



(12)发明专利申请

(10)申请公布号 CN 108986869 A

(43)申请公布日 2018. 12. 11

(21)申请号 201810833695.X

(22)申请日 2018.07.26

(71)申请人 南京群顶科技有限公司

地址 210019 江苏省南京市建邺区奥体大街68号新城科技园国际研发总部园1幢6-8层

(72)发明人 周镶玉 徐磊 张永磊 张琳琳

(74)专利代理机构 南京知识律师事务所 32207

代理人 张苏沛

(51)Int.Cl.

G11C 29/10(2006.01)

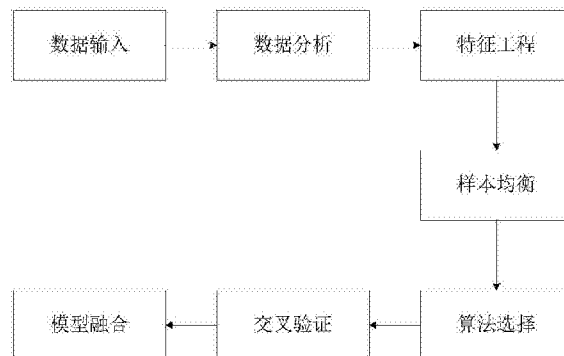
权利要求书2页 说明书6页 附图2页

(54)发明名称

一种使用多模型预测的磁盘故障检测方法

(57)摘要

本发明公开一种使用多模型预测的磁盘故障检测方法,通过时序数据处理手段提取磁盘SMART指标的多种特征,建立分类模型以预测磁盘状态;步骤一,数据输入:获取若干磁盘在一段时间内的监测数据构成的数据集;步骤二,SMART筛选:采用突变点检测的方式来选择SMART指标;步骤三,特征工程:SMART指标作为自定义特征提取模块的输入,提取SMART指标的特征,进而抽取出相应的参数配置,将参数配置作为参数传入特征提取模块,用以抽取训练集和测试集的特征集合;步骤四,平衡数据集:采用降维聚类,对占比较大的负样本进行降采样;步骤五,算法选择及建模:在步骤四的基础上训练分类模型,并测试当前磁盘属于正常状态或者需要被替换的故障状态。



1. 一种使用多模型预测的磁盘故障检测方法,其特征在于:通过时序数据处理手段提取磁盘SMART指标的多种特征,建立多个分类模型以预测磁盘状态;包括以下步骤:

步骤一:数据输入

获取若干磁盘在一段时间内的监测数据构成的数据集,磁盘故障数据标记为正样本,磁盘正常数据标记为负样本;所述监控数据中包括SMART指标;

步骤二:SMART筛选

采用突变点检测的方式来选择SMART指标作为后续特征提取模块的输入,将突变点定义为那些发生不可恢复的永久性突变的时刻点;

步骤三:特征工程

SMART指标作为自定义特征提取模块的输入,特征提取模块提取SMART指标的特征并对特征进行筛选得到有效的特征集;然后根据特征集从中抽取相应的特征提取模块参数配置,得到特征提取参数集;最后将参数配置作为参数传入特征提取模块,用以抽取训练集和测试集的特征集合;

步骤四:平衡数据集

针对负样本,在步骤三抽取特征后的数据集上,先采用降维之后分析负样本分布;再通过聚类,并根据聚类结果分布对负样本进行降采样;

步骤五:算法选择及建模

在步骤四的基础上选择算法分别训练分类模型;针对每个分类模型使用交叉验证的方式验证单个模型的分类性能;最后,对选择的多个算法的分类模型进行融合得到优化的融合模型,并使用融合模型测试当前磁盘属于正常状态或者需要被替换的故障状态。

2. 根据权利要求1所述的使用多模型预测的磁盘故障检测方法,其特征在于:

监控数据中还包括:时间戳、磁盘型号、序列号、存储容量字段。

3. 根据权利要求1所述的使用多模型预测的磁盘故障检测方法,其特征在于:

步骤二中,针对不同的SMART指标,通过假设检验,判断突变点是否为不可恢复的永久性突变,继而通过计算正样本中发生突变的磁盘比例,来比较不同的SMART指标的重要性,正样本中发生突变的磁盘比例高的SMART指标重要性高于正样本中发生突变的磁盘比例低的SMART指标重要性。

4. 根据权利要求1所述的使用多模型预测的磁盘故障检测方法,其特征在于,步骤三进一步包括如下步骤:

