



(12) 发明专利申请

(10) 申请公布号 CN 116720882 A

(43) 申请公布日 2023. 09. 08

(21) 申请号 202310620786.6

G06N 20/00 (2019.01)

(22) 申请日 2023.05.30

(71) 申请人 精英数智科技股份有限公司

地址 030000 山西省太原市小店区南中环
街529号清控创新基地D座10层

(72) 发明人 李彦 白博 高敏敏 潘涛
王世伟

(74) 专利代理机构 北京超凡宏宇专利代理事务
所(特殊普通合伙) 11463

专利代理师 安卫静

(51) Int. Cl.

G06Q 30/0202 (2023.01)

G06Q 50/02 (2012.01)

G06V 20/62 (2022.01)

H04W 4/021 (2018.01)

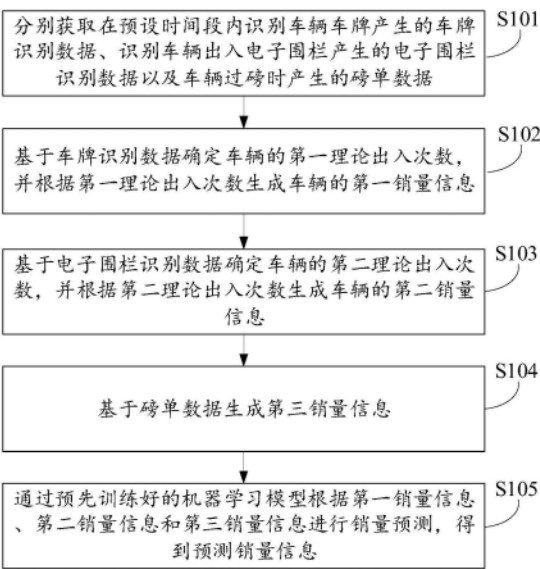
权利要求书2页 说明书13页 附图4页

(54) 发明名称

一种煤矿销量数据的预测方法、系统及电子设备

(57) 摘要

本发明提供了一种煤矿销量数据的预测方法、系统及电子设备,涉及数据管理技术领域,该方法首先获取在预设时间段内识别车辆车牌产生的车牌识别数据、识别车辆出入电子围栏产生的电子围栏识别数据以及车辆过磅时产生的磅单数据;而后基于车牌识别数据确定车辆的第一理论出入次数,并根据第一理论出入次数生成车辆的第一销量信息;再基于电子围栏识别数据确定车辆的第二理论出入次数,并根据第二理论出入次数生成车辆的第二销量信息;在基于磅单数据生成第三销量信息之后通过预先训练好的机器学习模型根据上述销量信息进行销量预测,得到预测销量信息。本发明能够提高运销流程中煤炭数据的准确性,进而提升煤炭监管效率。



1. 一种煤矿销量数据的预测方法,其特征在于,所述方法包括以下步骤:

分别获取在预设时间段内识别车辆车牌产生的车牌识别数据、识别车辆出入电子围栏产生的电子围栏识别数据以及车辆过磅时产生的磅单数据;

基于所述车牌识别数据确定所述车辆的第一理论出入次数,并根据所述第一理论出入次数生成所述车辆的第一销量信息;

基于所述电子围栏识别数据确定所述车辆的第二理论出入次数,并根据所述第二理论出入次数生成所述车辆的第二销量信息;

基于所述磅单数据生成第三销量信息;

通过预先训练好的机器学习模型根据所述第一销量信息、所述第二销量信息和所述第三销量信息进行销量预测,得到预测销量信息。

2. 根据权利要求1所述的煤矿销量数据的预测方法,其特征在于,所述车牌识别数据包括:车牌信息、出矿时记录的出矿时间、空车过磅时记录的第一空车时间和重车过磅时记录的第一重车时间;所述基于所述车牌识别数据确定所述车辆的第一理论出入次数的步骤,包括:

将所述第一重车时间和对应的所述第一空车时间之间的时长确定为第一消耗时长;

将所述第一消耗时长与预设的第一正常耗时范围进行比较,并根据比较结果确定所述第一理论出入次数。

3. 根据权利要求2所述的煤矿销量数据的预测方法,其特征在于,所述根据比较结果确定所述第一理论出入次数的步骤,包括:

当所述第一消耗时长小于或等于所述第一正常耗时范围的上限值时,确定所述第一理论出入次数为1;

