Software Engineering and Applications 软件工程与应用, 2024, 13(2), 155-164
Published Online April 2024 in Hans. https://www.hanspub.org/journal/sea
https://doi.org/10.12677/sea.2024.132016

Hans汉斯

基于三重集成采样的胎儿健康分类

刘　凯，刘靖宇

哈尔滨师范大学计算机科学与信息工程学院，黑龙江　哈尔滨

收稿日期：2024年2月28日；录用日期：2024年3月19日；发布日期：2024年3月27日

摘　要

胎心监护(CTG)作为监测和判断胎儿宫内健康情况的一种方法，已广泛应用于临床。产时胎心监护能连续动态的观察宫缩时胎儿在宫内的变化，提高对胎儿不良预后预测值，及时为临床做出前瞻性诊断，指导产程处理，以降低围生儿并发症和病死率。为了解决数据不平衡问题，研究采用了一种基于过采样的算法称为三重集成过采样，这一算法通过对少数类样本点进行聚类分析，将它们划分为核心点、边界点和噪音点，并随后采取不同的采样策略，扩展了边界区域的密度，并解决了少数类分布不均、新生成样本的模糊类边界等问题，旨在解决常规SMOTE过采样方法的缺点。实验结果表明，XGBoost算法在胎儿健康诊断任务中表现出色，达到了96.6%的准确率。

关键词

胎心监护，胎儿健康诊断，机器学习

Fetal Health Classification Based on Tri-Ensemble Sampling

Kai Liu, Jingyu Liu

College of Computer Science and Information Engineering, Harbin Normal University, Harbin Heilongjiang

Received: Feb. 28th, 2024; accepted: Mar. 19th, 2024; published: Mar. 26th, 2024

Abstract

Cardiotocography (CTG), as a method for monitoring and assessing the intrauterine health of fetuses, has been widely utilized in clinical practice. During childbirth, continuous monitoring of fetal heart rate dynamics allows for observation of fetal changes in utero during contractions, thereby enhancing the prediction of adverse fetal outcomes. This facilitates proactive clinical diagnosis, guides labor management, and reduces the incidence of neonatal complications and mortality. To address the issue of data imbalance, the study employed a oversampling-based algo-

文章引用: 刘凯, 刘靖宇. 基于三重集成采样的胎儿健康分类[J]. 软件工程与应用, 2024, 13(2): 155-164.
DOI: 10.12677/sea.2024.132016

rithm known as Tri-Ensemble Sampling. This algorithm conducts cluster analysis on minority class samples, categorizing them into core points, boundary points, and noise points. Subsequently, it adopts different sampling strategies to expand the density of boundary areas, resolving issues such as uneven distribution of minority classes and ambiguous boundaries of newly generated samples, aiming to overcome the shortcomings of conventional SMOTE oversampling methods. Experimental results demonstrate outstanding performance of the XGBoost algorithm in fetal health diagnosis tasks, achieving an accuracy of 96.6%.

Keywords

Cardiotocography (CTG), Fetal Health Diagnosis, Machine Learning

Copyright © 2024 by author(s) and Hans Publishers Inc.
This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).
http://creativecommons.org/licenses/by/4.0/
Open Access

1. 引言

据世界卫生组织统计，2020 年共有约 287,000 名妇女在怀孕和分娩期间及之后死亡，每天约有 800 名妇女死于与怀孕和分娩相关的可预防原因。这意味着大约每两分钟就有一名妇女死亡，其中绝大多数是可以预防的[1]。贫困国家承担了近 95%的孕产妇死亡人数，这种不均衡的死亡率分布反映了全球在获得医疗保健和治疗方面的差距。因此，确保孕妇和胎儿在妊娠期间的身体健康和稳定发育，准确预测腹中胎儿的健康情况至关重要。

本研究旨在解决胎心宫缩图(CTG)数据集中存在的不平衡问题，以提高胎儿健康分类的准确性和可靠性。具体而言，在医学领域 CTG 数据通常包含大量正常样本，但较少的可疑样本或病理样本，这导致了数据不平衡的问题。这种数据不平衡可能导致分类模型在识别可疑样本和病理样本时性能下降，模型更倾向于预测出现频率更高的正常样本。

合成少数过采样 SMOTE [2]。技术是一种广泛使用的过采样方法，用于解决不平衡数据集的问题。尽管它在很多情况下表现出色，但它也存在一些缺点：

1) 新样本生成取决于根样本和辅助样本的选择。若二者均属于少数类区域，合成的新样本被认为有效；但若其中一个是噪声样本，新样本可能跑到多数类区域，干扰正确分类，通常被视为无效。选择合适的根样本和辅助样本对于合成合理的新样本至关重要，以避免引入噪声和损害分类性能。