步骤31,故障磁盘、正常磁盘分别选择一段时间的历史SMART数据作为正样本、负样本的原始特征,用于后续的特征提取;

步骤32,将不同类型的SMART指标进行分组,对不同组SMART指标分别处理;

步骤33,经过分组处理后的SMART指标作为自定义特征提取模块的输入;通过交叉验证选出最具代表性的特征集合,并提取出该特征集合的特征提取模块参数配置,用以对测试集做相同的特征提取工作。

5. 根据权利要求1所述的使用多模型预测的磁盘故障检测方法,其特征在于:

步骤四中,针对负样本,在步骤三抽取特征后的数据集上,首先进行降维处理,若降维后有明显聚类则记录类别个数为k;选取距离k个聚类中心最近的磁盘数据作为训练集的负样本,进行聚类处理,抽取不同分布情况下的正常磁盘数据。

6. 根据权利要求1所述的使用多模型预测的磁盘故障检测方法,其特征在于:

为评估单个模型的分类性能,进行k折交叉验证;将数据集划分为若干个大小相似的互斥子集,每个子集保持数据分布的一致性;然后,每次用其中的1个子集作为测试集,余下的子集的并集作为训练集;以此得到若干组训练/测试集,从而可进行若干次训练和测试,最终以测试结果的均值作为最后得分,评估该模型的性能。

7. 根据权利要求1所述的使用多模型预测的磁盘故障检测方法,其特征在于:

采用Stacking模型集成方法,对多个二分类模型进行模型融合,得到一个优化的融合模型;

Stacking先从初始数据集训练出初级学习器,生成一个新数据集用于训练次级学习器,在产生新数据集时采用k折交叉验证;

初始训练集D被划分为k个大小相似的集合 D_1, D_2, \dots, D_k ;

令 D_j 和 $\bar{D}_j = D \setminus D_j$ 分别表示第j折的测试集和训练集;

给定T个初级学习算法,初级学习器 $Estimator_t^j$ 通过在 \bar{D}_j 上使用第t个学习算法而得;

再用训练得到的初级学习器预测 D_j 产生一个预测结果 $Predictions_t^j$,于是,在整个交叉验证过程结束后,从第t个初级学习器产生的次级训练集是

$D_t' = \{Predictions_t^1, Predictions_t^2, \dots, Predictions_t^k\}$,综合T个初级学习器产生的次级训练集,得到 $D' = \{D'_1, D'_2, \dots, D'_T\}$ 将用于学习次级学习器;

相应地,要产生测试集需要在预测 $Predictions_t^j$ 的同时也对测试集数据进行一次预测,得到 $Test_pred_t^j$,k折交叉验证结束后,对测试集的预测产生k个预测结果

$Test_pred_t' = \{Test_pred_t^1, Test_pred_t^2, \dots, Test_pred_t^k\}$,综合T个学习算法,可得最后的测试集为 $Test_pred'$ 。

一种使用多模型预测的磁盘故障检测方法

技术领域

[0001] 本发明涉及数据挖掘领域,具体涉及一种磁盘故障预测算法。

背景技术

[0002] 近年来,随着云存储等新兴技术的发展,海量数据存储技术发展越来越快,数据中心日益庞大。在数据中心的停机时间成本中,网络设备故障是重要因素,且磁盘作为数据最终存储的地方,是最重要的网络设备之一,也是最常发生故障的设备。随着数据的重要性不断增加,以磁盘为基础的存储设备故障所造成的事故影响越来越大,数据的恢复成本越来越高。

[0003] 磁盘故障一般分为两种:可预测的(predictable)和不可预测的(unpredictable)。不可预测的故障如芯片突然失效等瞬时故障,但像电机轴承磨损、盘片磁介质性能下降等都有一个过程,属于可预测的情况,如果发生这种问题,则需要提前发出告警,让使用者有足够的时间把重要资料转移到其它储存设备上。当磁盘发生故障时虽然可以通过副本或纠删码等机制来提高可靠性,但这些都属于硬盘故障后的数据恢复技术,是被动容错方式。且若多块磁盘出现故障,则会产生无法弥补的损失。

[0004] 针对主动容错方式,现在的硬盘厂商均遵循SMART标准。SMART(Self-Monitoring Analysis and Reporting Technology),自我监测、分析及报告技术,是一种自动的硬盘状态检测与预警系统和规范。通过对硬盘的各项指标信息进行监控并与厂商所设定的安全阈值进行比较,若超出阈值就自动向用户做出警告。但是基于阈值的故障预测方式太过简单,准确率较低且无法提前预测磁盘故障。

