

项目技术指标与测试报告

——基于复杂脑网络的轻度认知障碍计算机辅助诊断

一、概述

本项目针对功能和性能两方面以医生及相关研究者两类用户的需求进行了技术标准与指标的制定与测试。测试分为功能测试和性能测试，其中功能测试又分为主界面、预处理、网络分析、模型构建与诊断五个步骤进行测试，最终软件成功实现了每个模块的数据导入、参数设置，数据处理及数据存储等功能。而在性能测试方面，本项目用留一法测试了构建模型的交叉验证准确率（CV-accuracy），并测试了预测结果的准确率（accuracy）、精确率（precision）和召回率（recall）。结果显示，软件 MCICAD 在功能测试和性能测试中均表现优异，能够帮助医生实现基于复杂脑网络的 MCI 计算机辅助诊断，同时也能协助此领域的研究人员更方便且高效地进行实验。

二、项目标准与技术指标

- （1）基本功能正常：能够分别面向医生和研究人员成功完成预处理、脑网络构建、特征提取、特征选择、分类等功能。能够对各参数进行自定义，并且读取保存文件功能正常。
- （2）模型构建交叉验证准确率 $\geq 80\%$ 。
- （3）预测结果准确率 $\geq 75\%$ ，精确率 $\geq 75\%$ ，召回率 $\geq 75\%$ 。

三、项目功能测试方案

1. 测试环境

硬件 (计算机)	CPU	Intel(R) Core(TM) i5-7200U CPU @ 2.50GHz
	ROM	16G
软件	操作系统	Windows 10
	客户端	MATLAB R2017b

2. 测试样本数据(所有数据均从 ADNI 官网获取)

测试内容 样本(个)	构建模型功能测试 (Model Building)	诊断功能测试 (Diagnosis)
认知正常组(CN)	35	15
轻度认知障碍组(MCI)	35	15
总计	70	30

3. 测试计划

- （1）测试构建分类模型（含预处理、网络分析、模型构建）过程是否完善，模型构建用留一法求出交叉验证准确率（CV-accuracy）并检验是否达标。
- （2）测试模型应用（诊断）流程是否完善，包括默认模型测试和新建模型测试。计算并检验诊断结果准确率（accuracy）、精确率（precision）和召回率（recall）是否达标

4. 测试步骤

4.1 功能测试

- （1）主界面测试（MCICAD）

任务描述	需求功能名称	测试用例名称
测试每个子界面能否正常启用。	帮助（ReadMe）	单击弹出软件开发信息，自动打开用户指南
	完成并退出（Done）	单击关闭软件
	预处理（Preprocessing）	单击进入预处理界面

	网 络 分 析 （ Network Analysis）	单击进入网络分析界面
	模 型 构 建 （ Model Building）	单击进入模型构建界面
	诊断（Diagnosis）	单击进入诊断界面

(2) 预处理界面测试(Preprocessing)

任务描述	需求功能名称	测试用例名称
测试能否正常完成预处理功能并输出每个样本的时间序列及相关矩阵文件	预处理(Preprocessing)	导入原始 fMRI 数据，导入设置模板文件，单击运行并开始预处理

(3) 网络分析界面测试(Network Analysis)

任务描述	需求功能名称	测试用例名称
测试能否正常导入文件，测试能否正常设置输出路径	设置输入输出路径（ Input & Output Path）	单击选择输入文件所在文件夹，单击选择输出文件存储文件夹
测试默认值是否正常显示，测试能否对每个设定值进行更改	阈值设定（ Threshold Settings）	键入阈值的最小值、最大值以及步长
测试默认选项是否勾选完毕，测试能否对每个可选项进行选择与取消	特征提取（ Feature Extraction）	单击勾选框对每个可选项进行选择或取消
测试能否正常完成网络分析的运行并将输出文件存在用户设定的文件夹中	运行（Run）	单击开始执行网络构建及特征提取操作，自动弹出任务完成提醒
测试能否正常退出界面	完成并退出（Done）	单击退出界面

(4) 模型构建界面测试(Model Building)

