

DOI:10.13288/j.11-2166/r.2023.06.007

基于集成学习构建围绝经期综合征中医智能诊断模型

姚帅君, 闫敬来, 杜彩凤, 杨继红✉

山西中医药大学, 山西省晋中市高校园区大学街 121 号, 030619

[摘要] **目的** 基于集成学习中的模型融合方法建立围绝经期综合征中医智能化诊断模型。**方法** 检索中国知网 (CNKI)、万方数据知识服务平台 (WF)、维普中文科技期刊数据库 (VIP)、中国生物医学文献服务系统 (CBM) 以及古今医案云平台自建库至 2022 年 5 月 20 日收录的中医诊治围绝经期综合征的医案数据 1021 个, 以此建立数据集。将 1021 个数据样本按 7:3 比例分为训练集 714 个和测试集 307 个。利用随机森林、支持向量机、人工神经网络分别构建围绝经期综合征辨证诊断模型, 通过各分类器模型学习数据集并进行预测。通过网格搜索法进行模型的超参数调优, 然后通过异质集成的方法对三种分类器模型进行融合, 组成一个强学习器, 以准确率、精确率、召回率、f1 分数、受试者工作特征曲线下面积 (AUC 值) 等作为模型的评价指标。**结果** 分别构建了随机森林、支持向量机、神经网络、集成模型对其进行评估, 其准确率分别为 0.84、0.86、0.85、0.87; 精确率分别为 0.85、0.86、0.85、0.87; 召回率分别为 0.84、0.86、0.85、0.87; f1 分数分别为 0.84、0.86、0.85、0.87; AUC 值分别为 0.98、0.97、0.97、0.98。各指标均以集成模型最优。**结论** 集成模型能够综合各个弱监督模型的优点, 具有更好的稳定性, 其性能和泛化能力都优于单一学习器, 其结果对于临床中医诊断具有重要的参考价值。

[关键词] 围绝经期综合征; 集成学习; 智能诊断模型; 中医诊断; 人工智能

围绝经期综合征 (perimenopausal syndrome, PPS) 又称更年期综合征, 是指妇女绝经前后因性激素波动或减少所致的一系列躯体及精神、心理症状, 如月经紊乱、情志不宁、潮热面红、心悸失眠等^[1-2]。本病临床症状比较复杂, 轻重不一, 且持续时间数月甚至数年, 严重者可影响生活和工作, 危害妇女身心健康, 应对围绝经期的合理干预给予重视^[1-2]。根据 PPS 临床表现可将其归于中医学年老血崩、“脏燥”“百合病”的范畴^[3]。

辨证论治是中医认识疾病的基本原则, 也是中医区别于其他医学的重要特征^[4]。精准辨证是临床用药的基础, 也是临床疗效的保障。但中医辨证容易受个人主观因素的影响, 不确定性较强, 缺乏一定的客观性^[5]。机器学习以计算机为工具模拟人的思维方式, 能够通过学习知识解决问题, 目前在辅助医疗、医学影像、药物挖掘等领域发挥了十分重要的作用^[6]。在当前信息时代的大背景下, 将智能

技术与传统中医相结合, 开展中医智能化诊断研究, 对于中医药现代化有着深刻而长远的意义^[7]。统一规范的中医辨证模式是中医辨证智能化的基础, 同时也是中医现代化发展的重要内容之一^[8]。鉴于此, 本研究基于随机森林、支持向量机、BP 神经网络的机器学习方法, 对 PPS 中医智能化辨证进行研究, 并通过异质集成构建了智能辨证融合预测模型, 以期防治和干预 PPS 研究提供借鉴, 也为中医辅助诊断提供参考和依据。

1 资料

1.1 数据来源

本研究收集的医案数据来自两部分, 一部分来自中国知网 (CNKI)、维普中文科技期刊数据库 (VIP)、万方数据知识服务平台 (WF)、中国生物医学文献服务系统 (CBM) 建库以来至 2022 年 5 月 20 日收录的中医治疗 PPS 的临床文献; 另一部分来自古今医案云平台 2.3.5 (由中国中医科学院中医药信息研究所研发) 建库以来至 2022 年 5 月 20 日

基金项目: 国家自然科学基金 (81904247)

✉ 通讯作者: yangjh10@163.com

收录的有关 PPS 的中医病案。

1.2 医案筛选

各数据库检索表达式（按各数据库检索式呈现形式）如下。

中国知网：（主题：（围绝经期综合征+更年期综合征+经断前后诸证（精确））AND 主题：（医案+经验+病案+验案）（精确）），共检索得到文献 727 篇。

万方数据知识服务平台：（主题：（围绝经期综合征 or 更年期综合征 or 经断前后诸证）and 主题：（医案 or 经验 or 病案 or 验案）），共检索得到文献 1145 篇。

维普中文科技期刊数据库：题名或关键词=围绝经期综合征 or 更年期综合征 or 经断前后诸证 AND 题名或关键词=医案 or 经验 or 病案 or 验案，共检索得到文献 456 篇。

中国生物医学文献服务系统：（“围绝经期综合征”[全部字段：智能] OR “更年期综合征”[全部字段：智能] OR “经断前后诸证”[全部字段：智能]）AND （“医案”[全部字段：智能] OR “经验”[全部字段：智能] OR “病案”[全部字段：智能] OR “验案”[全部字段：智能]），共检索得到文献 785 篇。

在古今医案云平台以“围绝经期综合征”“更年期综合征”和“经断前后诸证”为主题词进行检索，共检索得到医案 559 则。

1.3 证型标准

依据《中西医结合内分泌代谢疾病诊治学》^[9]、《妇产科中西医结合诊疗技巧》^[10]、《中医妇科学》^[11]、《中医病证诊断疗效标准》^[12] 中有关 PPS 或经断前后诸证辨证分型的论述，将本病确定为肾阴虚、肾阳虚、肾阴阳两虚、肾虚肝旺、肝郁化火、肝气郁结、气滞血瘀、心脾两虚共 8 个证型。

