



(12) 发明专利申请

(10) 申请公布号 CN 118037332 A

(43) 申请公布日 2024.05.14

(21) 申请号 202410276842.3

G06F 18/241 (2023.01)

(22) 申请日 2024.03.12

G06N 3/045 (2023.01)

(71) 申请人 湖北中巽泰科技有限公司

G06N 3/0475 (2023.01)

地址 430000 湖北省武汉市洪山区友谊大道与纺机路交汇处盈国际A座塔楼7-14层1471号

G06N 3/094 (2023.01)

G06N 5/025 (2023.01)

(72) 发明人 黄以槐 袁红艳

(74) 专利代理机构 广东皓行知识产权代理事务所(普通合伙) 441037

专利代理师 臧天雨

(51) Int. Cl.

G06Q 30/0201 (2023.01)

G06F 18/2135 (2023.01)

G06F 18/23213 (2023.01)

G06F 18/2337 (2023.01)

权利要求书4页 说明书11页 附图1页

(54) 发明名称

一种管理营销数据的数据处理方法及系统

(57) 摘要

本发明属于数据处理技术领域,公开了一种管理营销数据的数据处理方法及系统。所述的方法包括如下步骤:采集历史营销大数据,对历史营销大数据进行数据降维和标签添加,得到包含历史营销服务类型标签和历史营销效果等级标签的营销效果分类样本集;构建若干营销服务数据库、若干营销效果子数据库以及营销效果分类模型;采集实时营销数据,根据关键营销指标集合,进行数据降维,获取数据降维后实时营销数据的实时营销服务类型标签和的实时营销效果等级标签,并将数据降维后实时营销数据存储至对应的营销服务数据库的对应的营销效果子数据库中。本发明解决了现有技术存在的步骤简单、效果差以及不利于数据管理的的问题。

采集历史营销大数据,进行数据降维,并为每一数据降维后历史营销数据设置对应的历史营销服务类型标签和历史营销效果等级标签,得到营销效果分类样本集

构建若干营销服务数据库和若干营销效果子数据库,并根据营销效果分类样本集,构建营销效果分类模型

采集实时营销数据,进行数据降维,获取对应实时营销服务类型标签和实时营销效果等级标签,并将数据降维后实时营销数据存储至对应的营销服务数据库的对应的营销效果子数据库中

1. 一种管理营销数据的数据处理方法,其特征在于:包括如下步骤:

采集历史营销大数据,对历史营销大数据进行数据降维,得到对应的关键营销指标集合和数据降维后历史营销数据集,并为每一数据降维后历史营销数据设置对应的历史营销服务类型标签和历史营销效果等级标签,得到包含若干历史营销服务类型标签和若干历史营销效果等级标签的营销效果分类样本集;

根据若干历史营销服务类型标签和若干历史营销效果等级标签,构建对应的若干营销服务数据库和若干营销效果子数据库,将每一营销效果分类样本存储至对应的营销服务数据库的对应的营销效果子数据库中,并根据营销效果分类样本集,使用深度学习算法,构建对应的营销效果分类模型;

采集实时营销数据,根据关键营销指标集合,对实时营销数据进行数据降维,得到对应的数据降维后实时营销数据,获取数据降维后实时营销数据的实时营销服务类型标签,使用营销效果分类模型,获取数据降维后实时营销数据的实时营销效果等级标签,并根据实时营销服务类型标签和实时营销效果等级标签,将数据降维后实时营销数据存储至对应的营销服务数据库的对应的营销效果子数据库中。

2. 根据权利要求1所述的一种管理营销数据的数据处理方法,其特征在于:所述的历史营销大数据包括若干历史营销数据,所述的历史营销数据包括历史营销策略文本数据和历史营销指标数据;

所述的历史营销指标数据包括历史营销成本指标、历史营销利润指标、历史营销网站流量指标、历史营销社交媒体指标、历史营销转化率指标、历史营销客户获取成本指标、历史销售业绩指标、历史市场竞争分析指标以及历史客户满意度指标;

所述的实时营销数据包括实时营销策略文本数据和实时营销指标数据;

所述的实时营销指标数据包括实时营销成本指标、实时营销利润指标、实时营销网站流量指标、实时营销社交媒体指标、实时营销转化率指标、实时营销客户获取成本指标、实时销售业绩指标、实时市场竞争分析指标以及实时客户满意度指标。

3. 根据权利要求2所述的一种管理营销数据的数据处理方法,其特征在于:采集历史营销大数据,对历史营销大数据进行数据降维,得到对应的关键营销指标集合和数据降维后历史营销数据集,并为每一数据降维后历史营销数据设置对应的历史营销服务类型标签和历史营销效果等级标签,得到包含若干历史营销服务类型标签和若干历史营销效果等级标签的营销效果分类样本集,包括如下步骤:

采集历史营销大数据,并对历史营销大数据进行预处理,得到预处理后历史营销数据集;

使用PCA方法,对预处理后历史营销数据集进行数据降维,得到对应的关键营销指标集合和数据降维后历史营销数据集;

提取每一数据降维后历史营销数据的历史关键词集合,并将历史关键词集合作为对应的历史检索标签;

根据历史关键词集合,设置每一数据降维后历史营销数据的历史营销服务类型标签;

使用FCM聚类算法,对数据降维后历史营销数据集进行聚类处理,得到若干聚类中心和对应的若干聚类集群;

对每一聚类中心设置对应的历史营销效果等级标签,并将历史营销效果等级标签扩散

至聚类中心对应的聚类集群,得到设置有历史营销效果等级标签的若干营销效果分类样本;

整合所有营销效果分类样本,得到包含若干历史营销服务类型标签和若干历史营销效果等级标签的营销效果分类样本集。

4. 根据权利要求3所述的一种管理营销数据的数据处理方法,其特征在于:使用PCA方法,对预处理后历史营销数据集进行数据降维,得到对应的关键营销指标集合和数据降维后历史营销数据集,包括如下步骤:

