖关系；RakEL 是在预测阶段随机选择一部分标签进行预

测，因此可能无法充分利用所有可用的标签信息；LP 的性能高度依赖于标签组合的质量和覆盖率，如果标签组合不充分或质量较低，会影响灵敏度和 G-mean；ML-KNN 在计算标签邻居时，对样本的密度敏感，如果不同标签类别的样本密度差异较大，可能会影响预测结果，导致灵敏度和 G-mean 较低；Ecc 通过集成多个分类器链进行预测，不同标签之间容易受到干扰，如果分类器链之间存在较高的耦合度，可能导致模型整体性能下降，影响灵敏度和 G-mean。

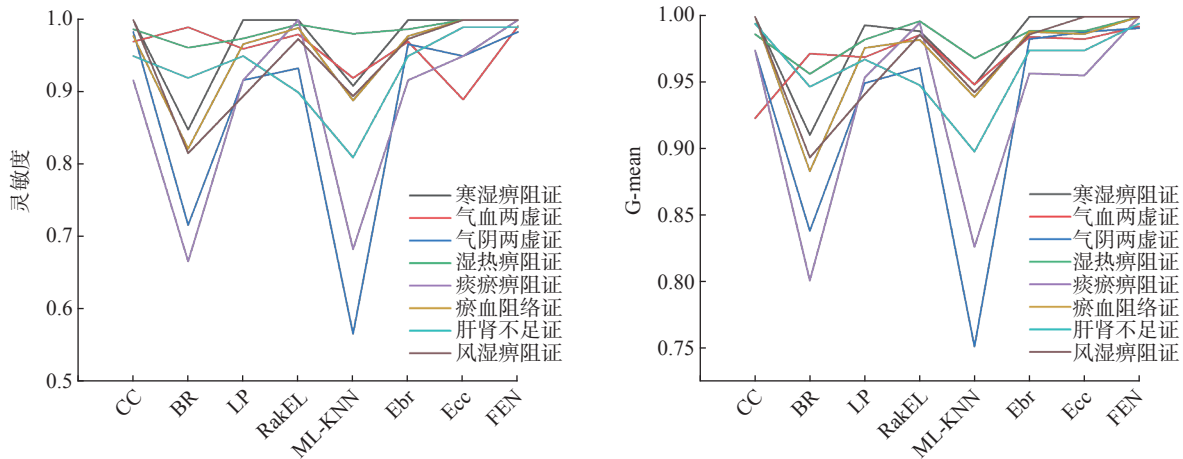


图 10 各多标签分类器对 RA 中医证候标签分类对比结果

Fig 10 Comparison results of each multi-label classifier for classification of traditional Chinese medicine syndrome labels of RA

RA: 类风湿关节炎; CC: 分类器链; BR: 二进制关联; LP: 标签幂集; RakEL: 随机k-标签集; ML-KNN: 多标签K最近邻; Ebr: 集成二进制关联; Ecc: 集成分类器链; FEN: 集成神经网络链; G-mean: 几何平均值。

图 11 所示为在 505 例测试集样本中不同分类器对每个标签分类的混淆矩阵。LP、BR 模型基于单标签分类器实现多标签分类，忽略了标签之间的关联性，导致证候标签预测错误，如 LP 在痰瘀痹阻证、风湿痹阻证、气血两虚证和瘀血阻络证上有 9 例或以上预测错误，BR 在气阴两虚证和瘀血阻络证上均有 7 例预测错误；分类器链、Ecc 模型虽为集成学习策略，但忽略了标签之间的冗余，导致部分证候标签出现偏差，如分类器链在痰瘀痹阻证上有 6 例预测错误、瘀血阻络证上有 7 例预测错误，Ecc 在湿热痹阻证上有 6 例预测错误、风湿痹阻证上有 4 例预测错误。表 5 列出了不同多标签分类器对 8 个标签分类的 F1 值，由于分类器链、BR、Ecc、Ebr 采用相关性学习和 / 或集成学习，在 8 个标签上均获得了较高的 F1 值；本研究模型 FEN 基于协方差理论的相关性并结合集成学习，相较于其