1.4 纳入标准

1) 符合上述证型标准的中医治疗 PPS 的医案文献；2) 文献中四诊信息记录完整的中医医案及临床研究。

1.5 排除标准

1) 中西医结合治疗的临床病例；2) 四诊信息记录不完整（如缺乏舌象和脉象）的医案；3) 合并其他疾病（如乳腺癌、糖尿病等）的 PPS 临床病例；4) 学位论文。

2 方法

2.1 数据库的建立

去除重复文献后，纳入符合标准的文献 963 篇，包含可供录入的医案 1021 则。将 1021 则医案中患者的四诊信息及其对应的证型数据手动录入到 Excel 2019 中，采用双人双录入方法，再核对以保证数据的准确性。最终得到 1021 则医案样本的数据信息，以此作为原始数据库。考虑到中医医案的记录具有一定的随意性，故对原始数据库中涉及的患者四诊信息，参考《中医临床常见症状术语规范》^[13] 进行数据规范化处理，如将睡眠差、失眠多梦统一为夜寐不安，大便溏、大便不成形统一为大便秘结。对于每个医案的辨证部分则按照上述证型标准进行规范，如肾阴虚、肾水不足、肾阴不足等统一为肾阴虚，肝失条达、肝郁气滞、肝气郁结等统一为肝气郁结。最终原始数据库纳入肾阴虚证 215 例，肾阳虚证 61 例，肾阴阳两虚证 141 例，肾虚肝旺证 255 例，肝郁化火证 148 例，肝气郁结证 79 例，气滞血瘀证 63 例，心脾两虚证 59 例，共 1021 例。

2.2 数据预处理

2.2.1 数据集的建立 在完成原始数据库的建立之后，进一步对患者临床指标及证型进行编码，将自然语言文本转换成计算机可以识别的数值内容。首先以规范后的患者四诊信息作为特征值，使用二分类编码对各列特征进行赋值，即出现该特征记为“1”，不出现该特征记为“0”，详见表 1。然后以数据库中每个样本所对应的证型作为目标值，通过标签编码^[14]（Label Encoding）对中医证型进行处理，将文本内容转换为数值信息，建立 PPS 数据集。

2.2.2 数据集的划分 利用 train_test_split 模块将 PPS 数据集的 70% 作为训练集（714 个），30% 作为测试集（307 个），通过设置参数 stratify=y 使得训练集和测试集中的各个证型数据与原 PPS 数据集保持相同的比例，然后通过设置参数 random_state=0 保证每次实验都使用相同的训练集和测试集，增加实验的可重现性。

2.3 实验环境

使用 PyCharm 2021.1 作为开发环境，随机森林、支持向量机、集成算法采用的框架为 Scikit-learn（1.0.2），BP 神经网络采用的框架为 PyTorch（1.10.2）与 Skorch（0.11.0）。Skorch 是具有 Scikit-learn 风格的 PyTorch 高级封装，可以实现 Scikit-learn 的高效兼容^[15-16]，为本研究的神经网络模型

表1 纳入分析的围绝经期综合征医案中医四诊信息

Table 1 TCM four examinations information of the included medical cases of perimenopausal syndrome

项目	内容
症状	胸部胀闷、口苦、口干、善太息、盗汗、烦热易汗出、夜寐不安、健忘、头晕、眼干燥、耳鸣、心烦、心悸、手足心热、四肢厥逆、下肢冷、畏寒、喜悲伤欲哭、易怒、抑郁、胁痛、经行乳胀、外阴干燥、外阴瘙痒、腰酸、气短、神疲、乏力、胃脘痞闷、纳呆、尿短黄、尿清、大便溏薄、大便秘结、大便溏结不调、经血挟块、月经量多、月经量少、月经量时多时少、月经先期、月经后期、月经先后不定期、经血非时而下、过早绝经
面色	面色红赤、面色㿔白、面色萎黄、颧红
舌象	舌红、舌淡红、舌淡白、舌暗红、舌紫、舌胖大、舌有齿痕、舌有瘀点、苔白、苔黄、苔薄、苔厚、少苔、无苔
脉象	脉弦、脉细、脉滑、脉虚、脉涩、脉沉、脉迟、脉数、脉缓、脉弱

提供了类似的Scikit-learn接口。模型解释采用的框架为SHAP (0.40.0)，使用SHAP值解释机器学习模型和特征重要性^[17]。

2.4 样本类别不平衡的处理

在PPS的各个证型中，肾虚肝旺证占比最多，而肾阳虚证、心脾两虚证和气滞血瘀证所占比例较小。因此本实验所使用的数据集为典型的多分类类别不平衡数据集，为了使模型的泛化能力不受到影响，本研究在实验过程中针对支持向量机、随机森林模型直接设置参数class_weight=“balanced”平衡样本的权重，对于BP神经网络则手动计算各类别的权重后赋值给weight参数。

2.5 分类器模型的构建

2.5.1 随机森林

随机森林 (random forest, RF) 是一种基于决策树的集成算法，目前在中医证候研究中有了一定的应用^[18]。RF算法首先在变量和数据的使用上进行随机化，生成很多分类树，再汇总分类树的结果，能够很好地预测多达几千个解释变量的作用，被誉为当前最好的机器学习算法之一^[19]。在本研究中通过调用sklearn中的RandomForestClassifier算法实现随机森林的建模，选取gini系数作为特征选择的衡量指标 (criterion)。

2.5.2 支持向量机

支持向量机 (support vector machine, SVM) 是20世纪90年代中期提出的一种基于统计学习的模式识别的机器学习算法，它的学习策略就是间隔最大化，可形式化为一个求解凸二次规划的问题，也等价于正则化的合页损失函数最小化问题^[20]。SVM在非线性时在输入空间中构造分离超平面是利用核函数映射将输入空间的样本映射到高维特征空间H中，而对于线性可分时就在输入空间中直接构造分离超平面，这使它能够很好地解决小样本、非线性、高维数和局部最小等问题^[21]。在本研究中通过调用sklearn中的svm.SVM算法实现支持向量机的建模，选取高斯径向基核函数