任务描述	需求功能名称	测试用例名称
测试能否正常导入文件	导 入 特 征 （ Load Features）	单击选择输入特征文件（.mat）
	导入标签(Load Labels)	单击选择输入标签（.xls 或.xlsx）
测试默认值是否正常显示，测试能否对每个设定值进行更改，测试需求功能能否正常运行（分别使用默认模型和新建模型进行测试）	特征选择（ Feature Selection）	单击选择特征类型，键入正则化参数（Lambda），键入特征数目，单击开始运行特征选择，自动弹出完成任务提醒
	建立模型(Build Model)	键入参数设置，单击开始建立模型，自动弹出完成任务提醒
测试结果能否正常显示，测试能否正常保存	结果（Results）	单击保存模型

模型		
测试是否能正常退出界面	完成并退出 (Done)	单击退出界面

(5) 诊断界面测试 (Diagnosis)

任务描述	需求功能名称	测试用例名称
测试能否正常导入文件	导入数据 (Load Data)	单击选择输入数据文件 (.mat)
	选择模型 (Load Model)	单击选择输入模型文件 (.mat)
测试诊断功能能否正常运行	诊断 (Diagnosis)	选择特征类型, 单击开始运行诊断, 自动弹出完成任务提醒
测试结果能否正常显示, 测试能否正常保存结果	结果 (Results)	单击保存诊断结果
测试是否能正常退出界面	完成并退出 (Done)	单击退出界面

4.2 性能测试

(1) 构建模型交叉验证准确率测试

测试样本数据:

	构建模型功能测试 (Model Building)
认知正常组 (CN)	35
轻度认知障碍 (MCI)	35
总计	70

测试方式:

任务描述	测试方法	相关标准
检测基于 AUC 特征的默认模型 CV-accuracy	基于留一法的交叉验证	CV-accuracy $\geq 80\%$
检测基于阈值特征的默认模型 CV-accuracy	基于留一法的交叉验证	CV-accuracy $\geq 80\%$

(2) 诊断结果的准确率、精确率和召回率测试

测试样本数据:

	诊断功能测试 (Diagnosis)
认知正常组 (CN)	15
轻度认知障碍 (MCI)	15

总计	30
----	----

测试方式：

真实值 \ 预测值	轻度认知障碍（MCI）	认知正常（CN）
被预测为 MCI	真正 (True Positive , TP)：被模型预测为正的正样本。	假正 (False Positive , FP)：被模型预测为正的负样本。
被预测为 CN	假负 (False Negative , FN)：被模型预测为负的正样本。	真负 (True Negative , TN)：被模型预测为负的负样本。