使用PCA方法,对预处理后历史营销数据集中所有历史营销指标数据进行矩阵转换,得到由若干预处理后历史营销指标数据行向量构成的历史营销指标数据矩阵;

对历史营销指标数据矩阵进行标准化处理,得到标准化处理后历史营销指标数据矩阵;

获取标准化处理后历史营销指标数据矩阵的协方差矩阵,并根据协方差矩阵,获取对应的转换矩阵;

根据标准化处理后历史营销指标数据矩阵和对应的转换矩阵,获取由若干备选主成分列向量构成的主成分矩阵;

将方差累计贡献率前85%对应的备选主成分列向量作为主成分列向量,并将主成分列向量对应的历史营销指标作为关键营销指标,得到关键营销指标集合;

根据关键营销指标集合,对每一预处理后历史营销数据进行数据降维,得到由若干历史关键营销指标数据构成的数据降维后历史营销数据;

遍历预处理后历史营销数据集中的所有预处理后历史营销数据,得到对应的数据降维后历史营销数据集。

5. 根据权利要求3所述的一种管理营销数据的数据处理方法,其特征在于:提取每一数据降维后历史营销数据的历史关键词集合,并将历史关键词集合作为对应的历史检索标签,包括如下步骤:

使用TF-IDF-CI算法,提取数据降维后历史营销数据中历史营销策略文本数据的若干特征词的类型间离散因子和类内离散因子;

提取对应的数据降维后历史营销数据的特征词的词频和逆向文本频率;

根据类型间离散因子、类内离散因子、词频以及逆向文本频率,获取若干特征词的权重;

根据特征词的权重进行排序,选取前M个特征词,作为对应的数据降维后历史营销数据的历史关键词集合,其中,M为预设历史关键词总数;

将历史关键词集合进行拼接,得到对应的数据降维后历史营销数据的历史检索标签;

遍历数据降维后历史营销数据集中所有数据降维后历史营销数据,得到每一数据降维后历史营销数据的历史关键词集合和对应的历史检索标签。

6. 根据权利要求5所述的一种管理营销数据的数据处理方法,其特征在于:根据历史关键词集合,设置每一数据降维后历史营销数据的历史营销服务类型标签,包括如下步骤:

采集营销知识大数据,并对营销知识大数据进行预处理,得到预处理后营销知识数据集;

根据预处理后营销知识数据集,构建对应的营销知识图谱;

根据历史关键词集合,使用营销知识图谱,进行营销服务类型分析,得到每一数据降维

后历史营销数据对应的历史营销服务类型标签。

7. 根据权利要求3所述的一种管理营销数据的数据处理方法,其特征在於:使用FCM聚类算法,对数据降维后历史营销数据集进行聚类处理,得到若干聚类中心和对应的若干聚类集群,包括如下步骤:

对数据降维后历史营销数据集进行初始化,得到每一数据降维后历史营销数据对聚类中心的初始的隶属度;

根据初始的隶属度,使用FCM聚类算法,对数据降维后历史营销数据集进行聚类,得到若干初始的聚类中心;

根据当前的隶属度,获取拉格朗日乘数法,获取合并函数值和变化值,若合并函数值大于合并函数值阈值,或变化值大于变化值阈值,则对当前的聚类中心和当前的隶属度进行迭代更新,得到更新的聚类中心和更新的隶属度,并重复本步骤,否则,将当前的聚类中心作为最终的聚类中心,并进入下一步骤;

根据每一数据降维后历史营销数据与若干最终的聚类中心的欧氏距离,对数据降维后历史营销数据集进行划分,得到每一最终的聚类中心和对应的聚类集群。

8. 根据权利要求3所述的一种管理营销数据的数据处理方法,其特征在於:根据若干历史营销服务类型标签和若干历史营销效果等级标签,构建对应的若干营销服务数据库和若干营销效果子数据库,将每一营销效果分类样本存储至对应的营销服务数据库的对应的营销效果子数据库中,并根据营销效果分类样本集,使用深度学习算法,构建对应的营销效果分类模型,包括如下步骤:

根据A类历史营销服务类型标签,构建对应的A个营销服务数据库,其中,A为历史营销服务类型标签种类数量;

根据B类历史营销效果等级标签,在每一营销服务数据库中构建对应的B个营销效果子数据库,其中,B为历史营销效果等级标签种类数量;

根据每一营销效果分类样本的历史营销服务类型标签和历史营销效果等级标签,将营销效果分类样本存储至对应的营销服务数据库的对应的营销效果子数据库中;

根据营销效果分类样本集,使用N-GAN-Attention-MLP算法,构建对应的营销效果分类模型。

9. 根据权利要求6所述的一种管理营销数据的数据处理方法,其特征在於:采集实时营销数据,根据关键营销指标集合,对实时营销数据进行数据降维,得到对应的数据降维后实时营销数据,获取数据降维后实时营销数据的实时营销服务类型标签,使用营销效果分类模型,获取数据降维后实时营销数据的实时营销效果等级标签,并根据实时营销服务类型标签和实时营销效果等级标签,将数据降维后实时营销数据存储至对应的营销服务数据库的对应的营销效果子数据库中,包括如下步骤:

采集实时营销数据,根据关键营销指标集合,对实时营销数据进行数据降维,得到对应的数据降维后实时营销数据;

使用TF-IDF-CI算法,提取数据降维后实时营销数据的实时关键词集合,并将实时关键词集合作为对应的实时检索标签;

根据实时关键词集合,使用营销知识图谱,匹配对应的实时营销服务类型标签;