他 7 个多标签分类器的 F1 值均值，在 8 个标签寒湿痹阻证、气血两虚证、气阴两虚证、湿热痹阻证、痰瘀痹阻证、瘀血阻络证、肝肾不足证和风湿痹阻证上的 F1 值分别提高了 1.34%、1.21%、3.64%、1.28%、4.70%、3.19%、2.35% 和 2.73%。FEN 在 8 个标签中以肝肾不足证、湿热痹阻证和痰瘀痹阻证的 F1 值较高，根据图 8 的协方差矩阵可知肝肾不足证和湿热痹阻证之间存在负相关，FEN 解决了具有负相关的标签之间的干扰，在这 2 个标签上的 F1 值最高，同时肝肾不足证和痰瘀痹阻证之间具有最强的正相关性，使在痰瘀痹阻证上也获得了较高的 F1 值。因此，本研究模型 FEN 对各标签分类的 F1 值优于其他多标签分类器，且能够合理地利用标签间的相关性，有效地提高了对每个标签分类的 F1 值。

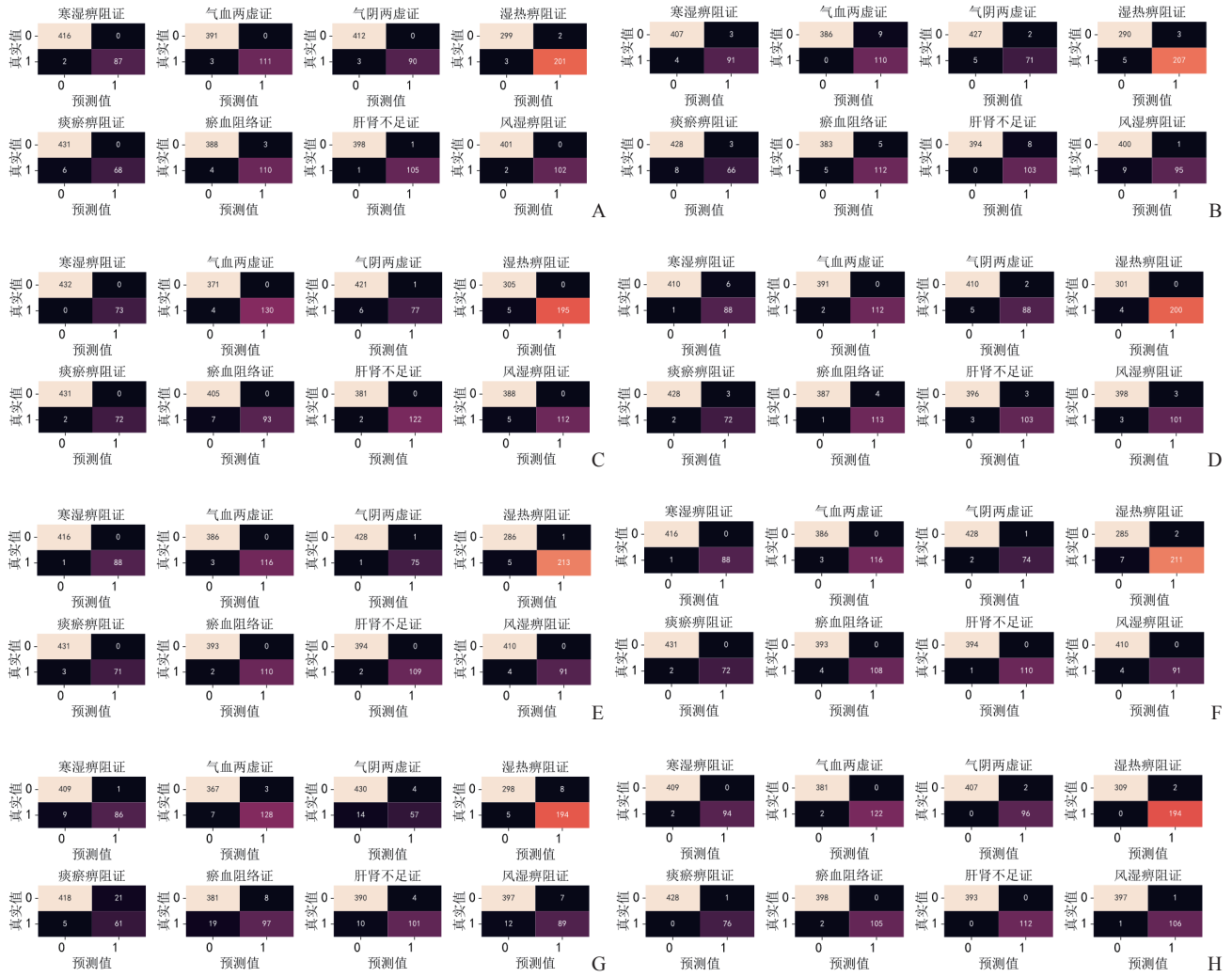


图 11 各分类器针对 RA 中医证候分类的混淆矩阵

Fig 11 Confusion matrix of each classifier for traditional Chinese medicine syndrome classification of RA

A: 分类器链; B: 标签募集; C: 二进制关联; D: 随机k-标签集; E: 集成分类器链; F: 集成二进制关联; G: 多标签K最近邻; H: 集成神经网络链. “0”表示没有该证候; “1”表示有该证候. RA: 类风湿关节炎.

表 5 各分类器针对 RA 中医证候分类的 F1 值

Tab 5 F1-scores of each classifier for traditional Chinese medicine syndrome classification of RA

中医证候	CC	LP	BR	RakEL	Ecc	Ebr	ML-KNN	FEN
寒湿痹阻证	0.988 9	0.963 0	1.000 0	0.948 5	0.994 5	0.994 5	0.945 1	0.989 4
气血两虚证	0.986 8	0.960 5	0.984 8	0.990 9	0.987 3	0.987 3	0.962 3	0.991 9