[0005] 现有的基于机器学习算法的故障检测方法,并没有充分的提取SMART指标的时序性特征。

发明内容

[0006] 针对现有技术方案存在的缺陷,本发明的目的是提供一种查准率查全率高,整体性能好的从时序数据特性的角度充分提取SMART特征并自动筛选有效特征的方法,用于磁盘故障预测。

[0007] 为达到上述目的,本发明采用的技术方案为:

[0008] 本发明提供的使用多模型预测的磁盘故障检测方法,是指通过时序数据处理手段提取磁盘SMART指标的多种特征,利用机器学习算法和相关理论,建立多个分类模型以预测磁盘状态,并根据Stacking方法对多个分类模型进行融合,有效提升了模型的准确性与稳定性。本发明的检测方法包括以下关键技术环节:

[0009] 步骤一:数据输入

[0010] 获取若干磁盘在一段时间内的监测数据构成的数据集,磁盘故障数据标记为正样本(包括磁盘停止工作、对指令无响应、或RAID系统报告不可读写等情况),磁盘正常数据标记为负样本。监控数据中包括:时间戳、磁盘型号、序列号、存储容量、SMART指标。

[0011] 步骤二:SMART筛选

[0012] 采用突变点检测的方式来选择SMART指标作为后续特征提取模块的输入,这里将突变点定义为那些发生不可恢复的永久性突变的时刻点。

[0013] 进一步的,针对不同的SMART指标,通过假设检验,判断突变点是否为不可恢复的永久性突变,继而通过计算正样本中发生突变的磁盘比例,来比较不同的SMART指标的重要性,正样本中发生突变的磁盘比例高的SMART指标重要性高于正样本中发生突变的磁盘比例低的SMART指标重要性。

[0014] 步骤三:特征工程

[0015] SMART指标是由传感器在一定时间间隔内采集的数据,因此每个SMART指标都是一条时间序列。磁盘状态不仅与当前SMART指标值有关,还与历史状态相关,因此需要一种紧凑的且包含尽可能多信息的表示方式。SMART指标作为自定义特征提取模块的输入,特征提取模块提取SMART指标的特征并对特征进行筛选得到有效的特征集;然后根据特征集从中抽取相应的特征提取模块参数配置,得到特征提取参数集;最后将参数配置作为参数传入特征提取模块,用以抽取训练集和测试集的特征集合;本发明中特征工程涉及以下环节:

[0016] 1.故障磁盘、正常磁盘分别选择一段时间的历史SMART数据作为正样本、负样本的原始特征,用于后续的特征提取;

[0017] 2.本发明中增加了对SMART的分析与分类,对不同类型的SMART指标进行分别处理。

[0018] 3.经过分组处理后提取的SMART指标更具有泛化表示能力,有助于提高模型的稳定性。例如对于累加值(如启停次数、重映射扇区数、累计通电时间等),作差分后取一段时间内的历史窗口或取最后一天的值;对于动态变化的值(如气流温度等)采用一段时间历史窗口内的原始值。作为自定义特征提取模块的输入;通过交叉验证选出最具代表性的特征集合,并提取出该特征集合的特征提取模块参数配置,用以对测试集做相同的特征提取工作。

[0019] 通过提取参数配置来对训练集和测试集做统一处理,可以减少特征维度从而减少过拟合,并且节约特征工程用时,提高效率。

[0020] 步骤四:平衡数据集

[0021] 由于正常磁盘的数据远远大于故障盘的数据,正样本占比极小,数据非常不平衡。因此为了平衡数据集,针对负样本(正常磁盘),在步骤三抽取特征后的数据集上,先采用降维之后分析负样本分布;再通过聚类,并根据聚类结果分布对占比较大的负样本进行降采样。

[0022] 首先进行PCA或tSNE等方法的降维处理,降至3维或2维后可视化,观察降维后的样本是否具有明显聚类,若有明显聚类则记录类别个数为k。采用KMeans无监督聚类等方法,选取距离k个聚类中心最近的磁盘数据作为训练集的负样本,可抽取不同分布情况(不同k值)下的正常磁盘数据,用这种方式选择的样本尽可能多地包含了负样本的信息,在平衡数据集的同时也降低了数据的冗余度。

[0023] 步骤五:算法选择及建模

[0024] 在步骤四的基础上使用梯度提升树及随机森林等算法训练分类模型,并测试当前磁盘属于正常状态或者需要被替换的故障状态。在步骤四的基础上选择算法分别训练分类