2) 未考虑多数类样本的分布情况。从处于类边界的少数类样本中合成新样本，其 k 近邻样本也处于类的边界，则经插值合成的少数类样本同样会落在两类的重叠区域，从而更加模糊两类的边界。

3) 少数类样本分布不均匀，则合成样本多数位于密集区，导致分类算法难以识别稀疏区的少数类样本，从而影响分类准确性。

为了解决这三个问题，我们对提出了一种新的方法三重集成过采样，旨在解决合成少数过采样 SMOTE 技术存在的一些问题。首先，三重集成过采样通过使用 DBSCAN 聚类算法[3]。对少数类样本进行聚类，将样本分为核心点、边界点和噪音点三类。对于噪点，选择不进行处理。利用 DBSCAN 聚类算法本身的特性，可以很好地发现离群点。对于可能存在的噪点，决定不生成新的样本，以避免引入噪声并解决上述新生成样本的质量问题。对于含多数样本的边界点，同样选择不进行处理。这些边界点含有多数类样本，证明它们处于决策边界附近。因此，对该类不进行处理可以使决策边界变得更加清晰，从

而解决上述类边界模糊问题。对于核心点，使用 SMOTE 算法进行过采样，以保留原有算法的密集区域。同时，对于不含多数样本的边界点，采用 ADASYN 算法进行过采样[4]。以扩展边界区域的密度。通过这种方式，在一定程度上解决了少数类分布问题。

2. 相关工作

在处理不平衡胎儿数据集的传统方法中，处理不平衡的胎儿数据集的方法主要使用 SMOTE 技术进行过采样，然后结合各种机器学习模型如 Adaptive Boosting (AdaBoost)，Categorical Boosting (CatBoost)，Decision Tree (DT)，Gradient Boosting (GB)，K-Nearest Neighbors (KNN)，Random Forest (RF)，Support Vector Machine (SVM)，Extreme Gradient Boosting (XGBoost)，Light Gradient Boosting Machine (LGBM)，Long Short-Term Memory (LSTM)和 Artificial Neural Network (ANN)以达到数据平衡和分类性能的优化目标。Md Takbir Alam 等人[5]使用 SMOTE 方法平衡胎儿数据集并结合六种机器学习模型，在随机森林模型实现了 97.51 的准确率。Sai Prasad Potharaju 等人[6]通过应用 SMOTE 技术来平衡数据集，从而提高基于树和惰性学习算法的分类效率。实验结果表明，平衡数据集的分类性能优于不平衡数据集。Chamidah 等人[7]采用 K-means 算法并结合 SVM 模型进行特征提取和重构以减少 CTG 输入数据的特征维度，在 CTG 胎儿数据集上实现了 90.64%的准确率。Haad Akmal 等人[8]为了解决 CTG 数据集的不平衡性，采用了 SMOTE 技术生成合成条目以抵消不平衡，采用了一种基于特征提取、特征选择和贝叶斯优化的机器学习模型，该模型在胎儿状态分类上实现了 96.62%精度和 94.96%的 CTG morphological patterns 分类准确度。Jayakumar Kaliappan 等人[9]采用了多种交叉验证技术以提高模型性能并确定性能最佳的算法。通过探索性数据分析，在应用交叉验证技术后，梯度提升和投票分类器实现 99%的准确率。Jayashree Piri 等人[10]采用了 SMOTE 过采样技术来平衡数据集，并采用一种基于拥挤距离的多目标蚁狮优化(MOALO-CD)进行特征选择，通过提高多种分类器的性能，找出了对胎儿健康的关键影响因素，实现 95%的准确率。Hoodbhoy 等人使用 CTG 数据进行训练，评估了十种不同机器学习分类模型的性能。实验结果表明，在胎儿状态的预测中，XGBoost 模型在测试数据上表现最出色，实现了 92%的精度。Yandi Chen 等人[11]通过主成分分析和可视化技术挖掘 CTG 数据的分布特征，结合深林算法和分类器，在外部公共数据集上实现了 92.64%的准确率、92.01%的平均 F1 值和 0.990 的曲线下面积(AUC)值。

然而，现有的胎儿监测模型大多侧重于使用 SMOTE 技术平衡样本和提高整体分类精度，通常忽视了 CTG 数据集样本不平衡的分布，导致可疑样本和病理样本的准确率在较低范围内，约在 45%至 82%和 66%至 94%之间。本研究提出三重集成过采样解决了样本不平衡问题，提高对可疑和病理样本的分类准确率，增加胎儿监测模型的可靠性和准确性。