气阴两虚证	0.983 7	0.953 1	0.956 1	0.961 7	0.987 0	0.980 5	0.863 6	0.989 9
湿热痹阻证	0.987 5	0.981 0	0.974 4	0.989 9	0.985 9	0.991 0	0.967 4	0.995 0
痰瘀痹阻证	0.957 8	0.923 4	0.986 3	0.988 4	0.979 1	0.983 0	0.824 3	0.993 5
瘀血阻络证	0.969 0	0.957 0	0.963 7	0.978 3	0.990 9	0.981 7	0.877 8	0.990 4
肝肾不足证	0.991 0	0.962 7	0.991 9	0.972 0	0.990 9	0.995 5	0.935 3	1.000 0
风湿痹阻证	0.990 4	0.951 0	0.980 1	0.971 0	0.978 5	0.978 5	0.903 4	0.991 0

RA: 类风湿关节炎; CC: 分类器链; LP: 标签募集; BR: 二进制关联; RakEL: 随机k-标签集; Ecc: 集成分类器链; Ebr: 集成二进制关联; ML-KNN: 多标签K最近邻; FEN: 集成神经网络链.

3 算法复杂度与特征贡献度分析

3.1 算法复杂度 对于基于问题转换法的多标签分类任务计算成本包括基本分类模型和问题转换

结构。其中，不同的问题转换结构可以选择不同的基分类器。本研究将 $F_y()$ 定义为一个基分类器 y 的时间复杂度。假设多标签分类数据集有 x 个实例，每个实例由 n 个维度和 k 个标签组成。对于 BR 模

型只需计算每个标签 k 次,因此 BR 模型的时间复杂度为 $O(kF_y(x,n))$ 。分类器链保留了 BR 的低复杂度,并扩展了特征空间提取标签相关性,因此分类器链的时间复杂度为 $O(kF_y(x,n+k))$ 。Ecc、Ebr 则分别生成 m 个分类器链和 BR,其时间复杂度分别为 $O(mkF_y(x,n+k))$ 、 $O(mkF_y(x,n))$ 。本研究模型 FEN 定义了基分类器 FEC,并生成链式结构分类器链 FECC,再采用集成学习扩展 FECC,其保持了与 Ecc 同样的时间复杂度,即 $O(mkF_y(x,n+k))$,但获得了更好的分类精度和泛化性能。