模型;针对每个分类模型使用交叉验证的方式验证模型的稳定性与准确性,根据交叉验证结果决定是否采用该算法;最后,对选择的多个算法的分类模型进行融合得到优化的融合模型,并使用融合模型测试当前磁盘属于正常状态或者需要被替换的故障状态。

[0025] 为评估单个模型的性能,进行k折交叉验证。将数据集划分为若干个大小相似的互斥子集,每个子集保持数据分布的一致性;然后,每次用其中的1个子集作为测试集,余下的子集的并集作为训练集;以此得到若干组训练/测试集,从而可进行若干次训练和测试,最终以这若干个测试结果的均值作为最后得分,并以此得分来评估该模型的性能。优选的,k折交叉验证可将数据集划分为5个大小相似的互斥子集。

[0026] 为了进一步提高模型的性能,本发明采用模型融合的方法。模型融合就是训练多个模型,然后将多个模型集成,最终融合得到一个性能最优的模型。本发明采用Stacking模型融合的方式来提高算法的精度和鲁棒性。

[0027] Stacking先从初始数据集训练出初级学习器,生成一个新数据集用于训练次级学习器,在产生新数据集时采用k折交叉验证的方式以减小过拟合风险。初始训练集D被划分为k个大小相似的集合 D_1, D_2, \dots, D_k 。令 D_j 和 $\bar{D}_j = D \setminus D_j$ 分别表示第j折的测试集和训练集。给定T个初级学习算法,初级学习器 $Estimator_t^j$ 通过在 \bar{D}_j 上使用第t个学习算法而得。再用训练得到的初级学习器预测 D_j 产生一个预测结果 $Predictions_t^j$,于是,在整个交叉验证过程结束后,从第t个初级学习器产生的次级训练集是 $D_t' = \{Predictions_1^j, Predictions_2^j, \dots, Predictions_k^j\}$,综合T个初级学习器产生的次级训练集,得到 $D' = \{D_1', D_2', \dots, D_T'\}$ 将用于学习次级学习器。相应地,要产生测试集需要在预测 $Predictions_t^j$ 的同时也对测试集数据进行一次预测,得到 $Test_pred_t^j$, k折交叉验证结束后,对测试集的预测产生k个预测结果 $Test_pred_t^j = \{Test_pred_t^1, Test_pred_t^2, \dots, Test_pred_t^k\}$,综合T个学习算法,可得最后的测试集为 $Test_pred'$ 。

[0028] 本发明具有以下有益效果:该方法能有效对磁盘SMART指标进行归因分析,选择有效原始特征;能够在正负样本不平衡的数据集中实现高准确率故障预测。

[0029] 1.相较于现有技术方案,本发明能够有效筛选出能够反映磁盘故障的SMART指标;

[0030] 2.增加了对SMART的分析与分类,对不同类型的SMART指标进行分别处理;

[0031] 3.利用特征工程高效利用了SMART时序数据,利用特征提取模块提取出更丰富有效的特征;特征提取模块可以提取百余种时序特征,并且针对在大量数据下利用该模块提取特征时的缺陷,提出了改进措施,以提高特征工程的效率;

[0032] 4.采用聚类降采样,解决了正负样本不平衡的问题;尽可能多地包含了负样本的信息,在平衡数据集的同时也降低了数据的冗余度;

[0033] 5.通过模型融合的方法,将多个模型集成,最终得到一个性能最优的模型。本发明采用Stacking模型融合的方式来提高算法的精度和鲁棒性。

附图说明

[0034] 图1为本发明实施例的使用多模型预测的磁盘故障预测方法流程图。

[0035] 图2为本发明实施例的使用多模型预测的磁盘故障预测方法特征工程处理流程图。

[0036] 图3为本发明实施例的使用多模型预测的磁盘故障预测方法5折交叉验证示意图。

[0037] 图4为本发明实施例的使用多模型预测的磁盘故障预测方法模型融合Stacking流程图。

具体实施方式

[0038] 为了便于本领域技术人员的理解,下面结合实施例与附图对本发明作进一步的说明。

[0039] 本实施例的磁盘故障预测方法,是指通过时序数据处理手段提取磁盘SMART指标的多种特征,利用机器学习算法和相关理论,建立二分类模型以预测磁盘状态。本实施例的磁盘故障预测算法流程如图1所示。包含以下关键技术环节:

[0040] 步骤一:数据收集

[0041] 优先采用Backblaze提供的数据集,该数据集中包含了超过30000块磁盘超过连续17个月的监测数据,当磁盘停止工作、对指令无响应、或RAID系统报告不可读写时,会被标记为正样本(即坏盘)。数据包含的字段有时间戳、磁盘型号、序列号、存储容量、SMART指标(好盘标记为0,坏盘标记为1)。

[0042] 虽然硬盘制造厂商均遵循SMART规范,但是厂商可以根据需要,使用不同的SMART组合或根据检测参数的多少增减SMART指标的数量,因此不同型号的磁盘之间,SMART指标的个数和组合有所区别,本实施例优先选择了数据集中占比最高的型号(ST4000DM000)作为训练和测试数据。

[0043] 步骤二:SMART筛选

[0044] 不同的厂家会使用不同的SMART指标,且部分SMART指标仅具有参考意义,不具备指向性,因此筛选出对硬盘故障有指示性的特征,及能够反映磁盘故障的指标尤为重要。

[0045] 本实施例采用突变点检测的方式来选择SMART指标作为后续特征提取模块的输入,这里将突变点定义为那些发生不可恢复的永久性突变的时刻点。针对不同的SMART指标,通过假设检验,判断突变点是否为不可恢复的永久性突变,继而通过计算正样本中发生突变的磁盘比例,来比较不同的SMART指标的重要性,比例越高,认为该SMART指标具有越丰富的判别信息。

[0046] 步骤三:特征工程

[0047] SMART指标是由传感器在一定时间间隔内采集的数据,因此每个SMART指标都是一条时间序列。磁盘状态不仅与当前SMART指标值有关,还与历史状态相关,因此需要一种紧凑的且包含尽可能多信息的表示方式。本实施例中特征工程涉及以下环节:

[0048] 1.由步骤二中突变点检测的过程中可知,坏盘中SMART指标发生突变的时刻几乎都在被替换之前的100天之内,因此为了利用突变点的丰富信息,选择坏盘寿命的倒数100天的历史数据作为正样本,好盘由于普遍使用时长较长,则可随机选择多个不同的起始点,同样选取100天的历史数据作为负样本。

[0049] 2.SMART指标包含几种不同类型的数据,有累加值如启停次数、重映射扇区数、累计通电时间等,有动态变化的值如气流温度等。本方法将不同类型的指标进行分组,对不同

组做不同的处理。例如对于累加值,作差分后取100天的历史窗口或取最后一天的值;对于动态变化的值采用100天历史窗口内的原始值。

[0050] 3.经过分组处理后的SMART指标中,作差分处理和取历史窗口原始值的指标将作为自定义特征提取模块的输入,特征提取模块提取了包括近似熵、自回归模型系数、自相关系数在内的几百余种特征。若对所有SMART指标抽取全量时序特征有如下缺陷:时间成本较高,并且抽取后的数据维度较高,容易过拟合。为克服这些缺陷,本实施例通过交叉验证来选出最具代表性的特征集合,并提取出该集合的参数配置,用以对测试集做相同的特征提取工作,具体操作如下:将整个数据集划分为训练集和测试集两个互斥子集。针对训练集,采用分层抽样将其分成若干部分(各部分正负样本比例相同),本实施例中将数据分为5个部分。首先进行初步特征提取:利用特征提取模块对每一部分数据进行时序特征的抽取,然后对提取的特征进行筛选;从经过初步提取的特征集中抽取出相应的特征提取模块的参数配置,可以得到5个不同的参数集,对这5个不同的参数集求取交集,以此作为特征提取模块的最终参数配置;并将此参数集作为参数传入特征提取模块,重新抽取训练集和测试集的特征集合。

[0051] 通过提取参数配置来对训练集和测试集做统一处理,可以减少特征维度从而减少过拟合,并且节约特征工程用时,提高效率。上述特征工程处理流程如图2所示。

[0052] 步骤四:平衡数据集

[0053] 由于正常磁盘的数据远远大于故障盘的数据,正样本占比极小,数据非常不平衡。因此为了平衡数据集,针对负样本(正常磁盘),在步骤三抽取特征后的数据集上,首先进行PCA或tSNE降维处理,降至3维或2维后可可视化,观察降维后的样本是否具有明显聚类,若有明显聚类则记录类别个数为k。采用KMeans无监督聚类,选取距离k个聚类中心最近的磁盘数据作为训练集的负样本,可抽取不同分布情况(不同k值)下的正常磁盘数据,用这种方式选择的样本尽可能多地包含了负样本的信息,在平衡数据集的同时也降低了数据的冗余度。