3. 材料和方法

3.1. 数据描述

本研究所使用的数据集来源于加州大学欧文分校机器学习存储库[12]。它记录了 2126 名处于妊娠期孕妇的胎儿心跳(FHR)和子宫收缩(UCs)情况，该数据集由 21 个特征，3 个类标签组成，这 3 类标签分别为正常、可疑、病理。

3.2. 数据分析

在医疗领域中，不平衡数据是一个具有挑战性的问题。通过可视化分析发现，在图 1 中，胎儿数据集存在样本不平衡的现象，而数据集不平衡可能会导致其中的分类性能降低[13]。通过主成分分析 PCA [14]降维处理，我们展示了对数据集的可视化，以更清晰地理解数据分布。

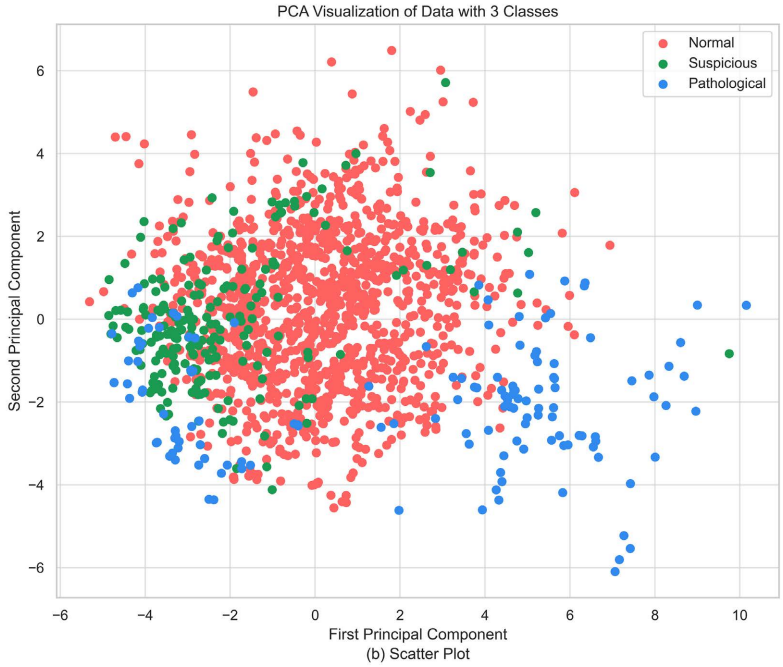

Figure 1. Distribution of fetal health classes
图 1. 胎儿健康类别的分布

如图 2 所示，在将原本的 21 个维度降至两个主成分后，我们可以观察到正常样本(用红色标识)在数据分布中分散在中央。而可疑样本(绿色)和病理样本(蓝色)之间存在较多的重叠区域。这种重叠可能导致在分类时容易混淆正常、可疑和病理样本。

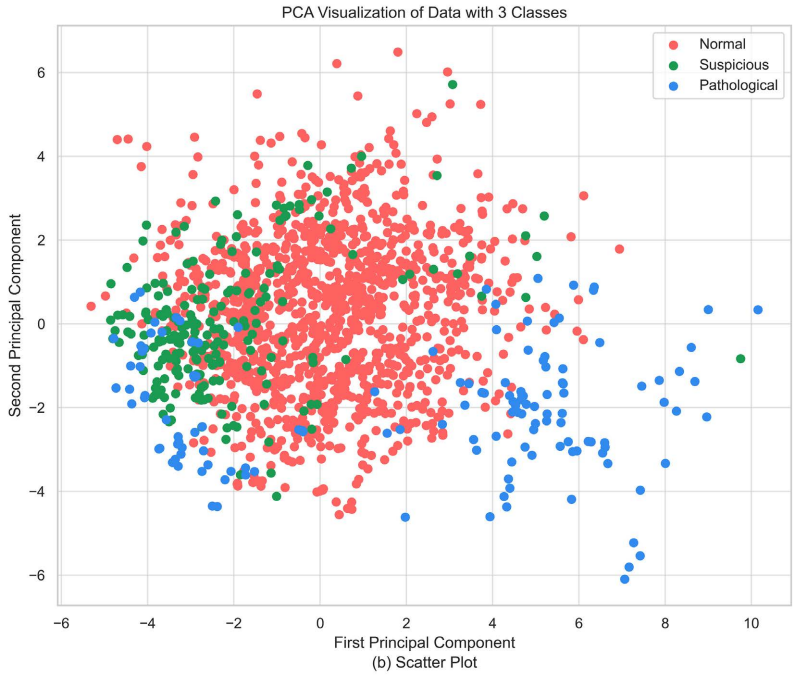

Figure 2. Overview of principal component analysis dimensionality reduction on the fetal dataset
图 2. 胎儿数据集的主成分分析降维概况