(rbf)解决非线性映射问题，设置参数probability=True输出概率值。

2.5.3 BP神经网络

BP神经网络 (backpropagation neural network) 是人工生物网络的一种，它能够通过模仿生物大脑神经元之间的联系，建立输入到输出的非线性映射^[22]。反向传播 (back propagation, BP) 算法是一个可有效求解梯度的算法，其本质就是链式求导法则的应用^[23]。BP神经网络模型的拓扑结构通常由输入层、输出层和隐含层构成，信号经输入层输入，通过隐含层计算，由输出层输出，如输出值和标记值之间存在误差，则将误差反向传播至输入层，并对神经元的权值进行调整，该迭代过程循环进行，直到达到停止条件^[24]。本研究中通过torch.nn模块搭建神经网络模型，其中输入层与隐含层之间、隐含层与隐含层之间选择Relu激活函数，输出层选择Softmax函数进行多分类输出，使用crossentropy loss (交叉熵损失函数) 作为模型的学习策略，优化器选择Adam (自适应矩估计算法)，训练迭代周期设定为50。然后调用Skorch中的NeuralNetClassifier类，将PyTorch神经网络模型转换为Scikit-learn风格的评估器，以便更好地进行集成和调参。

2.5.4 集成模型

集成学习 (ensemble learning) 通过构建并结合多个学习器来完成学习任务，一般分为同质集成和异质集成两种方法。同质集成只包含同种类型的个体学习器，而异质集成中的个体学习器则是由不同类型的算法产生^[25]。集成学习的组合策略一般有平均法、投票法、学习法三种，其中投票法是实际应用中最常用的组合策略。通过集成学习可以降低过拟合的风险，提高分类器的泛化能力，其原理和结构如图1所示。目前集成学习方法已被广泛运用于疾病辅助诊断系统、人体行为识别、人脸识别等方面^[26]。本研究利用sklearn.ensemble模块进行上述分类器的融合，通过该模块

的 VotingClassifier 算法构建三种模型融合的异质集成分类器, 选取软投票法 (Soft voting) 输出类概率。

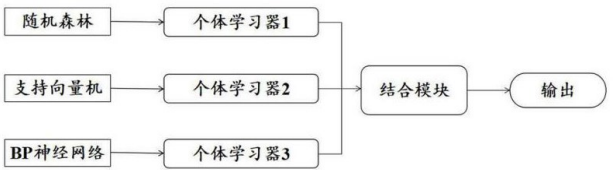


图 1 围绝经期综合征中医智能诊断集成学习示意图
Figure 1 Ensemble learning of TCM intelligent diagnosis of perimenopausal syndrome

2.6 超参数调优与模型评价

网格搜索是一种穷举搜索的超参数优化算法, 对于分类问题而言, 该方法将参数空间划分为若干网格, 通过循环遍历网格交叉点的所有参数组合来优化需要训练的模型, 同时计算对应模型的准确率, 当且仅当遍历完网格平面的所有节点时, 才能得到准确率最高的超参数组合, 即最优超参数组合^[27]。本研究模型的超参数优化均采用 scikit-learn 的 model_selection 库中的 GridSearchCV 模块通过参数列表 param_grid 寻找最优模型的超参数组合, 网格搜索使用分层 K 折 (K=5) 交叉验证。

模型评价使用准确率 (accuracy)、精确率 (precision)、召回率 (recall)、f1 分数 (f1 score)、混淆矩阵、受试者工作特征曲线 (ROC 曲线) 以及 ROC 曲线下面积 (AUC 值) 来评估模型的性能。AUC 值的计算方式为加权平均。

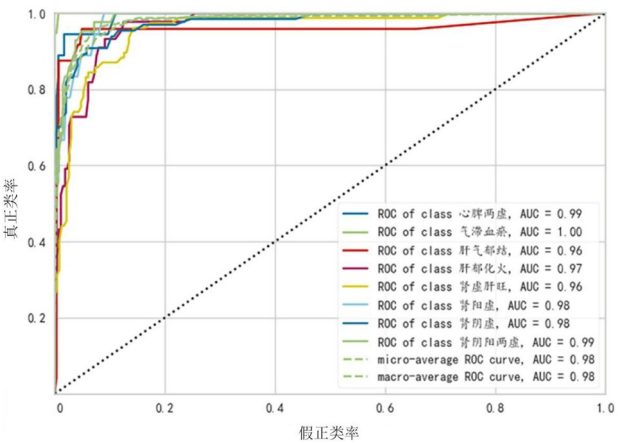
3 结果

3.1 模型评价

3.1.1 随机森林 随机森林算法输出的 PPS 各证型的 ROC 曲线和混淆矩阵见图 2 和图 3。结合图 2 和图 3 可以看出该模型在 PPS 的分类任务中共有 259 例分类正确, 其宏平均 AUC 值为 0.98。

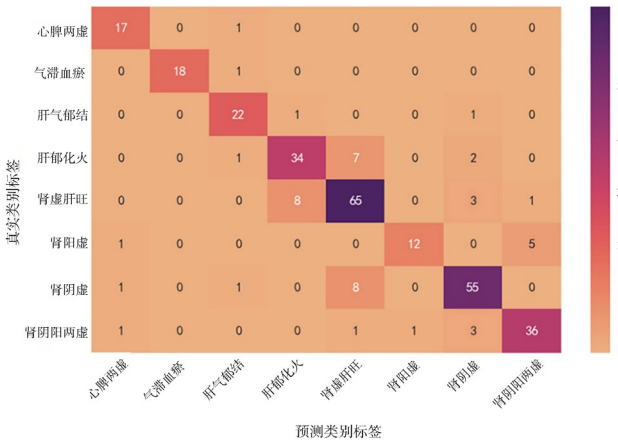
3.1.2 支持向量机 支持向量机算法输出的 PPS 各证型的 ROC 曲线和混淆矩阵见图 4 和图 5。结合图 4 和图 5 可以看出该模型在 PPS 的分类任务中共有 264 例分类正确, 其宏平均 AUC 值为 0.97。