[0054] 步骤五:算法选择及建模

[0055] 在步骤四的基础上使用梯度提升树及随机森林等算法训练分类模型,并测试当前磁盘属于正常状态或者需要被替换的故障状态。

[0056] 为评估单个模型的性能,进行k折交叉验证。将数据集划分为5个大小相似的互斥子集,每个子集保持数据分布的一致性,即采用分层抽样,保持正负样本的比例一致。然后,每次用其中的4个子集的并集作为训练集,余下的1个子集作为测试集;以此得到5组训练/测试集,从而可进行5次训练和测试,最终以这5个测试结果的均值作为最后得分,并以此得分来评估该模型的性能。单模型交叉验证流程如图3所示。

[0057] 为了进一步提高模型的性能,本实施例采用模型融合的方法。模型融合就是训练多个模型,然后按照一定的方法集成,最终得到一个性能最优的模型,属于集成学习的方法。集成学习的一般结构:先产生一组“个体学习器”,再用某种策略将它们结合起来。结合策略一般有:从结果文件中融合、Stacking和Blending。最简单便捷的方式就是从结果文件中进行融合,这样做不需要重新训练模型,只需把不同模型的测试结果按照某种策略,例如投票表决、加权平均等方式得出一个最终结果。而Stacking和Blending是更为强大的结合策略,叫做“学习法”,即通过另一个学习器来进行结合。学习法中,把个体学习器称作

初级学习器,用于结合的学习器叫做次级学习器或元学习器,本实施例选用Stacking作为结合策略。

[0058] Stacking先从初始数据集训练出初级学习器,生成一个新数据集用于训练次级学习器,在产生新数据集时采用k折交叉验证的方式以减小过拟合风险。初始训练集D被划分为k个大小相似的集合 D_1, D_2, \dots, D_k 。令 D_j 和 $\bar{D}_j = D \setminus D_j$ 分别表示第j折的测试集和训练集。给定T个初级学习算法,初级学习器 $Estimator_t^j$ 通过在 \bar{D}_j 上使用第t个学习算法而得。再用训练得到的初级学习器预测 D_j 产生一个预测结果 $Predictions_t^j$,于是,在整个交叉验证过程结束后,从第t个初级学习器产生的次级训练集是 $D'_t = \{Predictions_t^1, Predictions_t^2, \dots, Predictions_t^k\}$,综合T个初级学习器产生的次级训练集,得到 $D' = \{D'_1, D'_2, \dots, D'_T\}$ 将用于学习次级学习器。相应地,要产生测试集需要在预测 $Predictions_t^j$ 的同时也对测试集数据进行一次预测,得到 $Test_pred_t^j$, k折交叉验证结束后,对测试集的预测产生k个预测结果 $Test_pred_t^j = \{Test_pred_t^1, Test_pred_t^2, \dots, Test_pred_t^k\}$,综合T个学习算法,可得最后的测试集为 $Test_pred'$ 。本实施例选择的初级学习器选用SVM、RandomForest和Xgboost这三种分类器,次级学习器选用LogisticRegression算法,针对每一个初级学习器,均采用5折交叉验证的方式产生用于训练次级学习器的新数据集。以一种分类器为例,Stacking方法的流程如图4所示。

[0059] 本方法具有如下关键点:

- [0060] 1. 步骤二中用突变点检测的方式筛选能够反映磁盘故障的原始SMART指标;
 - [0061] 2. 步骤三中利用SMART分组处理手段对不同类型的指标做特定处理;
 - [0062] 3. 步骤三中提出一种可以提取百余种时序特征的特征提取模块,并且针对在大量数据下利用该模块提取特征时的缺陷,提出了改进措施,以提高特征工程的效率;
 - [0063] 4. 步骤四采用降维聚类的方法,来对占比较大的负样本进行降采样,以解决数据集不平衡带来的困难,该方式能够选出低冗余,具有代表性的负样本;
 - [0064] 5. 步骤五中本实施例采用Stacking模型融合的方式来提高算法的精度和鲁棒性。
- [0065] 以上的实施例仅为说明本发明的技术思想,不能以此限定本发明的保护范围,凡是按照本发明提出的技术思想,在技术方案基础上所做的任何改动,均落入本发明保护范围之内。