3.2 特征贡献度 主证和兼证的特征贡献度分析对于 RA 临床诊断和治疗具有重要的参考价值,本研究针对 1 683 例样本分析了 RA 中医证候分类中主证和兼证的特征贡献度。主证和兼证中均是主证的贡献度最高,分别为 46.0% 和 50.8%,该特征对于 RA 中医证候的诊断贡献度占比超过了 45%,在 RA 中医证候的诊断中可以作为良好的诊断指标。其次是次证的贡献度占比,分别达到了 17.7% 和 26.4%,脉象和舌诊(苔色、舌色、苔质、舌形)的贡献度占比有所不同,主证中的脉象、苔色、舌色、苔质和舌形的贡献度占比分别为 11.4%、11.1%、7.8%、4.2% 和 1.8%,兼证中的舌色、脉象、苔色、苔质和舌形的贡献度占比分别为 7.0%、6.9%、3.4%、2.8% 和 2.7%。由此可见,主证和次证特征均可作为判断 RA 中医证候分类的重要指标,这也与临床 RA 诊断经验相符。因为中医学提出一种“疾病-证候-症状”关联网络以探索 RA 中医证候的生物学内涵^[1],证候差异表达基因与其症状相关基因存在交集,因此主证、次证成为影响主证和兼证分类的主要因素。

4 总结

本研究提出了一种基于集成神经网络的模型 FEN,实现临床 RA 中医证候多标签分类。该模型构建基于 DNN 的基分类器提取临床 RA 样本的深层特征;根据协方差理论衡量标签相关性,调节分类器链的输入标签集,减少负相关或弱相关标签的干扰,合理学习局部标签的关联信息;采用集成学习方法构建多条分类器链,为每条链随机输入不同标签序列,减少不合理标签对分类的影响,提高了模型的分类精度和泛化性能。本实验结果表明,与其他多标签分类方法相比,FEN 模型具有较好的分

类精度和泛化性能。由于 FEN 模型采用了 DNN 和集成学习,与传统多标签分类模型相比,需要更多参数和更长训练时间,后期将重点研究如何减少模型参数量和训练时间。此外,本研究还分析了影响临床 RA 中医证候分类的特征贡献度,主证、次证是影响主证和兼证分类的主要因素,这为 RA 中医证候临床诊断提供了参考。

[参考文献]

- [1] 陈文佳,巩勋,刘蔚翔,等.从“病-证-症”关联网络探索类风湿关节炎中医证候的生物内涵[J].中国中药杂志,2022,47(3):796-806. DOI: 10.19540/j.cnki.cjcmm.20211105.501.
- [2] 王秀芳,赵衍振,樊冰,等.早期类风湿关节炎患者的临床特点与中医证候分布规律研究[J].风湿病与关节炎,2021,10(8):6-9. DOI: 10.3969/j.issn.2095-4174.2021.08.002.
- [3] 柳玉佳,王莘智,旷惠桃,等.类风湿关节炎中医证候、证素分布的临床研究[J].北京中医药大学学报,2020,43(1):79-83. DOI: 10.3969/j.issn.1006-2157.2020.01.014.
- [4] 陈永前,娄玉铃,王颂歌.冬季加重的类风湿关节炎中医证候规律分析[J].风湿病与关节炎,2022,11(7):24-27. DOI: 10.3969/j.issn.2095-4174.2022.07.005.
- [5] ALETAHA D, NEOGI T, SILMAN A J, et al. 2010 rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative[J]. Arthritis Rheum, 2010, 62(9): 2569-2581. DOI: 10.1002/art.27584.
- [6] HÜGLE M, KALWEIT G, HÜGLE T, et al. A dynamic deep neural network for multimodal clinical data analysis[M/OL]//SHABAN-NEJAD A, MICHALOWSKI M, BUCKERIDGE D L. Explainable AI in healthcare and medicine. Studies in computational intelligence, vol 914. Cham: Springer, 2020: 79-92. (2020-11-03)[2023-06-25]. https://doi.org/10.1007/978-3-030-53352-6_8.
- [7] LAW A, GHOSH A. Multi-label classification using a cascade of stacked autoencoder and extreme learning machines[J]. Neurocomputing, 2019, 358: 222-234. DOI: 10.1016/j.neucom.2019.05.051.
- [8] HASSAN M R, HUDA S, HASSAN M M, et al. Early detection of cardiovascular autonomic neuropathy: a multi-class classification model based on feature selection and deep learning feature fusion[J]. Inf Fusion, 2022, 77: 70-80. DOI: 10.1016/j.inffus.2021.07.010.
- [9] PLEITER B, TAJALLI B, KOFFAS S, et al. Tabdoor: backdoor vulnerabilities in transformer-based neural networks for tabular data[J/OL]. arXiv:2311.07550.