当所述第一消耗时长大于所述第一正常耗时范围的上限值时,确定所述第一理论出入次数为2。

4. 根据权利要求2所述的煤矿销量数据的预测方法,其特征在于,所述电子围栏识别数据至少包括:进入区域时间和离开区域时间;所述基于所述电子围栏识别数据确定所述车辆的第二理论出入次数的步骤,包括:

将所述离开区域时间和所述进入区域时间之间的时长确定为所述车辆在电子围栏识别下的第二消耗时长;

将所述第二消耗时长与预设的第二正常耗时范围进行比较;

当所述第二消耗时长在所述第二正常耗时范围内时,确定所述车辆的第二理论出入次数为1;

当所述第二消耗时长大于或等于所述第二正常耗时范围的上限值时,确定所述车辆的第二理论出入次数为2;

当所述第二消耗时长小于或等于所述第二正常耗时范围的下限值时,确定所述车辆的第二理论出入次数为0。

5. 根据权利要求1所述的煤矿销量数据的预测方法,其特征在于,所述通过预先训练好的机器学习模型根据所述第一销量信息、所述第二销量信息和所述第三销量信息进行销量预测,得到预测销量信息的步骤,包括:

将所述第一销量信息、所述第二销量信息和所述第三销量信息代入预测公式,得到预

测销量信息;所述预测公式包括:

$$T=\alpha S1+\beta S2+\gamma S3+d$$

其中,T表示所述预测销量信息,S1表示所述第一销量信息,S2表示所述第二销量信息,S3表示所述第三销量信息, α 、 β 和 γ 分别为预设的权重,d表示预设的误差值。

6.根据权利要求4所述的煤矿销量数据的预测方法,其特征在于,所述方法还包括:

以所述第一销量信息、所述第二销量信息、所述第三销量信息作为特征数据,以理论销量信息作为训练目标,对所述机器学习模型进行训练。

7.根据权利要求6所述的煤矿销量数据的预测方法,其特征在于,所述车牌识别数据还包括:入矿时间,所述磅单数据还包括:空车过磅的第二空车时间和重车过磅的第二重车时间,所述磅单数据包括:车辆的净重和核定载重;所述理论销量信息的获取方法包括:

将所述进入区域时间、所述入矿时间、所述第一空车时间和所述第二空车时间中的最小时间确定为开始时间;

将所述离开区间时间、所述出矿时间、所述第一重车时间和所述第二重车时间中的最大时间确定为离开时间;

将所述离开时间和所述开始时间之间的时长确定所述车辆在拉煤销量流程中的第三消耗时长;

将所述第三消耗时长与预设的第三正常耗时范围进行比较,并根据比较结果确定所述车辆的第三理论出入次数;其中,所述第三理论出入次数还用于所述机器学习模型的训练过程;

根据所述第三理论出入次数和所述车辆的核定载重,生成所述车辆在所述预设时间段内的理论销量信息。

8.根据权利要求7所述的煤矿销量数据的预测方法,其特征在于,所述方法还包括:

根据所述第三消耗时长与所述第三正常耗时范围之间的比较结果,标记车辆出入次数的类型为正常、异常或超时。

9.一种煤矿销量数据的预测系统,其特征在于,所述系统包括以下模块:

数据获取模块,用于分别获取在预设时间段内识别车辆车牌产生的车牌识别数据、识别车辆出入电子围栏产生的电子围栏识别数据以及车辆过磅时产生的磅单数据;

第一数据处理模块,用于基于所述车牌识别数据确定所述车辆的第一理论出入次数,并根据所述第一理论出入次数生成所述车辆的第一销量信息;

第二数据处理模块,用于基于所述电子围栏识别数据确定所述车辆的第二理论出入次数,并根据所述第二理论出入次数生成所述车辆的第二销量信息;

第三数据处理模块,用于基于所述磅单数据生成第三销量信息;

销量预测模块,用于通过预先训练好的机器学习模型根据所述第一销量信息、所述第二销量信息和所述第三销量信息进行销量预测,得到预测销量信息。

10.一种电子设备,其特征在于,包括:处理器和存储装置;所述存储装置上存储有计算机程序,所述计算机程序在被所述处理器运行时实现上述权利要求1至8任一项所述的煤矿销量数据的预测方法的步骤。

一种煤矿销量数据的预测方法、系统及电子设备

技术领域

[0001] 本发明涉及数据管理技术领域,尤其是涉及一种煤矿销量数据的预测方法、系统及电子设备。

背景技术

[0002] 目前,煤矿企业存在瞒报或漏报原煤产量、商品煤销量等问题,为此,亟需一种数智化监管煤炭产运销的方案。有效的煤炭产运销储管理需要依赖可靠的数据支持和科学依据,然而在实际操作中,数据的采集和填报存在人为篡改数据的可能,从而影响数据真实性,导致煤炭监管的准确性和效率都比较差。