根据实时营销服务类型标签,将数据降维后实时营销数据发送至对应的营销服务数据

库；

使用营销效果分类模型,对数据降维后实时营销数据进行营销效果分类,得到对应的实时营销效果等级标签；

根据实时营销效果等级标签,将数据降维后实时营销数据存储至营销服务数据库的对应的营销效果子数据库中。

10.一种管理营销数据的数据处理系统,用于实现如权利要求1-9任一所述的数据处理方法,其特征在于:所述的系统包括依次连接的数据初始化处理单元、数据库构建单元以及实时营销数据处理单元；

数据初始化处理单元,用于采集历史营销大数据,对历史营销大数据进行数据降维,得到对应的关键营销指标集合和数据降维后历史营销数据集,并为每一数据降维后历史营销数据设置对应的历史营销服务类型标签和历史营销效果等级标签,得到包含若干历史营销服务类型标签和若干历史营销效果等级标签的营销效果分类样本集；

数据库构建单元,用于根据若干历史营销服务类型标签和若干历史营销效果等级标签,构建对应的若干营销服务数据库和若干营销效果子数据库,将每一营销效果分类样本存储至对应的营销服务数据库的对应的营销效果子数据库中,并根据营销效果分类样本集,使用深度学习算法,构建对应的营销效果分类模型；

实时营销数据处理单元,用于采集实时营销数据,根据关键营销指标集合,对实时营销数据进行数据降维,得到对应的数据降维后实时营销数据,获取数据降维后实时营销数据的实时营销服务类型标签,使用营销效果分类模型,获取数据降维后实时营销数据的实时营销效果等级标签,并根据实时营销服务类型标签和实时营销效果等级标签,将数据降维后实时营销数据存储至对应的营销服务数据库的对应的营销效果子数据库中。

## 一种管理营销数据的数据处理方法及系统

### 技术领域

[0001] 本发明属于数据处理技术领域,具体涉及一种管理营销数据的数据处理方法及系统。

### 背景技术

[0002] 随着互联网和大数据技术的迅猛发展,企业的营销数据呈现出爆炸式增长。营销数据对于企业进行精准营销具有重要意义,需要对于营销数据进行妥善管理,以便于企业对营销数据进行后续的分析与应用。

[0003] 现有技术中,仅仅对营销数据的进行格式转换、压缩以及解压等简单的数据处理步骤,无法解决营销数据存在的高维、复杂以及噪声等问题,数据处理效果差,导致营销数据的数据价值低和管理困难。

### 发明内容

[0004] 为了解决现有技术存在的步骤简单、效果差以及不利于数据管理的问题,本发明目的在于提供一种管理营销数据的数据处理方法及系统。

[0005] 本发明所采用的技术方案为:

一种管理营销数据的数据处理方法,包括如下步骤:

采集历史营销大数据,对历史营销大数据进行数据降维,得到对应的关键营销指标集合和数据降维后历史营销数据集,并为每一数据降维后历史营销数据设置对应的历史营销服务类型标签和历史营销效果等级标签,得到包含若干历史营销服务类型标签和若干历史营销效果等级标签的营销效果分类样本集;

根据若干历史营销服务类型标签和若干历史营销效果等级标签,构建对应的若干营销服务数据库和若干营销效果子数据库,将每一营销效果分类样本存储至对应的营销服务数据库的对应的营销效果子数据库中,并根据营销效果分类样本集,使用深度学习算法,构建对应的营销效果分类模型;

采集实时营销数据,根据关键营销指标集合,对实时营销数据进行数据降维,得到对应的数据降维后实时营销数据,获取数据降维后实时营销数据的实时营销服务类型标签,使用营销效果分类模型,获取数据降维后实时营销数据的实时营销效果等级标签,并根据实时营销服务类型标签和实时营销效果等级标签,将数据降维后实时营销数据存储至对应的营销服务数据库的对应的营销效果子数据库中。

[0006] 进一步地,历史营销大数据包括若干历史营销数据,历史营销数据包括历史营销策略文本数据和历史营销指标数据;

历史营销指标数据包括历史营销成本指标、历史营销利润指标、历史营销网站流量指标、历史营销社交媒体指标、历史营销转化率指标、历史营销客户获取成本指标、历史销售业绩指标、历史市场竞争分析指标以及历史客户满意度指标;

实时营销数据包括实时营销策略文本数据和实时营销指标数据;

实时营销指标数据包括实时营销成本指标、实时营销利润指标、实时营销网站流量指标、实时营销社交媒体指标、实时营销转化率指标、实时营销客户获取成本指标、实时销售业绩指标、实时市场竞争分析指标以及实时客户满意度指标。

[0007] 进一步地,采集历史营销大数据,对历史营销大数据进行数据降维,得到对应的关键营销指标集合和数据降维后历史营销数据集,并为每一数据降维后历史营销数据设置对应的历史营销服务类型标签和历史营销效果等级标签,得到包含若干历史营销服务类型标签和若干历史营销效果等级标签的营销效果分类样本集,包括如下步骤:

采集历史营销大数据,并对历史营销大数据进行预处理,得到预处理后历史营销数据集;

使用PCA方法,对预处理后历史营销数据集进行数据降维,得到对应的关键营销指标集合和数据降维后历史营销数据集;

提取每一数据降维后历史营销数据的历史关键词集合,并将历史关键词集合作为对应的历史检索标签;

根据历史关键词集合,设置每一数据降维后历史营销数据的历史营销服务类型标签;

使用FCM聚类算法,对数据降维后历史营销数据集进行聚类处理,得到若干聚类中心和对应的若干聚类集群;

对每一聚类中心设置对应的历史营销效果等级标签,并将历史营销效果等级标签扩散至聚类中心对应的聚类集群,得到设置有历史营销效果等级标签的若干营销效果分类样本;