accuracy（准确率）： $(TP+TN)/(TP+FP+FN+TN)$

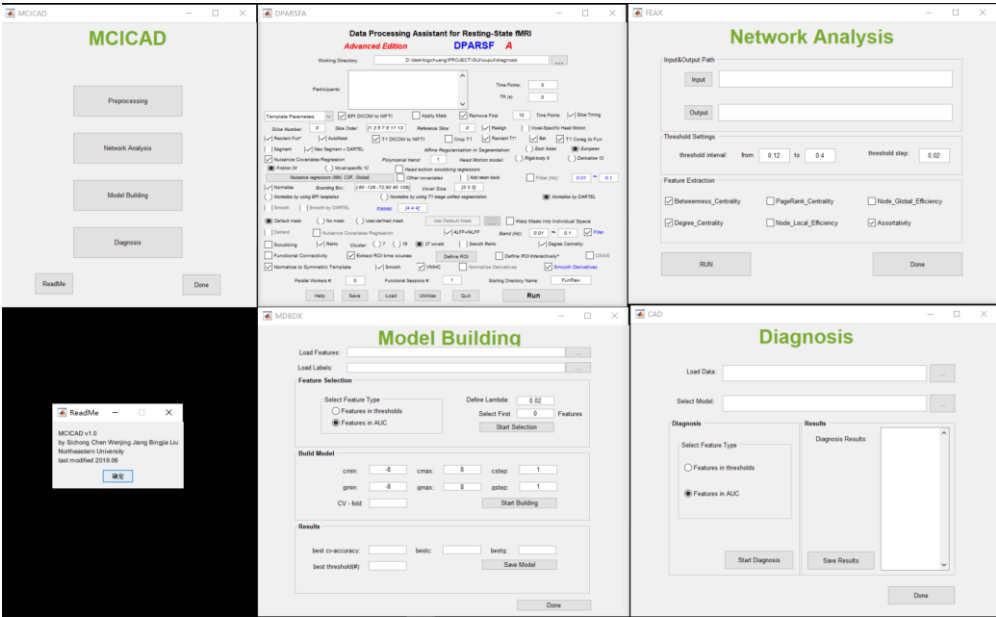
precision（精确率）： $TP/(TP+FP)$

recall（召回率）： $TP/(TP+FN)$

四、项目功能与性能测试记录

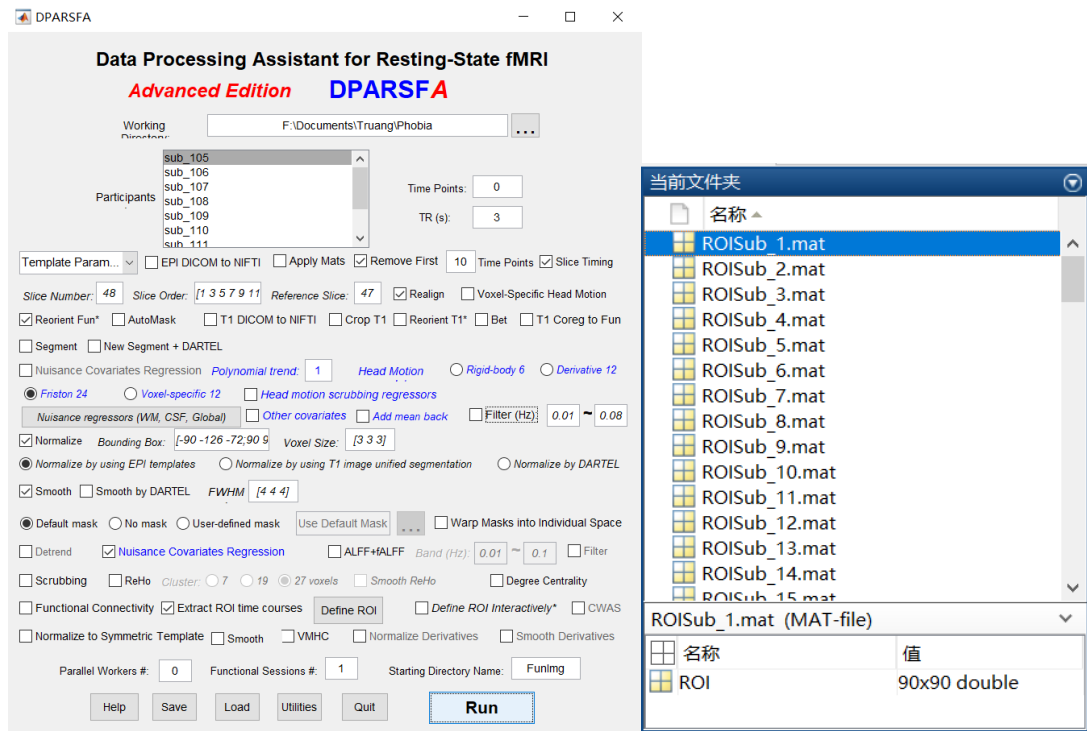
1. 功能测试记录

（1）主界面测试：



（主界面打开正常、各子 GUI 界面打开正常、ReadMe 信息读取正常）

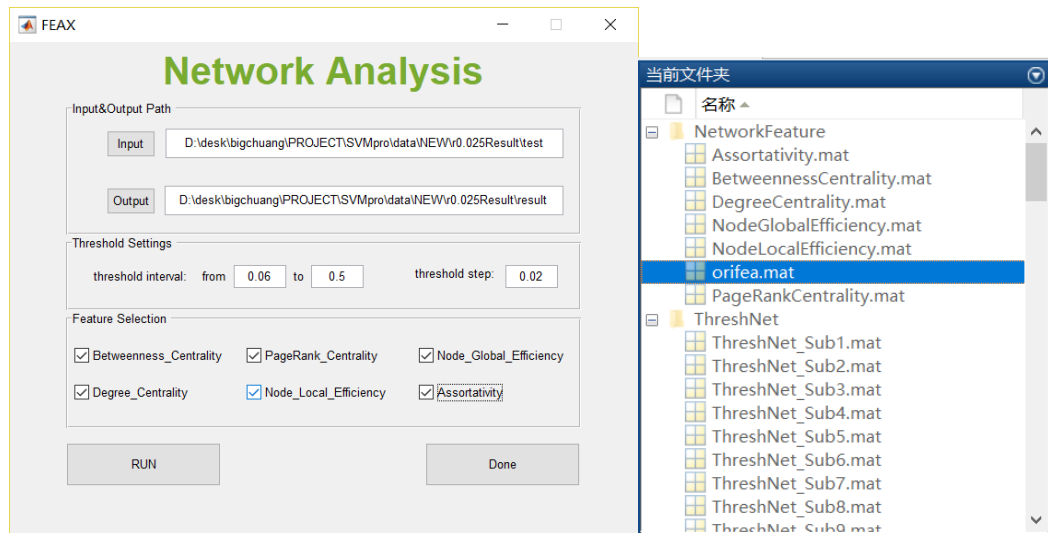
（2）预处理界面测试



(DPARSFA 导入文件功能和参数设置功能正常)