3.1.3 BP 神经网络 BP 神经网络算法输出的准确率-损失 (accuracy-loss) 曲线见图 6, 随着迭代次数的增加, 准确率逐渐上升, 损失逐渐下降, 在训练周期为 50 左右时曲线趋于平缓, 达到收敛。该算法输出的 PPS 各证型的 ROC 曲线和混淆矩阵见图 7 和图 8。结合图 7 和图 8 可以看出该模型在 PPS 的分类任务中共有 261 例分类正确, 其宏平均 AUC 值为 0.97。



注: 不同颜色的线条代表不同的证型类别; 图例中 ROC of class 代表不同类别的 ROC 曲线 (受试者工作特征曲线), AUC 为 ROC 曲线下面积。

图 2 围绝经期综合征中医智能诊断随机森林模型 ROC 曲线
Figure 2 ROC curve of TCM intelligent diagnosis support vector machine model of perimenopausal syndrome



注: 混淆矩阵表格中颜色越深代表数值越大。

图 3 围绝经期综合征中医智能诊断随机森林模型混淆矩阵
Figure 3 Confusion matrix of TCM intelligent diagnosis random forest model of perimenopausal syndrome

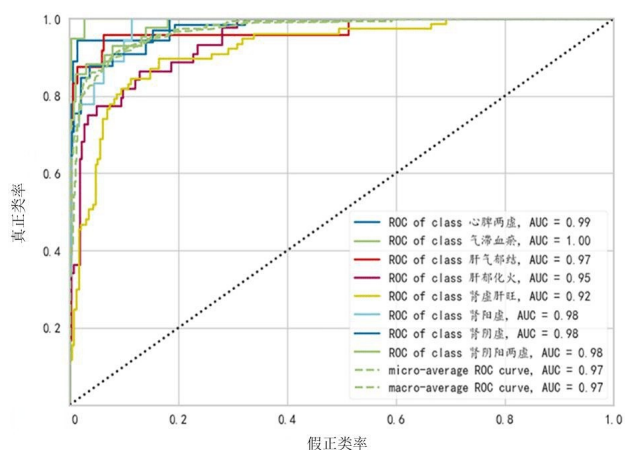
3.1.4 集成模型 集成算法输出的 PPS 各证型的 ROC 曲线和混淆矩阵见图 9 和图 10。结合图 9 和图 10 可以看出该模型在 PPS 的分类任务中共有 267 例分类正确, 其宏平均 AUC 值为 0.98。

3.2 模型比较

通过网格搜索得到的每个模型的最优参数如下。随机森林: n_estimators=150, max_depth=17, min_samples_leaf=1; 支持向量机: gamma=0.16, C=3.07; BP 神经网络: lr=0.0029, batch_size=17, num_units=48。在相同测试集下各模型的准确率、精确率、召回率、f1 分数以及 AUC 值如表 2 所示。

3.3 模型解释

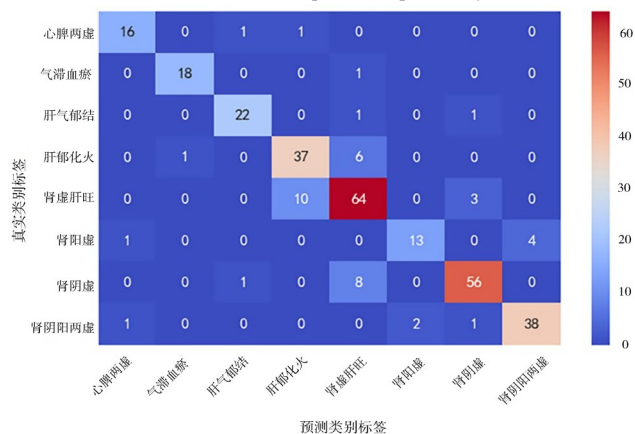
对于医疗行业而言, 模型的可解释性尤为重要



注：不同颜色的线条代表不同的证型类别；图例中 ROC of class 代表不同类别的 ROC 曲线（受试者工作特征曲线），AUC 为 ROC 曲线下面积。

图4 围绝经期综合征中医智能诊断支持向量机模型ROC曲线

Figure 4 ROC curve of TCM intelligent diagnosis support vector machine model of perimenopausal syndrome



注：混淆矩阵表格中颜色越深代表数值越大。

图5 围绝经期综合征中医智能诊断支持向量机模型混淆矩阵

Figure 5 Confusion matrix of TCM intelligent diagnosis support vector machine model of perimenopausal syndrome

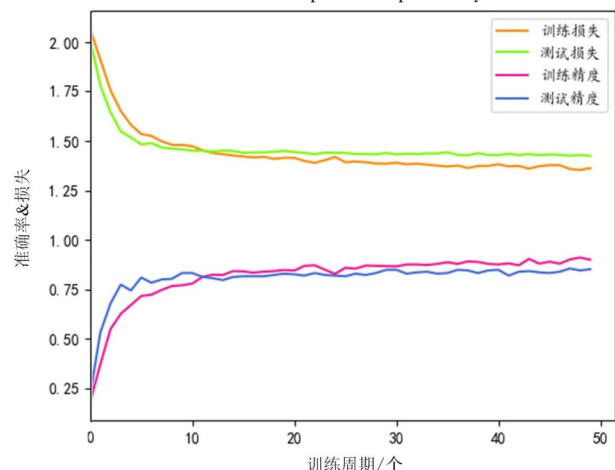
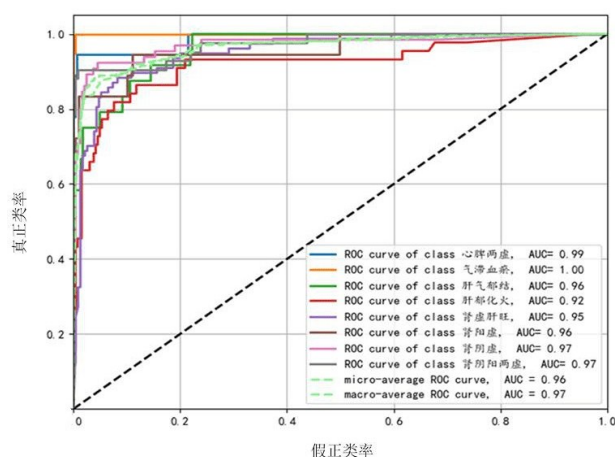