整合所有营销效果分类样本,得到包含若干历史营销服务类型标签和若干历史营销效果等级标签的营销效果分类样本集。

[0008] 进一步地,使用PCA方法,对预处理后历史营销数据集进行数据降维,得到对应的关键营销指标集合和数据降维后历史营销数据集,包括如下步骤:

使用PCA方法,对预处理后历史营销数据集中所有历史营销指标数据进行矩阵转换,得到由若干预处理后历史营销指标数据行向量构成的历史营销指标数据矩阵;

对历史营销指标数据矩阵进行标准化处理,得到标准化处理后历史营销指标数据矩阵;

获取标准化处理后历史营销指标数据矩阵的协方差矩阵,并根据协方差矩阵,获取对应的转换矩阵;

根据标准化处理后历史营销指标数据矩阵和对应的转换矩阵,获取由若干备选主成分列向量构成的主成分矩阵;

将方差累计贡献率前85%对应的备选主成分列向量作为主成分列向量,并将主成分列向量对应的历史营销指标作为关键营销指标,得到关键营销指标集合;

根据关键营销指标集合,对每一预处理后历史营销数据进行数据降维,得到由若干历史关键营销指标数据构成的数据降维后历史营销数据;

遍历预处理后历史营销数据集中的所有预处理后历史营销数据,得到对应的数据降维后历史营销数据集。

[0009] 进一步地,提取每一数据降维后历史营销数据的历史关键词集合,并将历史关键

词集合作为对应的历史检索标签,包括如下步骤:

使用TF-IDF-CI算法,提取数据降维后历史营销数据中历史营销策略文本数据的若干特征词的类间离散因子和类内离散因子;

提取对应的数据降维后历史营销数据的特征词的词频和逆向文本频率;

根据类间离散因子、类内离散因子、词频以及逆向文本频率,获取若干特征词的权重;

根据特征词的权重进行排序,选取前M个特征词,作为对应的数据降维后历史营销数据的历史关键词集合,其中,M为预设历史关键词总数;

将历史关键词集合进行拼接,得到对应的数据降维后历史营销数据的历史检索标签;

遍历数据降维后历史营销数据集中所有数据降维后历史营销数据,得到每一数据降维后历史营销数据的历史关键词集合和对应的历史检索标签。

[0010] 进一步地,根据历史关键词集合,设置每一数据降维后历史营销数据的历史营销服务类型标签,包括如下步骤:

采集营销知识大数据,并对营销知识大数据进行预处理,得到预处理后营销知识数据集;

根据预处理后营销知识数据集,构建对应的营销知识图谱;

根据历史关键词集合,使用营销知识图谱,进行营销服务类型分析,得到每一数据降维后历史营销数据对应的历史营销服务类型标签。

[0011] 进一步地,使用FCM聚类算法,对数据降维后历史营销数据集进行聚类处理,得到若干聚类中心和对应的若干聚类集群,包括如下步骤:

对数据降维后历史营销数据集进行初始化,得到每一数据降维后历史营销数据对聚类中心的初始的隶属度;

根据初始的隶属度,使用FCM聚类算法,对数据降维后历史营销数据集进行聚类,得到若干初始的聚类中心;

根据当前的隶属度,获取拉格朗日乘数法,获取合并函数值 and 变化值,若合并函数值大于合并函数值阈值,或变化值大于变化值阈值,则对当前的聚类中心和当前的隶属度进行迭代更新,得到更新的聚类中心和更新的隶属度,并重复本步骤,否则,将当前的聚类中心作为最终的聚类中心,并进入下一步骤;

根据每一数据降维后历史营销数据与若干最终的聚类中心的欧氏距离,对数据降维后历史营销数据集进行划分,得到每一最终的聚类中心和对应的聚类集群。

[0012] 进一步地,根据若干历史营销服务类型标签和若干历史营销效果等级标签,构建对应的若干营销服务数据库和若干营销效果子数据库,将每一营销效果分类样本存储至对应的营销服务数据库的对应的营销效果子数据库中,并根据营销效果分类样本集,使用深度学习算法,构建对应的营销效果分类模型,包括如下步骤:

根据A类历史营销服务类型标签,构建对应的A个营销服务数据库,其中,A为历史营销服务类型标签种类数量;

根据B类历史营销效果等级标签,在每一营销服务数据库中构建对应的B个营销效果子数据库,其中,B为历史营销效果等级标签种类数量;

根据每一营销效果分类样本的历史营销服务类型标签和历史营销效果等级标签,将营销效果分类样本存储至对应的营销服务数据库的对应的营销效果子数据库中;

根据营销效果分类样本集,使用N-GAN-Attention-MLP算法,构建对应的营销效果分类模型。

[0013] 进一步地,采集实时营销数据,根据关键营销指标集合,对实时营销数据进行数据降维,得到对应的数据降维后实时营销数据,获取数据降维后实时营销数据的实时营销服务类型标签,使用营销效果分类模型,获取数据降维后实时营销数据的实时营销效果等级标签,并根据实时营销服务类型标签和实时营销效果等级标签,将数据降维后实时营销数据存储至对应的营销服务数据库的对应的营销效果子数据库中,包括如下步骤:

采集实时营销数据,根据关键营销指标集合,对实时营销数据进行数据降维,得到对应的数据降维后实时营销数据;

使用TF-IDF-CI算法,提取数据降维后实时营销数据的实时关键词集合,并将实时关键词集合作为对应的实时检索标签;

根据实时关键词集合,使用营销知识图谱,匹配对应的实时营销服务类型标签;

根据实时营销服务类型标签,将数据降维后实时营销数据发送至对应的营销服务数据库;

使用营销效果分类模型,对数据降维后实时营销数据进行营销效果分类,得到对应的实时营销效果等级标签;

根据实时营销效果等级标签,将数据降维后实时营销数据存储至营销服务数据库的对应的营销效果子数据库中。

[0014] 一种管理营销数据的数据处理系统,用于实现数据处理方法,系统包括依次连接的数据初始化处理单元、数据库构建单元以及实时营销数据处理单元;