3.3. 使用三重集成过采样处理数据

3.3.1. DBSCAN 处理少数类样本

我们将数据集划分为多数类样本(记为 MM)和少数类样本(记为 MS)，多数类和少数类样本差(记为：D = MM-MS)，然后专注于处理少数类样本。为了更好地理解少数类样本的内在结构，我们采用了 DBSCAN 聚类算法对其进行聚类操作。DBSCAN 是一种强大的聚类算法，具有出色的离群点检测能力，它能够帮助我们找到潜在的异常样本或离群点。

通过 DBScan 算法，我们将少数类样本分为三个主要类别，包括核心点、边界点(含多数样本的边界点和不含多数样本的边界点)和噪音点，以更清晰地描述其内在分布特性。核心点代表着少数类样本中的主要聚类中心，边界点位于聚类边缘，而噪音点可能表示异常或不符合主要模式的样本。

3.3.2. 核心点

针对核心点，我们决定使用 SMOTE 算法合成少数类样本。SMOTE 将生成新的样本，以增加核心点聚类的样本数量，从而提高模型性能。

SMOTE algorithm steps：

- 遍历少数类样本的核心点并依次选择每个核心点样本 x_i，计算从 x_i 到其他核心点样本的欧氏距离，以获取其 k 个最近邻。
- 随机选择一些最近邻核心点样本，通过线性插值与核心样本 x_i 生成一个新样本。

$$x_{new} = x_i + \left(x_j - x_i\right) \tag{1}$$

其中 θ 是在(0, 1)范围内的随机数。

3.3.3. 不含多数类样本的边界点

针对不含多数类样本的边界点，我们选择采用 ADASYN 算法，它是一种自适应的过抽样技术。ADASYN 会更倾向于生成那些靠近决策边界的样本，以更好地捕捉边界点周围的信息[15]，提高模型对这些样本的性能。

3.3.4. 含多数类样本的边界点

对于含多数类样本的边界点，我们决定不进行额外处理，因为这些样本可能已经被主要的多数类样本覆盖，而不需要过多的合成或处理。

3.3.5. 噪音点

噪音点可能是异常值或者不符合主要模式的样本，因此我们决定不进行任何处理，保留它们的原始状态。它有助于维持数据集的完整性，同时专注于处理核心点和那些真正需要关注的边界点。

3.3.6. 平衡胎儿数据集

我们将通过各种处理方法生成的新样本点合并在一起，以创建一个全新的平衡数据集。这个平衡数据集将包括原始数据集中的多数类样本，以及经过 SMOTE 或其他合成方法处理的核心点，以及经过 ADASYN 处理的不含多数类样本的边界点。通过合并这些样本，我们确保了数据集的平衡性，使得多数类和少数类之间的样本数量差距被显著减小。

通过上述处理步骤，我们能够更好地处理多数类和少数类样本之间的不平衡问题，从而有效地提升了模型的整体性能。

3.4. 评估指标

混淆矩阵(Confusion Matrix)是用于评估二元分类模型性能的重要工具。在混淆矩阵中，我们计算以

下四个基本的性能指标：

- TP (True Positive)是真正例，表示正类样本中被正确预测为正类的数量。
- FN (False Negative)是假反例，表示正类样本中被错误预测为负类的数量。
- FP (False Positive)是假正例，表示负类样本中被错误预测为正类的数量。
- TN (True Negative)是真反例，表示负类样本中被正确预测为负类的数量。

混淆矩阵(Confusion Matrix)是用于评估二元分类模型性能的重要工具。在混淆矩阵中，我们计算以下四个基本的性能指标：

3.4.1. 准确率

预测正确的样本占样本总数的比例称为模型准确率。

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{2}$$

3.4.2. 精确率

精确率是用于评估分类模型性能的指标之一，它衡量了分类器评估为正例的样本中，实际为正例的比例。通常情况下，精确率越高，表示分类器在评估为正例的样本中，其准确率越高。精确率高意味着分类器更有能力准确地识别出真正的正例，而减少了将负例错误分为正例的情况。该指标可进行如下计算：

$$precision = \frac{TP}{TP + FN} \tag{3}$$

3.4.3. 召回率

召回率表示分类器正确评估的样本占所有阳性样本的比例，召回率越高则说明模型能够正确预测出的样本越多，模型的效果越好。

$$recall = \frac{TP}{TP + FN} \tag{4}$$

3.4.4. F1-score

F1 值(F1-score)是一种用来衡量分类模型精确度的指标，它通过综合考虑精确率和召回率来评估模型的性能，特别适用于处理不均衡数据。F1-score 的取值范围在 0 到 1 之间，越接近 1 表示模型性能越好。

$$F1 - Score = \frac{2 * precision * recall}{precision + recall} \tag{5}$$

通常情况下，上述四个性能指标值越高表示模型的分类性能较好，然而，在不平衡数据集中，模型可能由于产生过多的假阴性(False Negatives，FN)而表现出偏差。这意味着模型可能会漏掉一些真正属于少数类别的样本，导致对少数类的识别能力不足。因此，在评估不平衡数据集时，需要特别关注模型的召回率和其他适当的性能指标，以确保模型在关注的类别上表现良好。