- (2023-11-13) [2023-12-25]. <https://doi.org/10.48550/arXiv.2311.07550>.
- [10] SU J, ZHU M, MURTADHA A, et al. ZLPR: a novel loss for multi-label classification[J/OL]. arXiv:2208.02955. (2022-08-05) [2023-06-25]. <https://doi.org/10.48550/arXiv.2208.02955>.
- [11] MASUYAMA N, NOJIMA Y, LOO C K, et al. Multi-label classification via adaptive resonance theory-based clustering[J]. IEEE Trans Pattern Anal Mach Intell, 2023, 45(7): 8696-8712. DOI: 10.1109/TPAMI.2022.3230414.
- [12] ZHANG M L, ZHOU Z H. ML-KNN: a lazy learning approach to multi-label learning[J]. Pattern Recognit, 2007, 40(7): 2038-2048. DOI: 10.1016/j.patcog.2006.12.019.
- [13] WANG H, XU Y. Sparse elastic net multi-label rank support vector machine with pinball loss and its applications[J]. Appl Soft Comput, 2021, 104: 107232. DOI: 10.1016/j.asoc.2021.107232.
- [14] TSOUMAKAS G, KATAKIS I, VLAHAVAS I. Random k-labelsets for multilabel classification[J]. IEEE Trans Knowl Data Eng, 2011, 23(7): 1079-1089. DOI: 10.1109/TKDE.2010.164.
- [15] GONÇALVES T, QUARESMA P. A preliminary approach to the multilabel classification problem of Portuguese juridical documents[C/OL]//PIRES F M, ABREU S. Portuguese conference on artificial intelligence. EPIA 2003. Lecture notes in computer science, vol 2902. Berlin, Heidelberg: Springer, 2003: 435-444. [2023-06-25]. https://doi.org/10.1007/978-3-540-24580-3_50.
- [16] READ J, PFAHRINGER B, HOLMES G, et al. Classifier chains for multi-label classification[J]. Mach Lang, 2011, 85(3): 333-359. DOI: 10.1007/s10994-011-5256-5.
- [17] ESKANDARI S, GHASSABI S. Multi-label feature selection using adaptive and transformed relevance[J/OL]. arXiv:2309.14768. (2023-09-26) [2023-12-25]. <https://doi.org/10.48550/arXiv.2309.14768>.
- [18] LOU Q, DENG Z, XIAO Z, et al. Multilabel takagi-sugeno-Kang fuzzy system[J]. IEEE Trans Fuzzy Syst, 2022, 30(9): 3410-3425. DOI: 10.1109/TFUZZ.2021.3115967.
- [19] ZHANG X, ZHANG Q W, YAN Z, et al. Enhancing label correlation feedback in multi-label text classification via multi-task learning[J/OL]. arXiv:2106.03103. (2021-06-06) [2023-06-25]. <https://doi.org/10.48550/arXiv.2106.03103>.
- [20] LI S, XIA X, ZHANG H, et al. Multi-label noise transition matrix estimation with label correlations: theory and algorithm[J/OL]. arXiv:2309.12706. (2023-09-22) [2023-12-25]. <https://doi.org/10.48550/arXiv.2309.12706>.
- [21] MENCÍA E L, KULESSA M, BOHLENDER S, et al. Tree-based dynamic classifier chains[J]. Mach Lang, 2023, 112(11): 4129-4165. DOI: 10.1007/s10994-022-06162-3.
- [22] HAN Q, DU X, SUN Y, et al. Label dependencies-aware set prediction networks for multi-label text classification[J/OL]. arXiv:2304.07022. (2023-04-14) [2023-12-25]. <https://doi.org/10.48550/arXiv.2304.07022>.
- [23] BAKHSHI S, CAN F. Balancing efficiency vs. effectiveness and providing missing label robustness in multi-label stream classification[J/OL]. arXiv:2310.00665. (2023-10-01) [2023-12-25]. <https://doi.org/10.48550/arXiv.2310.00665>.

[本文编辑] 杨亚红