数据初始化处理单元,用于采集历史营销大数据,对历史营销大数据进行数据降维,得到对应的关键营销指标集合和数据降维后历史营销数据集,并为每一数据降维后历史营销数据设置对应的历史营销服务类型标签和历史营销效果等级标签,得到包含若干历史营销服务类型标签和若干历史营销效果等级标签的营销效果分类样本集;

数据库构建单元,用于根据若干历史营销服务类型标签和若干历史营销效果等级标签,构建对应的若干营销服务数据库和若干营销效果子数据库,将每一营销效果分类样本存储至对应的营销服务数据库的对应的营销效果子数据库中,并根据营销效果分类样本集,使用深度学习算法,构建对应的营销效果分类模型;

实时营销数据处理单元,用于采集实时营销数据,根据关键营销指标集合,对实时营销数据进行数据降维,得到对应的数据降维后实时营销数据,获取数据降维后实时营销数据的实时营销服务类型标签,使用营销效果分类模型,获取数据降维后实时营销数据的实时营销效果等级标签,并根据实时营销服务类型标签和实时营销效果等级标签,将数据降维后实时营销数据存储至对应的营销服务数据库的对应的营销效果子数据库中。

[0015] 本发明的有益效果为:

本发明公开了一种管理营销数据的数据处理方法及系统,对营销数据进行数据降维,解决了营销数据存在的高维、复杂以及噪声的问题,并减小数据体量和内存占用空间;构建营销服务数据库和营销效果子数据库将营销数据按照营销服务类型和营销效果等级

进行分类存储,加强营销数据之间的数据交流的同时,便于对营销数据进行管理和检索,提高了营销数据的数据价值;营销效果分类模型对营销数据进行自动、高效的营销效果分类,提高了营销数据分类和存储的实用性和准确性。

[0016] 本发明的其他有益效果将在具体实施方式中进一步进行说明。

## 附图说明

[0017] 图1是本发明中管理营销数据的数据处理方法的流程框图。

[0018] 图2是本发明中管理营销数据的数据处理系统的结构框图。

## 具体实施方式

[0019] 下面结合附图及具体实施例对本发明做进一步阐释。

[0020] 实施例1:

如图1所示,本实施例提供一种管理营销数据的数据处理方法,包括如下步骤:

S1:采集历史营销大数据,对历史营销大数据进行数据降维,得到对应的关键营销指标集合和数据降维后历史营销数据集,并为每一数据降维后历史营销数据设置对应的历史营销服务类型标签和历史营销效果等级标签,得到包含若干历史营销服务类型标签和若干历史营销效果等级标签的营销效果分类样本集,包括如下步骤:

S1-1:采集历史营销大数据,并对历史营销大数据进行预处理,得到预处理后历史营销数据集;

历史营销大数据包括若干历史营销数据,历史营销数据包括历史营销策略文本数据和历史营销指标数据;

历史营销指标数据包括历史营销成本指标、历史营销利润指标、历史营销网站流量指标、历史营销社交媒体指标、历史营销转化率指标、历史营销客户获取成本指标、历史销售业绩指标、历史市场竞争分析指标以及历史客户满意度指标;

预处理包括对历史营销大数据进行数据格式转换、重复数据筛选以及错误数据剔除;

S1-2:使用主成分分析(Principal Component Analysis,PCA)方法,对预处理后历史营销数据集进行数据降维,得到对应的关键营销指标集合和数据降维后历史营销数据集,包括如下步骤:

S1-2-1:使用PCA方法,对预处理后历史营销数据集中所有历史营销指标数据进行矩阵转换,得到由若干预处理后历史营销指标数据行向量构成的历史营销指标数据矩阵

$X = [x_1, x_2, \dots, x_i, \dots, x_n]^T$ , 其中,  $x_i$  为第*i*个预处理后历史营销指标数据行向量,*i*为行向量指示量,*n*为预处理后历史营销数据总数;

S1-2-2:对历史营销指标数据矩阵进行标准化处理,得到标准化处理后历史营销指标数据矩阵;

公式为:

$$X' = \frac{(X - \mu)}{\sigma}$$

[0021] 式中,  $X'$  为标准化处理后历史营销指标数据矩阵;  $\mu$  为历史营销指标数据矩阵的均值;  $\sigma$  为历史营销指标数据矩阵的方差;

S1-2-3: 获取标准化处理后历史营销指标数据矩阵的协方差矩阵, 并根据协方差矩阵, 获取对应的转换矩阵;

公式为:

$$D = \frac{1}{n} Y Y^T = P \left( \frac{1}{n} X' X'^T \right) P^T = E^T \left( \frac{1}{n} X' X'^T \right) E$$

[0022] 式中,  $D$  为标准化处理后历史营销指标数据矩阵的协方差矩阵;  $Y$  为主成分矩阵;  $P$  为转换矩阵;  $E$  为单位特征向量矩阵;  $n$  为预处理后历史营销数据总数;

S1-2-4: 根据标准化处理后历史营销指标数据矩阵和对应的转换矩阵, 获取由若干备选主成分列向量构成的主成分矩阵  $Y = [y_1, y_2, \dots, y_l, \dots, y_L]$ ; 其中,  $y_l$  为第  $l$  备选主成分列向量,  $l$  为预处理后历史营销数据的数据维度指示量,  $L$  为数据维度总数;

公式为:

$$Y = P X'$$

[0023] 式中,  $Y$  为主成分矩阵;  $P$  为转换矩阵;  $X'$  为标准化处理后历史营销指标数据矩阵;

S1-2-5: 将方差累计贡献率前85%对应的备选主成分列向量作为主成分列向量, 并将主成分列向量对应的历史营销指标作为关键营销指标, 得到关键营销指标集合;