(预处理功能、存储数据功能正常)

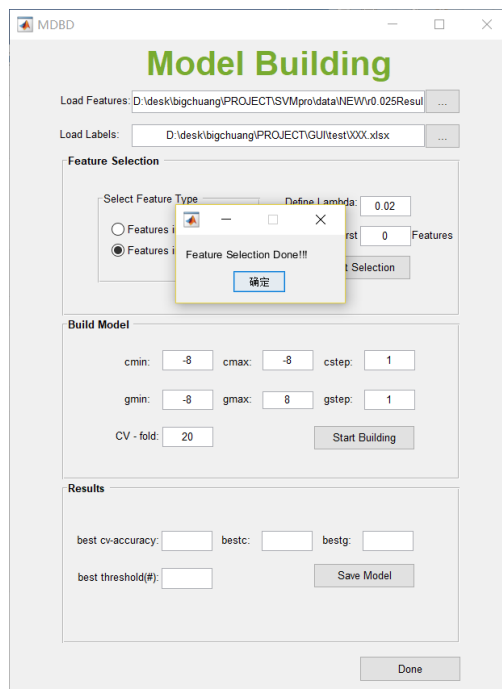
(3) 网络分析界面测试



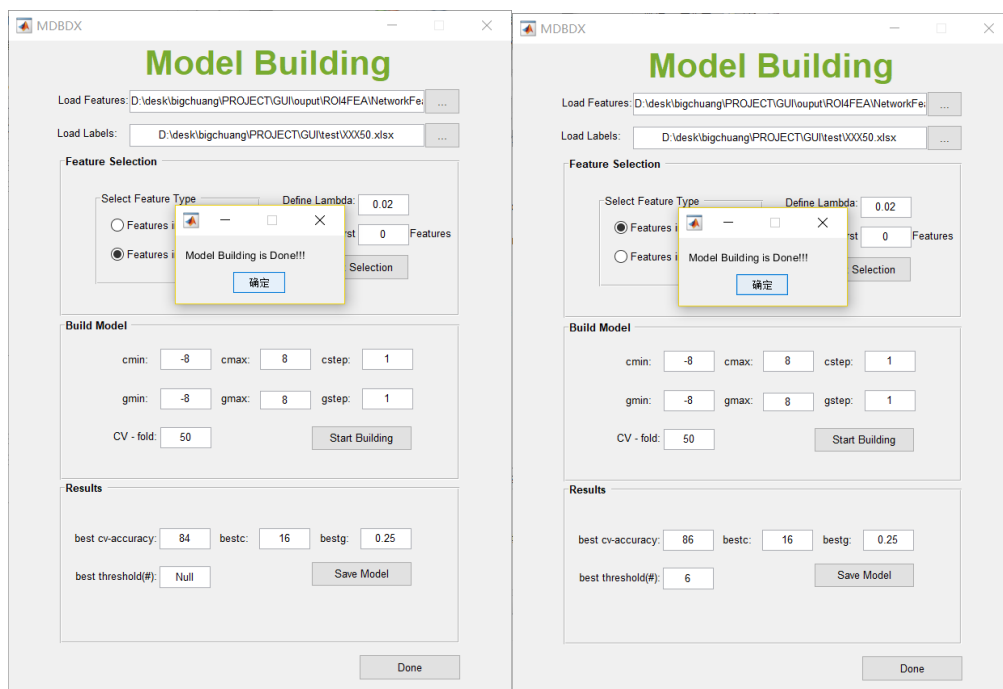
(导入数据、参数设置功能正常)

(特征提取、储数据功能正常)

(4) 模型构建界面测试



(导入数据、参数设置，参数传递和特征选择功能正常)



(模型构建功能及参数传递功能正常)