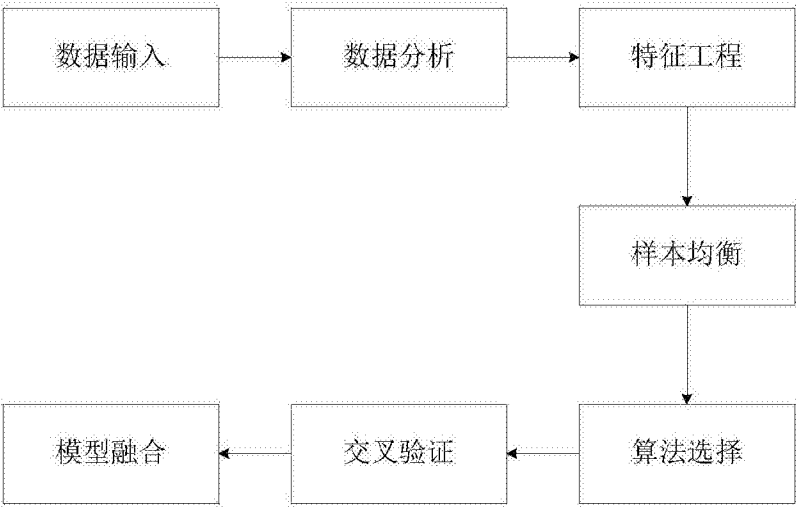


图1

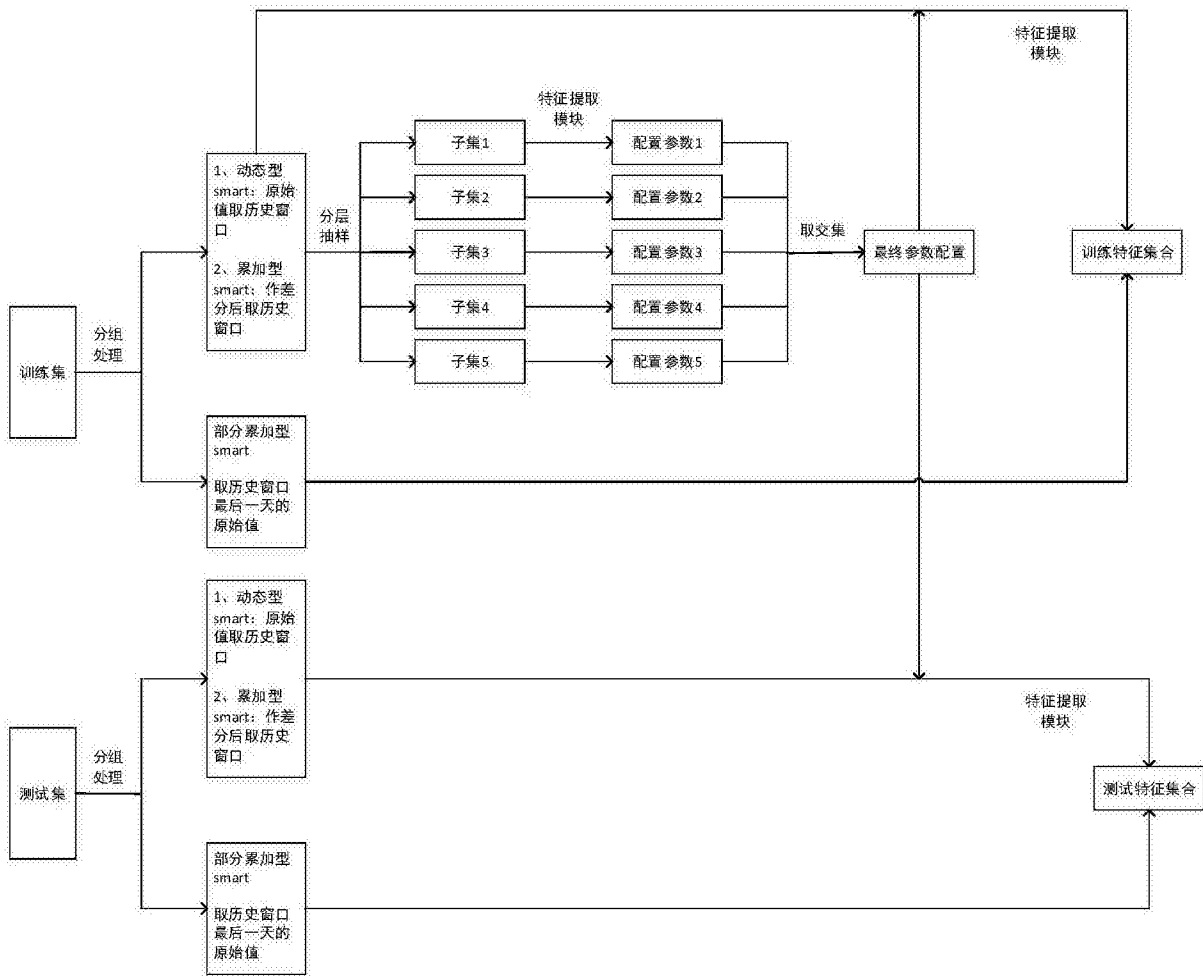


图2