图6 围绝经期综合征中医智能诊断神经网络模型准确率-损失曲线

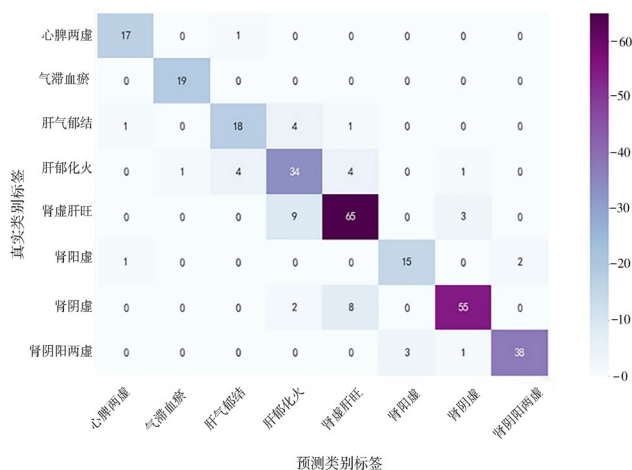
Figure 6 Accuracy-loss curve of TCM intelligent diagnosis neural network model of perimenopausal syndrome



注：不同颜色的线条代表不同的证型类别；图例中 ROC of class 代表不同类别的 ROC 曲线（受试者工作特征曲线），AUC 为 ROC 曲线下面积。

图7 围绝经期综合征中医智能诊断神经网络模型ROC曲线

Figure 7 ROC curve of TCM intelligent diagnosis neural network model of perimenopausal syndrome

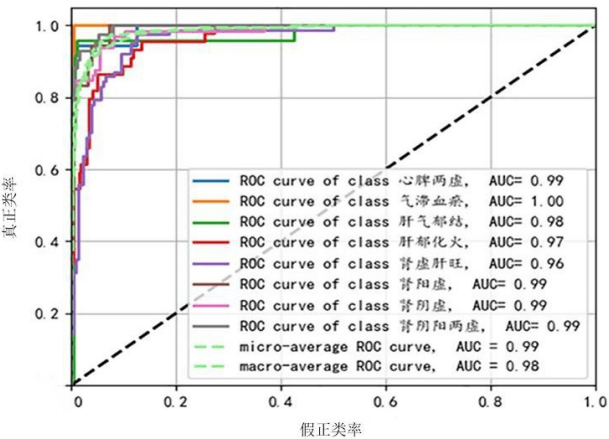


注：混淆矩阵表格中颜色越深代表数值越大。

图8 围绝经期综合征中医智能诊断神经网络模型混淆矩阵

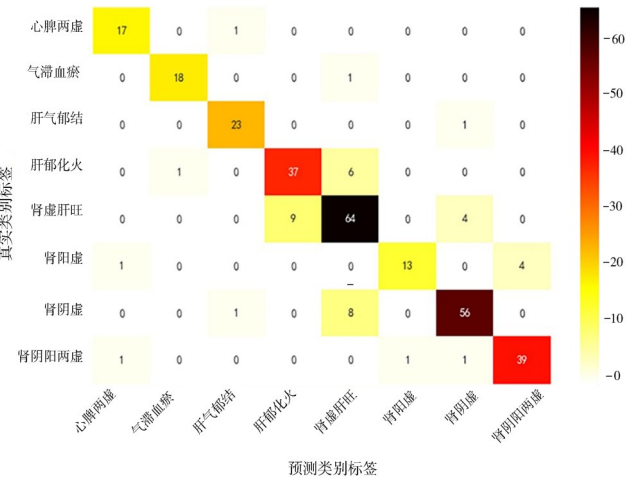
Figure 8 Confusion matrix of TCM intelligent diagnosis neural network model of perimenopausal syndrome

和必要，医生、护士都依赖新的医疗技术帮助他们监控和决策患者护理。因此，良好的可解释机器学习模型有助于提高临床工作者解决问题的能力，从而提高患者的护理质量^[28]。SHAP值（SHapley Additive exPlanations）是由 Lundberg 和 Lee^[29] 于 2017 年提出的用于解释任何机器学习模型输出的方法。SHAP 值源自博弈论中的夏普利值 (shapley value)，该方法量化了模型中每个特征对观察结果最终预测的贡献，使用基于所有可能的特征子集组合（包括给定特征）预测模型^[30]。本研究首先求得每个样本的特征对应的 SHAP 值，然后使用 SHAP 值的平均值作为该特征变量的重要性值，从而得到全局解



注：不同颜色的线条代表不同的证型类别；图例中 ROC of class 代表不同类别的 ROC 曲线（受试者工作特征曲线），AUC 为 ROC 曲线下面积。

图 9 围绝经期综合征中医智能诊断集成模型 ROC 曲线
Figure 9 ROC curve of TCM intelligent diagnosis ensemble model of perimenopausal syndrome



注：混淆矩阵表格中颜色越深代表数值越大。
图 10 围绝经期综合征中医智能诊断集成模型混淆矩阵
Figure 10 Confusion matrix of TCM intelligent diagnosis ensemble model of perimenopausal syndrome

表 2 围绝经期综合征不同分类器模型性能评估指标
Table 2 Performance evaluation indicators of different classifier models of perimenopausal syndrome