发明内容

[0003] 有鉴于此,本发明的目的在于提供一种煤矿销量数据的预测方法、系统及电子设备,该方法能够提高煤炭运输销售流程中数据的真实性、准确性,进而提升煤炭监管效率。

[0004] 第一方面,本发明实施例提供了一种煤矿销量数据的预测方法,所述方法包括以下步骤:

[0005] 分别获取在预设时间段内识别车辆车牌产生的车牌识别数据、识别车辆出入电子围栏产生的电子围栏识别数据以及车辆过磅时产生的磅单数据;

[0006] 基于车牌识别数据确定车辆的第一理论出入次数,并根据第一理论出入次数生成车辆的第一销量信息;

[0007] 基于电子围栏识别数据确定车辆的第二理论出入次数,并根据第二理论出入次数生成车辆的第二销量信息;

[0008] 基于磅单数据生成第三销量信息;

[0009] 通过预先训练好的机器学习模型根据第一销量信息、第二销量信息和第三销量信息进行销量预测,得到预测销量信息。

[0010] 在一些实施方式中,车牌识别数据包括:车牌信息、出矿时记录的出矿时间、空车过磅时记录的第一空车时间和重车过磅时记录的第一重车时间;基于车牌识别数据确定车辆的第一理论出入次数的步骤,包括:

[0011] 将第一重车时间和对应的第一空车时间之间的时长确定为第一消耗时长;

[0012] 将第一消耗时长与预设的第一正常耗时范围进行比较,并根据比较结果确定第一理论出入次数。

[0013] 在一些实施方式中,根据比较结果确定第一理论出入次数的步骤,包括:

[0014] 当第一消耗时长小于或等于第一正常耗时范围的上限值时,确定第一理论出入次数为1;

[0015] 当第一消耗时长大于第一正常耗时范围的上限值时,确定第一理论出入次数为2。

[0016] 在一些实施方式中,电子围栏识别数据至少包括:进入区域时间和离开区域时间;基于电子围栏识别数据确定车辆的第二理论出入次数的步骤,包括:

[0017] 将离开区域时间和进入区域时间之间的时长确定为车辆在电子围栏识别下的第二消耗时长；

[0018] 将第二消耗时长与预设的第二正常耗时范围进行比较；

[0019] 当第二消耗时长在第二正常耗时范围内时，确定车辆的第二理论出入次数为1；

[0020] 当第二消耗时长大于或等于第二正常耗时范围的上限值时，确定车辆的第二理论出入次数为2；

[0021] 当第二消耗时长小于或等于第二正常耗时范围的下限值时，确定车辆的第二理论出入次数为0。

[0022] 在一些实施方式中，通过预先训练好的机器学习模型根据第一销量信息、第二销量信息和第三销量信息进行销量预测，得到预测销量信息的步骤，包括：

[0023] 将第一销量信息、第二销量信息和第三销量信息代入预测公式，得到预测销量信息；预测公式包括：

[0024] $T = \alpha S1 + \beta S2 + \gamma S3 + d$

[0025] 其中，T表示预测销量信息，S1表示第一销量信息，S2表示第二销量信息，S3表示第三销量信息， α 、 β 和 γ 分别为预设的权重，d表示预设的误差值。

[0026] 在一些实施方式中，该方法还包括：以第一销量信息、第二销量信息、第三销量信息作为特征数据，以理论销量信息作为训练目标，对机器学习模型进行训练。

[0027] 在一些实施方式中，车牌识别数据还包括：入矿时间，磅单数据还包括：空车过磅的第二空车时间和重车过磅的第二重车时间，磅单数据包括：车辆的净重和核定载重；理论销量信息的获取方法包括：