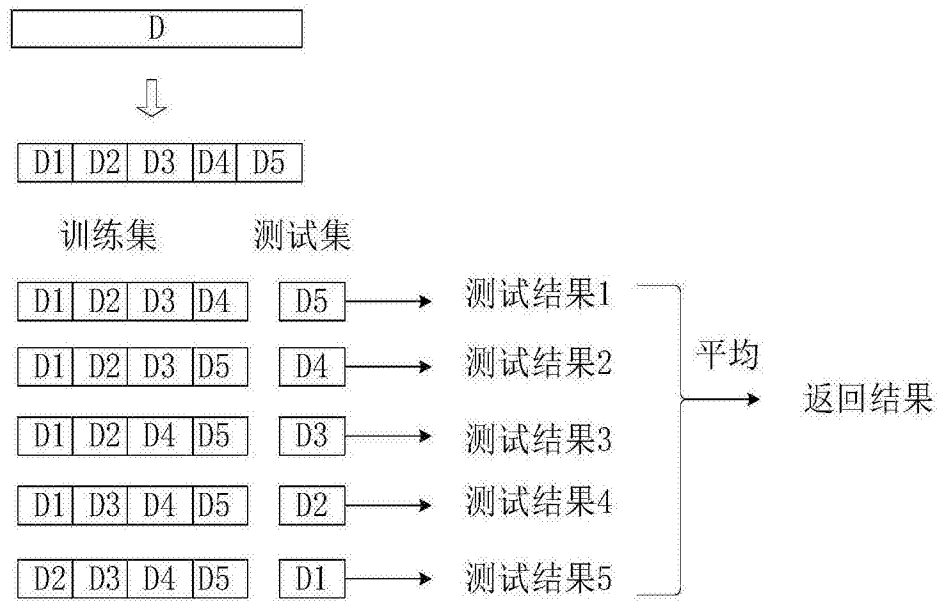


图3

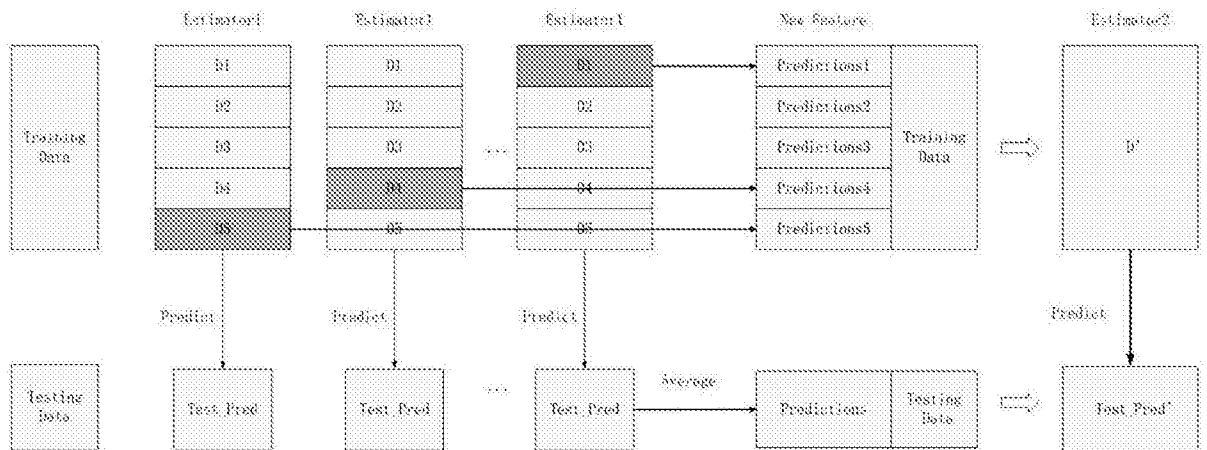


图4