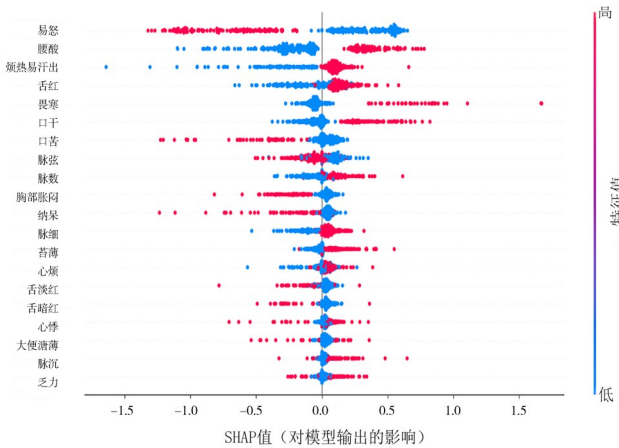
模型	准确率	精确率	召回率	f1 分数	AUC 值
随机森林	0.84	0.85	0.84	0.84	0.98
支持向量机	0.86	0.86	0.86	0.86	0.97
BP 神经网络	0.85	0.85	0.85	0.85	0.97
集成模型	0.87	0.87	0.87	0.87	0.98

注：AUC 值，受试者工作特征曲线下面积。

释，以此来明确模型中贡献度较大的特征。

从表 2 可以看出，集成模型输出的证型分类结果更具有可解释性，对此，本研究在集成模型的基

础上，调用 SHAP 库的 Explainer API 输出了 PPS 辨证模型排名前 20 特征重要性，见图 11。该图结合了特征重要度和特征的影响，y 轴上的位置由特征决定，x 轴上的位置由 SHAP 值决定^[31]。图中每一个点代表一个样本，点越离散代表该特征对模型的影响越显著，点越集中代表该特征对模型的影响越小。从图中可以看出，症状表现中贡献度较大的特征有易怒、腰酸、烦热易汗出、畏寒、口干、口苦等；舌象中贡献度较大的特征有舌淡红、舌暗红；脉象中贡献度较大的特征有脉弦、脉数、脉细、脉沉。同时我们也可以看出排名前 20 的特征变量中没有出现月经情况，表明女性的月经情况对模型的输出无显著影响。



注：图中纵坐标代表每一个特征；样本点颜色越红代表特征值越大，颜色越蓝代表特征值越小。
图 11 围绝经期综合征中医智能诊断 SHAP 特征摘要图
(排名 ≥ 20 的特征变量)

Figure 11 Summary of SHAP summary plot for TCM intelligent diagnosis of perimenopausal syndrome

4 讨论

《素问·阴阳应象大论篇》云：“年四十，而阴气自半也，起居衰矣。”《素问·上古天真论篇》亦云：“女子七岁，肾气盛，齿更发长；二七而天癸至，任脉通，太冲脉盛，月事以时下……七七任脉虚，太冲脉衰少，天癸竭，地道不通，故形坏而无子也。”这表明妇女经断之年肾气渐衰、天癸将竭，导致机体阴阳失衡。同时这一时期由于雌激素的减少、心理压力等因素的影响，也可能出现烦躁易怒的情况^[32]。综合这两方面的原因，这一时期女性大多会形成阴虚阳亢、水不涵木的功能状态，因此肾虚肝旺是围绝经期综合征最为常见病机。

中医辨证通过望闻问切收集四诊信息，运用司外揣内的思维模式对疾病进行诊断，这种思维模式

与机器学习中的分类问题相类似,二者均体现了对知识接受、规律总结以及实践运用的过程^[33-34]。同时,中医辨证的取类比象过程与人工智能信息采集过程相类似,二者之间有着良好的结合属性^[35]。因此将人工智能算法应用于中医诊断领域,可为中医四诊客观化提供科学方法,优化中医辨证的过程。

本研究基于随机森林、支持向量机、BP神经网络分别对PPS数据集进行了预测,并利用软投票法对三种学习器进行了融合,形成了集成模型,且在类别不平衡数据集上,该模型的性能和泛化能力优于单一的学习器。对于单一学习器而言,支持向量机有着更好的表现,其原因可能是,支持向量机能够将低维空间中线性不可分的样本映射到高位空间,构造的分离超平面能使训练集中的样本距其尽可能远,这更加适合中医证候的非线性和多维性特点^[36]。而集成学习的思想则是综合多个弱监督模型的优点,根据多个分类器的预测结果给出最终的预测结果。因此,即便某个弱学习器预测错误,其他学习器也能将错误纠正^[37]。研究表明,集成后的预测结果一般来说都会比单个学习器好,甚至在很多情况下模型性能会有质的飞跃^[38],这与本研究得出的结论是一致的,表明通过集成算法建立PPS中医辅助诊断辨证模型具有方法学上的可行性,其预测结果对临床中医诊断具有重要的参考价值。目前,已有学者将集成算法应用到中医诊断领域,如周璐等^[39]利用Scikit-learn构建了“方-证要素对应”的复合结构智能辨证模型,模型的输出准确率高于单一算法建模结果。任雪等^[40]使用主动集成学习的中医智能诊断模型在肺炎咳喘数据集上进行验证,其预测准确率高于传统机器学习算法,也高于当前主流的深度学习方法。张玉玺^[41]利用Weka构建了结直肠癌集成辨证模型,模型预测准确率达到93.35%,其结果高于三种个体学习器。

从图10可以看出,集成模型在心脾两虚证、气滞血瘀证、肝气郁结证的预测上有着良好的表现,这三种证型分别有1例分类错误。但由于肝郁化火证、肾虚肝旺证、肾阴虚证这三种证型的症状表现有一定的相似性,故模型在肾虚肝旺和肝郁化火、肾虚肝旺和肾阴虚的分类表现上不是十分理想。考虑其原因在于数据集规模偏小,导致BP神经网络的准确率偏低,没有充分体现出其应有的优势。其次,本研究仅将医案中症状的有无纳入研究范围,未能考虑症状的严重程度及前后变化。因此在今后的研究中,仍需进一步扩大数据集,将每个

症状的轻重程度分别进行编码,并对特征变量进行优化,删去对模型的输出无影响的特征,同时与现代医学的生化指标相结合,以期更好地提高辨证模型的性能和算法准确率,为中医智能诊断提供方法学上的参考,为PPS的智能辨证提供更有力的支持。