公式为:

$$G = \frac{\sum_{l=1}^K \lambda_l}{\sum_{l=1}^L \lambda_l}$$

[0024] 式中,  $G$  为方差累计贡献率;  $\lambda_l$  第  $l$  个备选主成分  $y_l$  的方差;  $l$  为数据维度指示量;  $L$  为数据维度总数;  $K$  为主成分总数;

主成分列向量  $\{y_1, y_2, \dots, y_l, \dots, y_K\}$  对应的历史营销指标作为关键营销指标;

S1-2-6: 根据关键营销指标集合, 对每一预处理后历史营销数据进行数据降维, 得到由若干历史关键营销指标数据构成的数据降维后历史营销数据;

S1-2-7: 遍历预处理后历史营销数据集中的所有预处理后历史营销数据, 得到对应的数据降维后历史营销数据集;

S1-3: 提取每一数据降维后历史营销数据的历史关键词集合, 并将历史关键词集合作为对应的历史检索标签, 包括如下步骤:

S1-3-1: 使用 (TF-IDF-CI, Term Frequency-Inverse Document Frequency-Class Information) 算法, 提取数据降维后历史营销数据中历史营销策略文本数据的若干特征词的类型离散因子和类内离散因子;

公式为:

$$CI_{ac} = \frac{2}{\pi} \arctan(S(t_j))$$

$$S(t_j) = \sqrt{\frac{\sum_{k=1}^{|C|} (TF(t_j, c_k) - \overline{TF}(t_j))^2}{|C| - 1}}$$

[0025] 式中,  $CI_{ac}$  为类间离散因子;  $S(t_j)$  为特征词  $t_j$  的标准差;  $c_k$  为特征词  $t_j$  的类别;  $|C|$  为类别总数;  $TF(t_j, c_k)$  为特征词  $t_j$  和在类别  $c_k$  中出现的频率;  $\overline{TF}(t_j)$  为  $t_j$  在各类别中出现的频率;  $t_j$  为特征词指示量;  $k$  为类别指示量;

$$CI_{ic} = \frac{2}{\pi} \arctan(s(t_j, c_k))$$

$$s(t_j, c_k) = TF(t_j, c_k) \cdot \frac{N(c_k)}{n(t_j, c_k)} - TF(t_j, c_k)$$

[0026] 式中,  $CI_{ic}$  为类内离散因子;  $s(t_j, c_k)$  为类别  $c_k$  中普遍出现的特征词  $t_j$  的数量;  $n(t_j, c_k)$  为包括特征词  $t_j$  的类别  $c_k$  的文档数量;  $N(c_k)$  为包括类别  $c_k$  的文档总数量;

S1-3-2: 提取对应的数据降维后历史营销数据的特征词的词频和逆向文本频率;

S1-3-3: 根据类间离散因子、类内离散因子、词频以及逆向文本频率, 获取若干特征词的权重;

公式为:

$$TF - IDF - DF = TF \times IDF \times \frac{1 - CI_{ic}}{1 - CI_{ac}}$$

[0027] 式中,  $TF - IDF - DF$  为特征词  $t_j$  的权重;  $TF$  为词频;  $IDF$  为逆向文本频率;  $CI$  为离散因子;

S1-3-4: 根据特征词的权重进行排序, 选取前  $M$  个特征词, 作为对应的数据降维后历史营销数据的历史关键词集合, 其中,  $M$  为预设历史关键词总数;

S1-3-5: 将历史关键词集合进行拼接, 得到对应的数据降维后历史营销数据的历史检索标签;

S1-3-6: 遍历数据降维后历史营销数据集中所有数据降维后历史营销数据, 得到每一数据降维后历史营销数据的历史关键词集合和对应的历史检索标签;

S1-4: 根据历史关键词集合, 设置每一数据降维后历史营销数据的历史营销服务类型标签, 包括如下步骤:

S1-4-1: 采集营销知识大数据, 并对营销知识大数据进行预处理, 得到预处理后营销知识数据集;

S1-4-2: 根据预处理后营销知识数据集, 构建对应的营销知识图谱, 包括如下步骤:

S1-4-2-1: 提取每一预处理后营销知识数据的若干命名实体;

S1-4-2-2: 提取每一预处理后营销知识数据的若干实体关系;

S1-4-2-3:根据若干命名实体和若干实体关系,进行知识融合,并构建对应的营销知识图谱;

S1-4-3:根据历史关键词集合,使用营销知识图谱,进行营销服务类型分析,得到每一数据降维后历史营销数据对应的历史营销服务类型标签;

S1-5:使用模糊C均值(Fuzzy C-Means,FCM)聚类算法,对数据降维后历史营销数据集进行聚类处理,得到若干聚类中心和对应的若干聚类集群,包括如下步骤:

S1-5-1:对数据降维后历史营销数据集进行初始化,得到每一数据降维后历史营销数据对聚类中心的初始的隶属度;

S1-5-2:根据初始的隶属度,使用FCM聚类算法,对数据降维后历史营销数据集进行聚类,得到若干初始的聚类中心;

S1-5-3:根据当前的隶属度,获取拉格朗日乘数法,获取合并函数值和变化值,若合并函数值大于合并函数值阈值,或变化值大于变化值阈值,则对当前的聚类中心和当前的隶属度进行迭代更新,得到更新的聚类中心和更新的隶属度,并重复本步骤,否则,将当前的聚类中心作为最终的聚类中心,并进入下一步骤;