4. 结果分析

我们通过分析模型的分类报告来评估预测模型的准确性，使用了准确率(Accuracy, ACC)、精确率(Precision, PRE)、召回率(Recall)、F1 值等评价指标来验证模型的性能。在研究过程中，我们使用了八个不同的分类器，分别是 Adaptive boosting (AdaBoost)，Categorical boosting (CatBoost)，Decision tree (DT)，Gradient Boosting (GB)，K-nearest neighbor (KNN)，random forest (RF)，Support vector machine (SVM)，

Extreme Gradient Boosting (XGBoost)，用于诊断胎儿的健康状况。此外，我们还获得了每个分类器的混淆矩阵(Confusion Matrix, CM)以更详细地评估其性能。

4.1. Smote 过采样胎儿健康分类

图 3 详细介绍了使用各种技术实现的混淆矩阵(CM)。CM 用于对胎儿健康状况进行三元分类，N、S 和 P 分别代表正常、可疑和病理。表 1 展示了八种传统机器学习算法的性能分析报告，其中 Predicted 表示模型预测的结果，Actual 表示实际的结果。

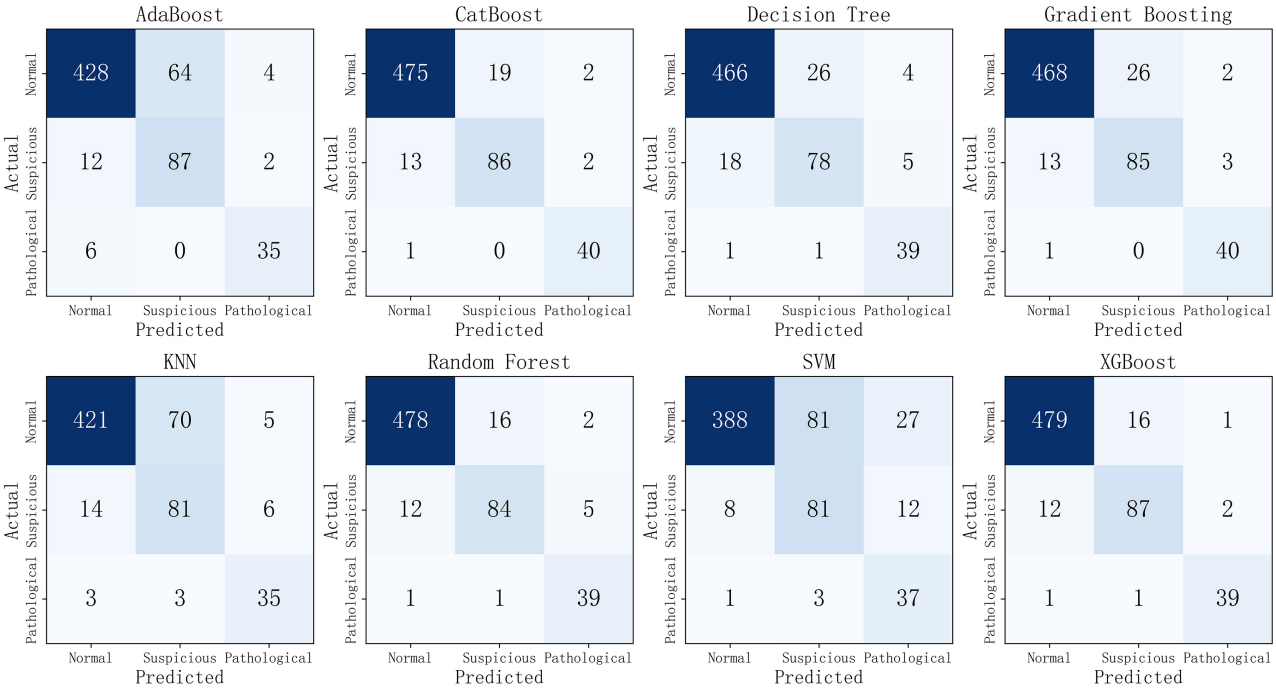

Figure 3. Comparison and analysis of machine learning models based on confusion matrix with SMOTE oversampling
图 3. 基于 SMOTE 过采样的混淆矩阵对机器学习模型进行比较和分析

Table 1. Analysis of fetal state performance of eight machine learning models under SMOTE oversampling
表 1. 八种机器学习模型在 SMOTE 过采样下胎儿状态表现的分析

