

“心血管类诊疗（医用电子类、康复仪器）”赛道

命题项目 2（算法类）

（面向本科生与研究生开放）

一、题目

$N \rightarrow N+1$ ：可识别未知类型心电数据的 AI 算法

二、概述

心电异常识别是心电信号处理领域的一项重点工作，近年来，得益于人工智能技术的飞速发展，机器学习算法应用于心电异常识别层出不穷，特别是 2019 年吴恩达团队在 *Nature Medicine* 期刊发表了心电异常 AI 分类模型工作，心电 AI 计算成为研究热点，各种算法相继提出，大大推动了心电异常识别工作的进展。

然而，几乎所有的心电 AI 算法都是针对给定的 N 种类型心电数据，通过模型训练，进而在未知测试数据上实现 N 种分类，即 $N \rightarrow N$ 的 AI 学习，这大大限制了具有自主认知功能的心电 AI 发展。目前还未出现：给定训练心电数据有 N 类，可以训练一个模型，能够识别某个测试数据不属于这 N 类中一种，即 $N \rightarrow N+1$ 的 AI 学习。只有这样，才能发展具备自主认知能力、拓展自主认知边界，实现真正的医学 AI。

本赛题以常见的心电正常和房颤信号识别为背景，希望参赛者能够通过对正常、房颤两类心电信号的 AI 学习，能够使得模型能够不仅仅识别测试数据中的正常和房颤心电，还能识别除正常和房颤外的其他类信号，从而构建一个对不在标签范围内的异常心电信号仍有一定识别能力的心电分类模型。

三、任务要求

基于本赛题提供的正常和房颤 2 类心电数据，要求参赛者设计一种 AI 算法，能够识别出测试数据库中正常、房颤、其他非正常/房颤 3 类心电数据类型，实现模型具有 $N \rightarrow N+1$ 的自主学习能力。

提交作品应包括如下要素：1. 选题背景意义；2. 研究目标；3. 算法设计原理 4. 算法实施方案（包括算法流程图、主要程序等）；5. 算法结果；6. 分析与讨论；7. 总结与展望。

四、竞赛数据

本赛题数据为单导联心电信号，采样率为 400 Hz，信号时间长度 10 s，数据格式是 .mat 文件。赛题提供了有标签的训练集数据和无标签的测试集数据，其中训练集数据包括正常心电（标签为 1，10000 条，train_nor 数据集）和房颤心电（标签为 2，10000 条，train_af 数

数据集）共 2 类，全部有标签，共计 20000 条；测试集数据包括正常心电、房颤心电、其他类型心电（见下表，非正常或房颤心电），共计 2000 条，无标签标注。参赛者需要设计算法准确识别测试集中的 3 类信号，并依据下表给每个信号打标签。

心电种类	正常 (Normal)	房颤 (AF)	其他 (Others)
标签	1	2	3

数据网址：https://pan.baidu.com/s/1qa_0P1KJCUibp2YYifrYgg

提取码：x71r

五、评估指标

采用识别准确率 Acc 值进行算法性能评价， $Acc = \frac{\text{正确识别的样本数}}{\text{总样本数}}$ 。例如有 10 个信号，标签分别为[1 1 1 2 2 2 3 3 3 3]，算法识别结果为[1 2 1 3 2 2 3 1 3 3]，则 $Acc=70\%$ 。

六、评分标准

评分考察内容	得分项	分数
结果准确性	组委会根据竞赛团队提交的测试集预测结果进行客观评分。	20
算法科学性和创新性	算法的设计思路和原理是否科学、合理，算法的创新点是否清晰、明确、突出。	30
算法有效性和展示度	算法流程是否清晰、规范，算法代码是否完整、真实、可靠，算法有效性验证效果如何。	30
答辩现场表现	答辩 PPT 展示是否完整、规范、美观，答辩选手陈述是否清晰、回答问题是否准确。	20
总分		100

七、注意事项

1. 本赛题提交分为 2 个阶段：

阶段 1 初赛阶段：自赛题发布至竞赛结束前 1 周，参赛者需要将测试集的预测结果按照组委会规定的要求整理成.csv 文件（命名为：`answer_队名_提交次数.csv`），然后以邮件形式发送至本赛题组委会邮箱 shelab2024@163.com，邮件主题命名为：赛道+队名。每个队伍在阶段 1 有 5 次提供测试机会，组委会将通过邮件反馈分数，并在竞赛网站上公布即时成绩。比如一个参赛队队名为：ABC，第 3 次提交预测结果，相应文件名为：`answer_ABC_3.csv`。

阶段 2 复赛阶段：竞赛结束前 1 周至竞赛结束，参赛队选择性能最好的一组代码和预测结果（命名为：`answer_队名_best.csv`），将训练代码和预测结果打包，发送至本赛题组委会

邮箱 shelab2024@163.com，组委会运行参赛队代码，核实其结果真实性，文件中需包含训练源代码以及解释如何进行复现 **readme** 文件。对于无法提供代码复核以及代码无法复现的队伍，组委会有权取消比赛资格。组委会将在网站更新最终排行榜（以每支队伍提交的最好成绩计入最终排名）。

2. 提交的设计报告撰写条理清晰、简洁明了，格式正确，参考文献标注准确，其中正文前言和问题引入部分不超过 2 页，算法设计、结果展示、讨论等部分不超过 8 页，报告整体不超过 12 页。

3. 不得抄袭、不得作弊，否则一律取消评审资格。