综上,通过机器学习算法挖掘新知识,对中医理论进行定性和定量分析已成为当前研究的热点^[42],构建智能辨证诊断系统将促进人工智能与中医的结合,使中医诊断更加客观化和标准化,这对于加快中医科技创新、促进中医药现代化有着重要的意义和深远的影响^[43]。

参考文献

- [1]刘碧岩.妇产科学[M].长春:吉林科学技术出版社,2019:312-315.
- [2]柳河.新编内分泌代谢病学[M].长春:吉林科学技术出版社,2019:318-326.
- [3]刘雁峰,王铁峰.中西医结合治疗围绝经期综合征述评[J].北京中医药大学学报,2022,45(1):15-20.
- [4]张婷婷,王亚强,温川飙,等.论智能中医辨证可解释研究的挑战和研究思路[J].中医杂志,2022,63(6):506-510.
- [5]许梦白,刘雁峰,赵宗耀,等.基于人工智能的不孕症中医辨证模型的构建与应用[J].中华中医药杂志,2021,36(9):5532-5536.
- [6]高婉卿,程宁,李力松,等.机器学习在方剂研究中的应用概述[J].中国中医药信息杂志,2021,28(6):131-137.
- [7]杨佳澄.基于群体智能的中医辨证诊断研究[D].兰州:兰州交通大学,2019.
- [8]杨涛,朱学芳.中医辨证智能化研究现状及发展趋势[J].南京中医药大学学报,2021,37(4):597-601.
- [9]方朝晖.中西医结合内分泌代谢疾病诊治学[M].北京:中国中医药出版社,2013:186-188.
- [10]范红霞,王霞灵,等.妇产科中西医结合诊疗技巧[M].广州:广东科技出版社,2005:341-343.
- [11]刘敏如.中医妇科学[M].成都:四川科学技术出版社,2007:96-100.
- [12]国家中医药管理局.中医病证诊断疗效标准[S].北京:中国医药科技出版社,2012:241.
- [13]黎敬波,马力.中医临床常见症状术语规范[M].北京:中国医药科技出版社,2005.
- [14]朱兆鑫,杨珊珊,彭安杰,等.基于自适应矩估计的BP神经网络对中医痛经证型分类的研究[J].世界科学技术-中医药现代化,2021,23(12):4560-4568.
- [15]李靖.基于注意力神经网络的Z玻色子标注[D].大连:大连理工大学,2021.
- [16]丁光磊.面向研究过程的任务通用型深度学习框架的

- 设计与实现[D]. 北京:北京邮电大学,2020.
- [17]MOLNAR C. Interpretable Machine Learning. A Guide for Making Black Box Models Explainable [EB/OL]. (2019-02-21) [2022-06-30]. <https://christophm.github.io/interpretable-ml-book/>.
- [18]刘金涛,尚津锋,吴宇峰,等. 基于随机森林算法的我国东北和西南地区名老中医治疗失眠异同比较[J]. 中国中医药信息杂志,2022,29(7):1-8.
- [19]蔡晓路. 基于随机森林的类风湿关节炎证型判别模型研究[D]. 北京:北京中医药大学,2016.
- [20]李航. 统计学习方法[M]. 北京:清华大学出版社,2019:111.
- [21]许明东,马晓聪,岳桂华,等. 基于支持向量机的高血压中医证候与血脂、尿酸、空腹血糖关系的研究[J]. 时珍国医国药,2016,27(12):3063-3065.
- [22]徐亮,陈守强,侯建辉,等. 基于BP神经网络的中医辨证模型构建方法探讨[J]. 世界中医药,2016,11(2):335-338.
- [23]廖星宇. 深度学习入门之PyTorch[M]. 北京:电子工业出版社,2017:86-87.
- [24]付姝菲,刘宏艳,任海燕,等. BP神经网络在中医药研究领域中的应用[J]. 医学信息,2021,34(12):12-18.
- [25]周志华. 机器学习[M]. 北京:清华大学出版社,2016:171.
- [26]高泽伦. 基于机器学习的疾病诊断模型研究[D]. 兰州:西北师范大学,2021.
- [27]张蓓蓓,胡敏. 基于网格搜索改进随机森林的顾客满意度预测[J]. 北京信息科技大学学报(自然科学版),2021,36(4):50-53,58.
- [28]陈珂锐,孟小峰. 机器学习的可解释性[J]. 计算机研究与发展,2020,57(9):1971-1986.
- [29]LUNDBERG S, LEE SI. A unified approach to interpreting model predictions [EB/OL]. (2017-12-04) [2022-05-01]. <https://proceedings.neurips.cc/paper/2017/file/8a20a8621978632d76c43dfd28b67767-Paper.pdf>.
- [30]LI R, SHINDE A, LIU A, et al. Machine learning-based interpretation and visualization of nonlinear interactions in prostate cancer survival[J]. JCO Clin Cancer Inform, 2020:637-646. doi: 10.1200/CCI.20.00002.
- [31]罗妍,王枫,叶文玲. 基于XGBoost和SHAP的急性肾损伤可解释预测模型[J]. 电子与信息学报,2022,44(1):27-38.
- [32]吕春梅,徐一丹,申思,等. 围绝经期女性生活质量影响因素的路径分析[J]. 中华现代护理杂志,2016,22(28):4025-4029.
- [33]舒琛洁,梁浩,刘淑明,等. 机器学习算法对证候要素“气虚”辅助诊断的性能评估[J]. 北京中医药大学学报,2021,44(10):928-934.
- [34]曹云,卢毅,陈建新,等. 基于机器学习的胃食管反流病中医智能辨证模型的应用[J]. 北京中医药大学学报,2019,42(10):869-874.
- [35]徐云浩,王洋,陶文娟,等. 中医辨证的象思维属性及对微观辨证的指导价值[J]. 中医杂志,2022,63(10):901-904.
- [36]张伟妃,张彦博,张瑞义,等. 基于多种信息处理方法的大肠癌中医证候模型的构建[J]. 中华中医药杂志,2021,36(2):1034-1038.
- [37]吴嘉乐. 异质集成学习器在鸢尾花卉分类中的应用[J]. 中国设备工程,2018(20):229-231.
- [38]邹永杰. 基于特征提取的分类集成在脾虚证诊断中的应用[D]. 广州:暨南大学,2008.
- [39]周璐,李光庚,孙燕,等. 复合结构智能化辨证选方模型的构建[J]. 世界中医药,2018,13(2):479-483.
- [40]任雪,郭燕. 基于主动集成学习的中医智能诊断模型及构建方法[J]. 中国循证医学杂志,2019,19(9):1118-1123.
- [41]张玉玺. 基于集成学习的结直肠癌辨证分型模型研究[D]. 贵州:贵州大学,2020.
- [42]罗思言,王心丹,饶向荣. 人工智能在中医诊断中的应用进展[J]. 中国医学物理学杂志,2022,39(5):647-654.
- [43]文志华,夏帅帅,刘东波,等. 中医智能辨证诊断技术的演进与问题探讨[J]. 世界科学技术-中医药现代化,2021,23(11):4298-4304.