MLmodel	Precision			Recall			F1-score			ACC
	N	S	P	N	S	P	N	S	P	
AdaBoost	0.960	0.576	0.854	0.863	0.861	0.854	0.909	0.690	0.854	0.862
CatBoost	0.971	0.819	0.909	0.958	0.851	0.976	0.964	0.835	0.941	0.942
Decision Tree	0.961	0.743	0.812	0.940	0.772	0.951	0.950	0.757	0.876	0.914
Gradient Boosting	0.971	0.766	0.889	0.944	0.842	0.976	0.957	0.802	0.930	0.929
KNN	0.961	0.526	0.761	0.849	0.802	0.854	0.901	0.635	0.805	0.842
Random Forest	0.974	0.832	0.848	0.964	0.832	0.951	0.969	0.832	0.897	0.942
SVM	0.977	0.491	0.487	0.782	0.802	0.902	0.869	0.609	0.632	0.793
XGBoost	0.974	0.837	0.929	0.966	0.861	0.951	0.970	0.849	0.940	0.948

在采用 SMOTE 过采样算法后，大多数模型在推理正常类别的方面表现较好，精确率普遍在 96.0%至 97.7%之间，然而，在推理可疑和病理类别方面，这些模型的性能普遍下降，可疑样本精准度在 49.1%至 83.7%之间，病理样本的精确率在 48.7%至 92.9%之间。简而言之，模型对这些样本的预测准确率较低。存在一些难以区分的可疑和病理样本，导致模型无法准确判断其是否为可疑和病理。在混淆矩阵中，分类错误的过多集中在，实际为正常样本预测为可疑样本和实际为可疑样本预测为正常样本上，模型容易将可疑和病理情况识别为正常情况，导致少数类数据过多地倾向于多数类数据，从而使模型更容易捕获大部分正例，但也导致了大量的假正例。从召回率数据中可以看出，不同模型在不同类别的召回率存在差异。在某些情况下，模型在正常类别上的召回率较高，但在可疑和病理类别上的召回率较低。模型的召回率通常在正常类别上最高，这可能是因为正常样本的数量较多，模型更容易检测到它们。F1 分数综合考虑了精确率和召回率，因此提供了更全面的性能评估。一般而言，F1 分数高的模型对于平衡精确率和召回率表现良好。可以看出 Random Forest 和 XGBoost 算法在大多数情况下表现较好，具有较高的 F1 分数。这八个机器学习模型的准确率普遍在 79.3%~94.8%之间，尤其在医疗领域中，这种表现被认为远未达到令人满意的程度。胎儿健康诊断要求极高的准确性和可靠性。错误的诊断可能导致误判胎儿的健康状况，从而延误治疗或给母婴带来不必要的风险。

XGBoost 的准确率最高(94.8%)，在可疑和病理状态下，它表现出最高的精确率，能够准确地识别出异常情况。相较于其他机器学习模型，XGBoost 在准确率、召回率和 F1 分数等关键评估指标上均普遍优于其他模型。然而，由于 CTG 解读主要依赖于产科医生对跟踪的分析，导致数据的解释具有主观性[16]。