公式为:

$$J_t = \sum_{i'=1}^m \sum_{j'=1}^c u_{i'j'}^\alpha \cdot d_{i'j'} + \sum_{i'=1}^m \lambda_{i'} \left( \sum_{j'=1}^c u_{i'j'}^\alpha - 1 \right)$$

$$\Delta J_t = J_t - J_{t-1}$$

[0028] 式中,  $J_t$ 、 $J_{t-1}$ 为第 $t$ 、 $t-1$ 时刻的合并函数值; $\Delta J_t$ 为变化值; $\lambda_{i'}$ 为第 $i'$ 数据降维后历史营销数据的特征值; $d_{i'j'}$ 为第 $i'$ 个数据降维后历史营销数据到第 $j'$ 个当前的聚类中心的距离; $u_{i'j'}$ 为第 $i'$ 数据降维后历史营销数据对于第 $j'$ 聚类中心的当前的隶属度; $\alpha$ 为超参数; $m$ 为数据总数; $c$ 为聚类中心总数; $j'$ 为聚类中心指示量; $i'$ 为数据降维后历史营销数据指示量;

$$\bar{u}_{i'j'} = \frac{1}{\sum_{k'=1}^c \left( \frac{d_{i'j'}}{d_{i'k'}} \right)^{\frac{2}{\alpha-1}}}$$

[0029] 式中,  $\bar{u}_{i'j'}$ 为第 $i'$ 数据降维后历史营销数据对于第 $j'$ 聚类中心的更新的隶属度; $d_{i'j'}$ 、 $d_{i'k'}$ 为第 $i'$ 个数据降维后历史营销数据到第 $j'$ 、 $k'$ 个更新的聚类中心的距离; $k'$ 为聚类中心指示量;

$$z_{j'} = \frac{\sum_{i'=1}^m \bar{u}_{i'j'}^{-\alpha} x'_{i'}}{\sum_{i'=1}^m \bar{u}_{i'j'}^{-\alpha}}$$

[0030] 式中,  $z_{j'}$  为第  $j'$  更新的聚类中心;  $x_{i'}$  为第  $i'$  数据降维后历史营销数据;  $\bar{u}_{i'j'}$  为第  $i'$  数据降维后历史营销数据对于第  $j'$  聚类中心的更新的隶属度;

S1-5-4: 根据每一数据降维后历史营销数据与若干最终的聚类中心的欧氏距离, 对数据降维后历史营销数据集进行划分, 得到每一最终的聚类中心和对应的聚类集群;

S1-6: 对每一聚类中心设置对应的历史营销效果等级标签, 并将历史营销效果等级标签扩散至聚类中心对应的聚类集群, 得到设置有历史营销效果等级标签的若干营销效果分类样本;

S1-7: 整合所有营销效果分类样本, 得到包含若干历史营销服务类型标签和若干历史营销效果等级标签的营销效果分类样本集;

S2: 根据若干历史营销服务类型标签和若干历史营销效果等级标签, 构建对应的若干营销服务数据库和若干营销效果子数据库, 将每一营销效果分类样本存储至对应的营销服务数据库的对应的营销效果子数据库中, 并根据营销效果分类样本集, 使用深度学习算法, 构建对应的营销效果分类模型, 包括如下步骤:

S2-1: 根据A类历史营销服务类型标签, 构建对应的A个营销服务数据库, 其中, A为历史营销服务类型标签种类数量;

S2-2: 根据B类历史营销效果等级标签, 在每一营销服务数据库中构建对应的B个营销效果子数据库, 其中, B为历史营销效果等级标签种类数量;

S2-3: 根据每一营销效果分类样本的历史营销服务类型标签和历史营销效果等级标签, 将营销效果分类样本存储至对应的营销服务数据库的对应的营销效果子数据库中;

S2-4: 根据营销效果分类样本集, 使用N-生成对抗网络 (Generative Adversarial Network, GAN) -Attention-多层感知机 (Multilayer Perceptron, MLP) 算法, 构建对应的营销效果分类模型, 包括如下步骤:

S2-4-1: 将营销效果分类样本集中的若干营销效果分类样本, 按照7:3的比例划分为模型训练样本集和模型测试样本集;

S2-4-2: 使用N-GAN-Attention-MLP算法, 构建初始的营销效果分类模型, 并将模型训练样本集输入初始的营销效果分类模型, 进行优化训练, 得到优化的营销效果分类模型, 其中, N为关键营销指标总数;

S2-4-3: 将模型测试样本集输入优化的营销效果分类模型, 进行模型测试, 得到对应的若干预测营销效果等级标签;

S2-4-4: 根据每一预测营销效果等级标签和对应的历史营销效果等级标签进行对比统计, 得到模型测试准确率;

S2-4-5: 若模型测试准确率大于预设模型测试准确率阈值, 则输出最优的营销效果分类模型, 否则, 继续进行优化训练;

S3: 采集实时营销数据, 根据关键营销指标集合, 对实时营销数据进行数据降维, 得到对应的数据降维后实时营销数据, 获取数据降维后实时营销数据的实时营销服务类型标签, 使用营销效果分类模型, 获取数据降维后实时营销数据的实时营销效果等级标签, 并根据实时营销服务类型标签和实时营销效果等级标签, 将数据降维后实时营销数据存储至对应的营销服务数据库的对应的营销效果子数据库中, 包括如下步骤:

S3-1:采集实时营销数据,根据关键营销指标集合,对实时营销数据进行数据降维,得到对应的数据降维后实时营销数据;

实时营销数据包括实时营销策略文本数据和实时营销指标数据;

实时营销指标数据包括实时营销成本指标、实时营销利润指标、实时营销网站流量指标、实时营销社交媒体指标、实时营销转化率指标、实时营销客户获取成本指标、实时销售业绩指标、实时市场竞争分析指标以及实时客户满意度指标;

S3-2:使用TF-IDF-CI算法,提取数据降维后实时营销数据的实时关键词集合,并将实时关键词集合作为对应的实时检索标签;

S3-3:根据实时关键词集合,使用营销知识图谱,匹配对应的实时营销服务类型标签;