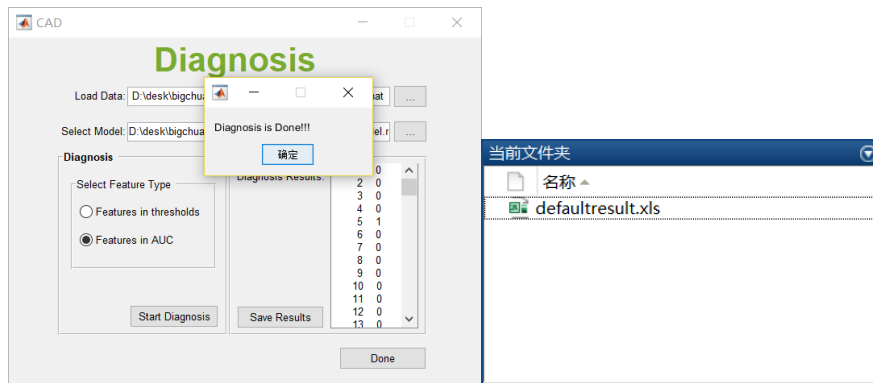
当前文件夹	名称	值
	auctest50model.mat	
	thdtest50model.mat	
auctest50model.mat (MAT-file)		
	MODEL	1x1 struct
	INDEX	190x1 double
	BESTTHD	1

当前文件夹	名称	值
	auctest50model.mat	
	thdtest50model.mat	
thdtest50model.mat (MAT-file)		
	MODEL	1x1 struct
	INDEX	121x1 double
	BESTTHD	6

(模型存储功能正常)

(5) 诊断界面测试

测试默认 AUC 模型 (defaultaucmodel):



(导入数据, 参数设置和分类功能正常)

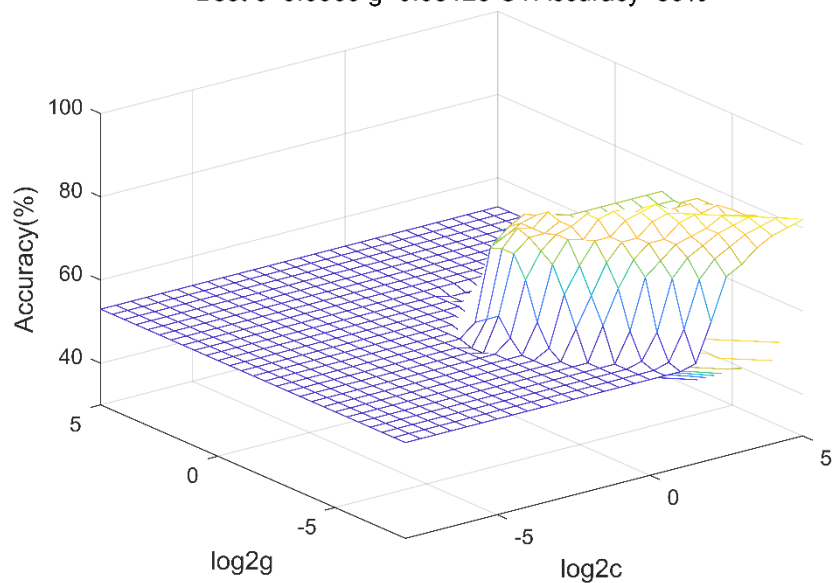
(数据存储功能正常)

2. 性能测试记录

2.1 模型构建交叉验证准确率测试记录

(1) 基于 AUC 特征的默认模型 (defaucmodel)

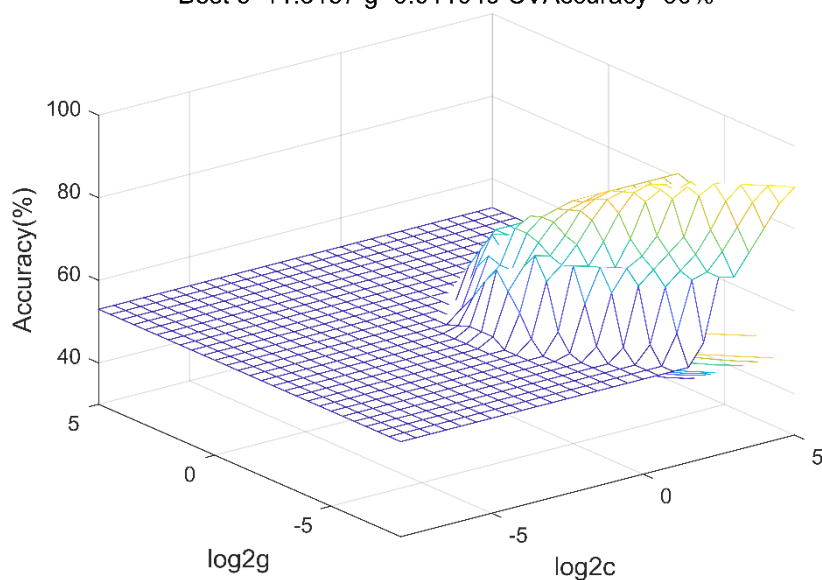
Parameter Selection Result in 3D(by SVM method)[GridSearchMethod]
Best $c=5.6569$ $g=0.03125$ $CVAccuracy=83\%$