4.2. 三重集成过采样过采样胎儿状态分类

图 4 显示了每个机器学习学习模型的混淆矩阵，表 2 显示了使用三重集成过采样算法执行不平衡处理后从每个分类器获得的模型性能指标。

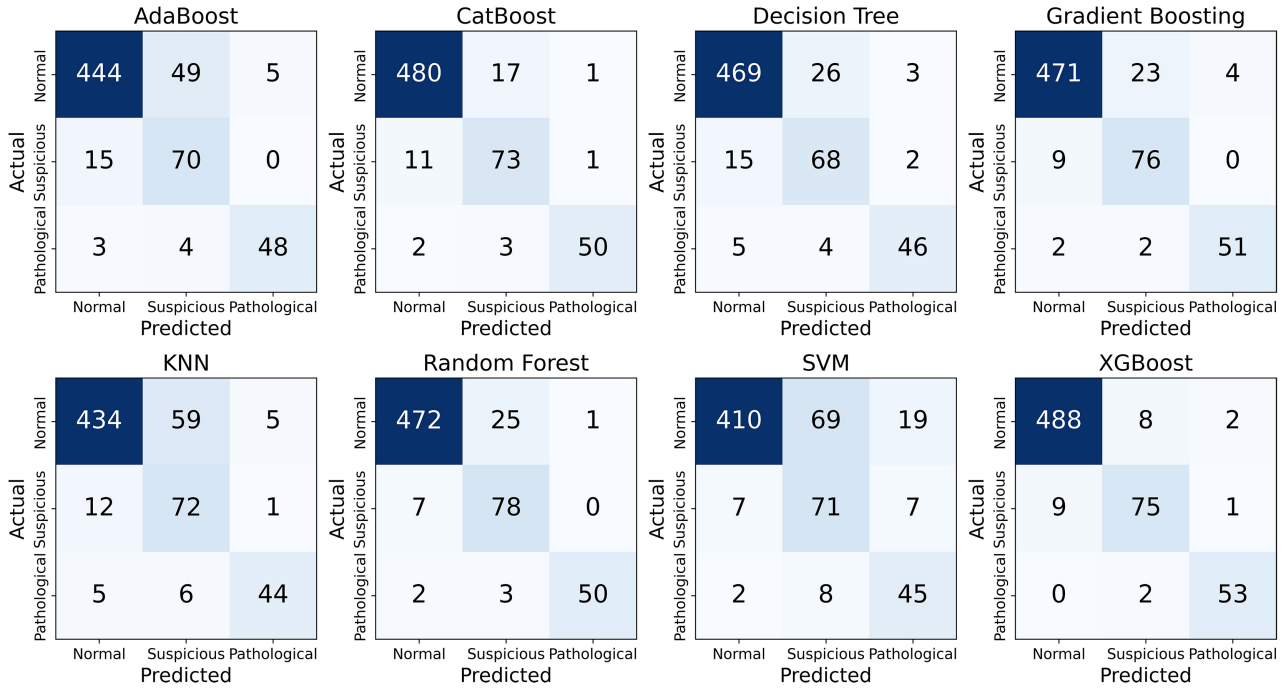

Figure 4. Comparison and analysis of machine learning models based on confusion matrix with Tri-Ensemble sampling
图 4. 基于三重集成过采样的混淆矩阵对机器学习模型进行比较和分析

Table 2. Analysis of fetal state performance of eight machine learning models under Tri-Ensemble sampling
表 2. 八种机器学习模型在三重集成过采样下胎儿状态表现的分析

ML model	precision			Recall			F1-score			ACC
	N	S	P	N	S	P	N	S	P	
AdaBoost	0.961	0.569	0.906	0.892	0.824	0.873	0.925	0.673	0.889	0.881
CatBoost	0.974	0.785	0.962	0.964	0.859	0.909	0.969	0.820	0.935	0.945
Decision Tree	0.959	0.694	0.902	0.942	0.800	0.836	0.950	0.743	0.868	0.914
Gradient Boosting	0.977	0.752	0.927	0.946	0.894	0.927	0.961	0.817	0.927	0.937
KNN	0.962	0.526	0.880	0.871	0.847	0.800	0.915	0.649	0.838	0.862
Random Forest	0.981	0.736	0.980	0.948	0.918	0.909	0.964	0.817	0.943	0.940
SVM	0.979	0.480	0.634	0.823	0.835	0.818	0.894	0.609	0.714	0.824
XGBoost	0.982	0.882	0.946	0.980	0.882	0.964	0.981	0.882	0.955	0.966

从总体上来看，采用三重集成过采样之后，准确率、精准度、召回率和 F1 分数都提升了 1%~2%的精度，这表明该策略在提高模型性能方面是有效的。具体来说，这些指标的提升表明模型在识别正常样本、病理样本以及可疑样本方面的性能都有所改进，能够更好地改善模型的性能。相交于 SMOTE 更有效，更能全面地处理了数据不平衡的挑战，提高了模型对正常样本、可疑样本和病理样本的分类能力。

5. 结束语

在医疗保健领域，解决不平衡数据分类问题的技术之一是 SMOTE 过采样技术。然而，现有的 SMOTE 算法存在一些局限性，它主要关注少数类样本的插值生成，而忽视了多数类样本的分布情况。这可能导致新生成的样本仍然位于类边界区域，进一步增加了类别边界的模糊性。然而，当少数类样本分布不均匀时，SMOTE 算法倾向于在密集区域生成更多的样本[17]，导致稀疏区域的少数类样本被忽视，从而影响分类准确性。针对这些挑战，我们提出了三重集成过采样算法。通过融合 DBSCAN、SMOTE 和 ADASYN 算法，该方法能够更好地处理少数类样本分布不均匀的情况，生成高质量且均衡的过采样数据。三重集成过采样算法通过对不同类型的样本点采取不同操作，解决了 SMOTE 算法存在的问题，为医疗数据挖掘提供了新的解决方案，进一步提高机器学习模型在高度不平衡的医疗数据上的性能，推动生物医学数据挖掘的发展。

参考文献

[1] Espinoza, I.S.R. (2020) Obstetric Hemorrhage, Its Role in Maternal Morbidity and Mortality and the Importance of Its Diagnosis, Prevention and Timely Management. *Mexican Journal of Medical Research ICSA*, **8**, 37-44. https://doi.org/10.29057/mjmr.v8i15.3906

[2] Chawla, N.V., Bowyer, K.W., Hall, L.O., *et al*. (2002) SMOTE: Synthetic Minority over-Sampling Technique. *Journal of Artificial Intelligence Research*, **16**, 321-357. https://doi.org/10.1613/jair.953

[3] Ester, M., Kriegel, H.P., Sander, J., *et al*. (1996) A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise. *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining*, Portland, 2-4 August 1996, 226-231.