Research on Intelligent Diagnosis Model of Traditional Chinese Medicine based on Ensemble Learning: Taking Perimenopausal Syndrome as an Example

YAO Shuaijun, YAN Jinglai, DU Caifeng, YANG Jihong

Shanxi University of Traditional Chinese Medicine, Jinzhong, 030619

ABSTRACT Objective To set up a traditional Chinese medicine (TCM) intelligent diagnosis model of perimenopausal syndrome based on the model fusion method in ensemble learning. **Methods** The databases including CNKI, WanFang Data, VIP, CBM as well as the ancient and modern medical case cloud platform were searched from inception date to May 20th, 2022, to collect the 1021 data samples of medical cases of perimenopausal syndrome treated by TCM and then establish the data set. The 1021 data samples were divided into training set ($n = 714$) and test set ($n = 307$) by a ratio of 7 : 3. Random forest, support vector machine and artificial neural network were used to build the

syndrome differentiation diagnosis models of perimenopausal syndrome, and the above classifier models were used to learn the data set and make predictions. Then the hyper-parameters of the models were optimized by grid search method, and the three classifier models were fused by heterogeneous integration method to form a strong learner. And the accuracy, precision, recall, f1 score, and the area under the curve (AUC) value were taken as the evaluation indicators of models. **Results** The accuracy rate of random forest, support vector machine, neural network and ensemble model was 0.84, 0.86, 0.85, and 0.87, respectively; the corresponding precision rates were 0.85, 0.86, 0.85 and 0.87; the recall rates were 0.84, 0.86, 0.85, and 0.87, respectively; the f1 scores were 0.84, 0.86, 0.85 and 0.87; the AUC values were 0.98, 0.97, 0.97 and 0.98. The ensemble model was optimal in all indicators. **Conclusion** The ensemble model which has better stability can synthesize the advantages of each weak supervision model, and its performance and generalization ability are better than that of a single learner, thereby having significant reference value for clinical diagnosis in TCM.

Keywords perimenopausal syndrome; ensemble learning; intelligent diagnostic model; traditional Chinese medicine diagnosis; artificial intelligence

(收稿日期: 2022-07-22; 修回日期: 2022-12-02)

[编辑: 焦 爽]

· 信 息 窗 ·

“传承精华 守正创新——加快推进新时代中医药高质量发展” 专家创新大讲堂在京举行 ——聚焦前沿推进中医药高质量发展

2月19—20日,由人力资源社会保障部、国家中医药管理局主办的“传承精华 守正创新——加快推进新时代中医药高质量发展”专家创新大讲堂在京举行。这是专家创新大讲堂开办以来首次聚焦中医药主题。人力资源社会保障部党组成员、副部长汤涛,国家卫生健康委党组成员、国家中医药管理局党组书记余艳红出席开幕式并致辞。国家中医药管理局副局长、党组成员秦怀金主持开幕式。

余艳红在致辞中表示,党的二十大擘画了中国式现代化全面推进中华民族伟大复兴的宏伟目标,部署了促进中医药传承创新发展的重点任务。推进中国式现代化,要求我们必须加快推进中医药现代化。希望各位专家学者坚持开放包容,紧扣中医药高质量发展这个主题,充分展示中医药领域各方面的最新研究成果,激发创新智慧,汇聚创新力量,促进学科交叉,着力营造百花齐放百家争鸣的良好学术氛围,凝聚起共同推动中医药振兴发展的磅礴力量。

大讲堂深聚焦中医药传承创新发展的重大问题、重大成果、重要进展、重大策略和重要举措,开展高端学术交流,展示学术研究进展与成果,旨在启发专家学者和青年人才对中医药传承创新的战略性、前瞻性思考,助力取得更多原创性、突破性成果,推动高层次人才队伍建设,促进中医药传承创新发展。

中国工程院院士、国家中医药管理局副局长、党组成员、中国中医科学院院长黄璐琦,中国工程院院士程京,中国工程院院士、国医大师王琦分别作题为《中药资源创新及其可持续利用研究》《工程科学艺术添翼助力中医中药发展》《中医基础理论的重构与时代发展》的特邀报告。中国工程院院士肖培根、石学敏、吴以岭、刘良、田金洲、肖伟等24位专家学者围绕中医药科技创新、中医医疗服务创新与体系建设、中医药经典传承与发展、新时代中医药人才培养、中药质量提升与产业发展、中医药疫病防治等6个主题作专题报告。

人力资源社会保障部、国家中医药管理局有关负责同志,部分院士、国医大师和全国名中医代表,有关高校、科研院所、医疗机构、中药企业负责同志,岐黄学者、全国中医临床优秀人才、青年岐黄学者、创新骨干人才、西学中骨干人才、全国老中医药专家学术经验继承工作指导老师及继承人以及博士后等专家人才代表,共2万余人通过现场和云端分会场参加。

(来源:中国中医科学院)