S3-4:根据实时营销服务类型标签,将数据降维后实时营销数据发送至对应的营销服务数据库,包括如下步骤:

S3-4-1:若存在与实时营销服务类型标签相匹配的历史营销服务类型标签,则将数据降维后实时营销数据发送至匹配的历史营销服务类型标签对应的营销服务数据库,否则,进入下一步骤;

S3-4-2:根据实时营销服务类型标签,构建新增的营销服务数据库,并将数据降维后实时营销数据发送至新增的营销服务数据库;

S3-4-3:根据B类历史营销效果等级标签,在新增的营销服务数据库中构建对应的B个营销效果子数据库;

S3-5:使用营销效果分类模型,对数据降维后实时营销数据进行营销效果分类,得到对应的实时营销效果等级标签,包括如下步骤:

S3-5-1:将数据降维后实时营销数据输入营销效果分类模型;

S3-5-2:使用营销效果分类模型的N个基于GAN网络的生成器,提取数据降维后实时营销数据在N个角度的数据特征;

基于Transformer结构作为GAN网络的生成器,用于在多个角度提取数据降维后实时营销数据的数据特征,具备对不同尺度信息进行特征提取的泛化能力;Attention机制对GAN网络的若干个多尺度生成器的权重进行动态调节,使得这些多尺度生成器具备不同的权重;

S3-5-3:对N个角度的数据特征进行特征融合,得到对应的融合数据特征;

MLP网络作为全连接层,特点是层与层之间的神经元全部相互连接,用于对N个角度的降维后实时营销数据的数据特征进行特征融合,得到多角度融合特征,提高数据表征能力;

S3-5-4:根据融合数据特征,进行营销效果分类,得到对应的实时营销效果等级标签;

基于softmax函数构建的分类器,用于根据多角度融合特征,进行实时营销效果等级标签预测,得到实时营销效果等级标签;

S3-6:根据实时营销效果等级标签,将数据降维后实时营销数据存储至营销服务数据库的对应的营销效果子数据库中。

[0031] 实施例2:

如图2所示,本实施例提供一种管理营销数据的数据处理系统,用于实现数据处理方法,系统包括依次连接的数据初始化处理单元、数据库构建单元以及实时营销数据处理单元;

数据初始化处理单元,用于采集历史营销大数据,对历史营销大数据进行数据降维,得到对应的关键营销指标集合和数据降维后历史营销数据集,并为每一数据降维后历史营销数据设置对应的历史营销服务类型标签和历史营销效果等级标签,得到包含若干历史营销服务类型标签和若干历史营销效果等级标签的营销效果分类样本集;

数据库构建单元,用于根据若干历史营销服务类型标签和若干历史营销效果等级标签,构建对应的若干营销服务数据库和若干营销效果子数据库,将每一营销效果分类样本存储至对应的营销服务数据库的对应的营销效果子数据库中,并根据营销效果分类样本集,使用深度学习算法,构建对应的营销效果分类模型;

实时营销数据处理单元,用于采集实时营销数据,根据关键营销指标集合,对实时营销数据进行数据降维,得到对应的数据降维后实时营销数据,获取数据降维后实时营销数据的实时营销服务类型标签,使用营销效果分类模型,获取数据降维后实时营销数据的实时营销效果等级标签,并根据实时营销服务类型标签和实时营销效果等级标签,将数据降维后实时营销数据存储至对应的营销服务数据库的对应的营销效果子数据库中。

[0032] 本发明公开了一种管理营销数据的数据处理方法及系统,对营销数据进行数据降维,解决了营销数据存在的高维、复杂以及噪声的问题,并减小数据体量和内存占用空间;构建营销服务数据库和营销效果子数据库将营销数据按照营销服务类型和营销效果等级进行分类存储,加强营销数据之间的数据交流的同时,便于对营销数据进行管理和检索,提高了营销数据的数据价值;营销效果分类模型对营销数据进行自动、高效的营销效果分类,提高了营销数据分类和存储的实用性和准确性。

[0033] 本发明不局限于上述可选的实施方式,任何人在本发明的启示下都可得出其他各种形式的产品。上述具体实施方式不应理解成对本发明的保护范围的限制,本发明的保护范围应当以权利要求书中界定的为准,并且说明书可以用于解释权利要求书。

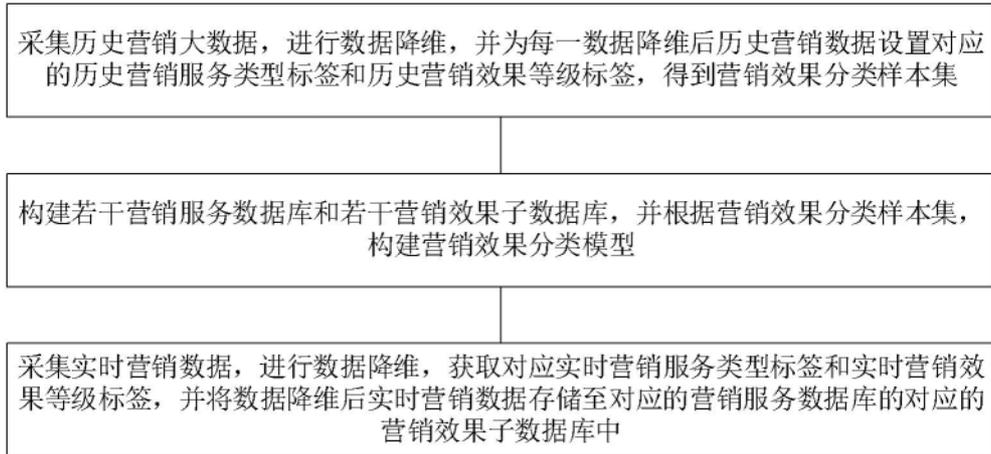


图1



图2