[4] He, H., Bai, Y., Garcia, E.A., *et al*. (2008) ADASYN: Adaptive Synthetic Sampling Approach for Imbalanced Learning. 2008 *IEEE International Joint Conference on Neural Networks* (*IEEE World Congress on Computational Intelligence*), Hong Kong, 1-8 June 2008, 1322-1328.

[5] Alam, M.T., Khan, M.A.I., Dola, N.N., *et al*. (2022) Comparative Analysis of Different Efficient Machine Learning Methods for Fetal Health Classification. *Applied Bionics and Biomechanics*, **2022**, Article ID: 6321884. https://doi.org/10.1155/2022/6321884

[6] Potharaju, S.P., Sreedevi, M., Ande, V.K., *et al*. (2019) Data Mining Approach for Accelerating the Classification Accuracy of Cardiotocography. *Clinical Epidemiology and Global Health*, **7**, 160-164. https://doi.org/10.1016/j.cegh.2018.03.004

[7] Chamidah, N. and Wasito, I. (2015) Fetal State Classification from Cardiotocography Based on Feature Extraction Using Hybrid K-Means and Support Vector Machine. 2015 *International Conference on Advanced Computer Science and Information Systems* (*ICACSIS*), Depok, 10-11 October 2015, 37-41. https://doi.org/10.1109/ICACSIS.2015.7415166

[8] Akmal, H., Hardalaç, F. and Ayturan, K. (2023) A Fetal Well-Being Diagnostic Method Based on Cardiotocographic Morphological Pattern Utilizing Autoencoder and Recursive Feature Elimination. *Diagnostics*, **13**, Article 1931. https://doi.org/10.3390/diagnostics13111931

[9] Kaliappan, J., Bagepalli, A.R., Almal, S., *et al*. (2023) Impact of Cross-Validation on Machine Learning Models for Early Detection of Intrauterine Fetal Demise. *Diagnostics*, **13**, Article 1692. https://doi.org/10.3390/diagnostics13101692

[10] Piri, J., Mohapatra, P. and Dey, R. (2021) Multi-Objective Ant Lion Optimization Based Feature Retrieval Methodology for Investigation of Fetal Well Being. 2021 *Third International Conference on Inventive Research in Computing Applications* (*ICIRCA*), Coimbatore, 2-4 September 2021, 1732-1737. https://doi.org/10.1109/ICIRCA51532.2021.9544860

[11] Chen, Y., Guo, A., Chen, Q., *et al*. (2021) Intelligent Classification of Antepartum Cardiotocography Model Based on Deep Forest. *Biomedical Signal Processing and Control*, **67**, Article ID: 102555. https://doi.org/10.1016/j.bspc.2021.102555

[12] Campos, D. and Bernardes, J. (2010) Cardiotocography. UCI Machine Learning Repository.

[13] Thabtah, F., Hammoud, S., Kamalov, F., *et al*. (2020) Data Imbalance in Classification: Experimental Evaluation. *Information Sciences*, **513**, 429-441. https://doi.org/10.1016/j.ins.2019.11.004

[14] Beattie, J.R. and Esmonde-White, F.W.L. (2021) Exploration of Principal Component Analysis: Deriving Principal Component Analysis Visually Using Spectra. *Applied Spectroscopy*, **75**, 361-375. https://doi.org/10.1177/0003702820987847

[15] Islam, Z., Abdel-Aty, M., Cai, Q., *et al*. (2021) Crash Data Augmentation Using Variational Autoencoder. *Accident Analysis & Prevention*, **151**, Article ID: 105950. https://doi.org/10.1016/j.aap.2020.105950

[16] Sahin, H. and Subasi, A. (2015) Classification of the Cardiotocogram Data for Anticipation of Fetal Risks Using Machine Learning Techniques. *Applied Soft Computing*, **33**, 231-238. https://doi.org/10.1016/j.asoc.2015.04.038

[17] Douzas, G., Bacao, F. and Last, F. (2018) Improving Imbalanced Learning through a Heuristic Oversampling Method Based on K-Means and SMOTE. *Information Sciences*, **465**, 1-20. https://doi.org/10.1016/j.ins.2018.06.056