(CV-accuracy = 83%)

(2) 基于阈值特征的默认模型 (defthdmodel)

Parameter Selection Result in 3D(by SVM method)[GridSearchMethod]
Best c=11.3137 g=0.011049 CVAccuracy=90%



(CV-accuracy = 90%)

2.2 诊断结果的准确率、精确率和召回率测试记录

(1) 使用基于 AUC 特征的默认模型 (defaucmodel) 的测试结果

真实值 \ 预测值	轻度认知障碍 (MCI)	认知正常 (CN)
被预测为 MCI	真正 (TP): 14 个	假正 (FP): 4 个
被预测为 CN	假负 (FN): 1 个	真负 (TN): 11 个

accuracy (准确率) = $(TP+TN)/(TP+FP+FN+TN) = 83.3\%$

precision (精确率) = $TP/(TP+FP) = 77.8\%$

recall (召回率) = $TP/(TP+FN) = 93.3\%$

(2) 使用基于阈值特征的默认模型 (defthdmodel) 的测试结果

真实值 \ 预测值	轻度认知障碍 (MCI)	认知正常 (CN)
被预测为 MCI	真正 (TP): 13 个	假正 (FP): 1 个
被预测为 CN	假负 (FN): 2 个	真负 (TN): 14 个

accuracy (准确率) = $(TP+TN)/(TP+FP+FN+TN) = 90.0\%$

precision (精确率) = $TP/(TP+FP) = 92.9\%$

recall (召回率) = $TP/(TP+FN) = 86.7\%$

(3) 两组默认模型测试结果对比

	accuracy (准确率)	precision (精确率)	recall (召回率)
aucmodel	83.3%	77.8%	93.3%
thdmodel	90.0%	92.9%	86.7%

五、小结

根据功能测试结果可知，构建分类模型（含预处理、网络分析、模型构建）过程完善，模型应用（诊断）流程完善，默认模型测试和基于 30 个样本的新建模型均能正常使用，参数设置，数据存取等功能正常，用于研究时提供了较高的便捷性与自由度。

由性能测试结果可知，基于 AUC 特征的默认模型通过留一法验证得到交叉验证准确率 $CV\text{-accuracy} = 83\%$ ，基于阈值特征的默认模型的交叉验证准确率 $CV\text{-accuracy} = 90\%$ ，两个结果均高于技术指标 80%。同时，对测试样本的诊断结果各项数据皆达到标准：使用基于 AUC 特征的默认模型（以下简称 aucmodel）进行测试时， $accuracy$ （准确率）= 83.3%， $precision$ （精确率）= 77.8%， $recall$ （召回率）= 93.3%。使用基于阈值特征的默认模型（以下简称 thdmodel）进行测试时， $accuracy$ （准确率）= 90.0%， $precision$ （精确率）= 92.9%， $recall$ （召回率）= 86.7%。

两个默认模型对测试样本的分类准确率、精确率和召回率都超出了标准，值得注意的是，由于 aucmodel 的训练数据为一个阈值区间内特征值的积分，它相较于 thdmodel 具有更高的泛化能力，利于应对更多不同的数据。同时，其召回率更是达到了 93.3% 的高标准，这意味着 aucmodel 能够精准地判别出大部分的 MCI 患者。

对于 thdmodel，其训练数据为具体阈值下的特征值，故在最佳阈值下能表现出更高的分类准确率（90.0%）和精确率（92.9%），这意味着 thdmodel 能在测试数据的阈值被限定的情况下得到准确率极高的分类结果。