[0028] 将进入区域时间、入矿时间、第一空车时间和第二空车时间中的最小时间确定为开始时间；

[0029] 将离开区域时间、出矿时间、第一重车时间和第二重车时间中的最大时间确定为离开时间；

[0030] 将离开时间和开始时间之间的时长确定车辆在拉煤销量流程中的第三消耗时长；

[0031] 将第三消耗时长与预设的第三正常耗时范围进行比较，并根据比较结果确定车辆的第三理论出入次数；其中，第三理论出入次数还用于机器学习模型的训练过程；

[0032] 根据第三理论出入次数和车辆的核定载重，生成车辆在预设时间段内的理论销量信息。

[0033] 在一些实施方式中，该方法还包括：根据第三消耗时长与第三正常耗时范围之间的比较结果，标记车辆出入次数的类型为正常、异常或超时。

[0034] 第二方面，本发明实施例提供了一种煤矿销量数据的预测系统，系统包括以下模块：

[0035] 数据获取模块，用于分别获取在预设时间段内识别车辆车牌产生的车牌识别数据、识别车辆出入电子围栏产生的电子围栏识别数据以及车辆过磅时产生的磅单数据；

[0036] 第一数据处理模块，用于基于车牌识别数据确定车辆的第一理论出入次数，并根据第一理论出入次数生成车辆的第一销量信息；

[0037] 第二数据处理模块，用于基于电子围栏识别数据确定车辆的第二理论出入次数，并根据第二理论出入次数生成车辆的第二销量信息；

[0038] 第三数据处理模块,用于基于磅单数据生成第三销量信息;

[0039] 销量预测模块,用于通过预先训练好的机器学习模型根据第一销量信息、第二销量信息和第三销量信息进行销量预测,得到预测销量信息。

[0040] 第三方面,发明实施例还提供一种电子设备,包括存储器、处理器,存储器中存储有可在处理器上运行的计算机程序,其中,处理器执行计算机程序时实现上述第一方面提到的煤矿销量数据的预测方法的步骤。

[0041] 本发明实施例带来了至少以下有益效果:

[0042] 本发明提供了一种煤矿销量数据的预测方法、系统及电子设备,该方法在进行煤矿数据处理时需要分别获取在预设时间段内识别车辆车牌产生的车牌识别数据、识别车辆出入电子围栏产生的电子围栏识别数据以及车辆过磅时产生的磅单数据,以上多维度的煤炭数据提供了可靠的数据基础,保证数据的真实性和全面性。其中,基于物联网识别数据能够较好地实现对煤炭的重量等数据的实时监测和录入,确保煤炭销售数据的真实性。基于车牌识别数据确定车辆的第一理论出入次数,并根据第一理论出入次数生成车辆的第一销量信息;该过程对车辆入出矿和过磅时的车牌进行识别,获取车辆信息,对每辆车进行跟踪,从而实现对车辆的全程跟踪和监管,有效防止企业瞒报或漏报车次信息,保障数据真实性。基于电子围栏识别数据确定车辆的第二理论出入次数,并根据第二理论出入次数生成车辆的第二销量信息;以上内容利用电子围栏技术实现电子煤矿管控:在各个关键节点设置电子围栏,加强对进出矿山的车辆流量和货物流向的监管和管控,避免非法运输及数据造假等问题的产生,提高运销数据的真实性和准确性。在基于磅单数据生成第三销量信息之后,通过预先训练好的机器学习模型根据第一销量信息、第二销量信息和第三销量信息进行销量预测,得到预测销量信息。本方案基于以上多维度数据推算企业真实销量,各维度之间相互比对,检验数据的准确性和有效性,并通过对多个关键指标进行数据挖掘和分析,能够有效地监管煤炭产运销的全流程,提高监管的效率和精度,避免管中存在的真实性问题和企业瞒报漏报问题发生。因此,本发明能够提高运销流程中煤炭数据的真实性、准确性,进而提升煤炭监管效率。

[0043] 本发明的其他特征和优点将在随后的说明书中阐述,或者,部分特征和优点可以从说明书推知或毫无疑义的确定,或者通过实施本发明的上述技术即可得知。

[0044] 为使本发明的上述目的、特征和优点能更明显易懂,下文特举较佳实施方式,并配合所附附图,作详细说明如下。

附图说明

[0045] 为了更清楚地说明本发明具体实施方式或现有技术中的技术方案,下面将对具体实施方式或现有技术描述中所需要使用的附图作简单地介绍,显而易见地,下面描述中的附图是本发明的一些实施方式,对于本领域普通技术人员来讲,在不付出创造性劳动的前提下,还可以根据这些附图获得其他的附图。

[0046] 图1为本发明实施例提供的一种煤矿销量数据的预测方法的流程图;

[0047] 图2为本发明实施例提供的一种煤矿销量数据的预测方法中AI识别数据处理的流程图;

[0048] 图3为本发明实施例提供的一种煤矿销量数据的预测方法中电子围栏识别数据处

理的流程圖；

[0049] 圖4為本發明實施例提供的一種煤礦銷量數據的預測方法中理論銷量信息获取的流程圖；

[0050] 圖5為本發明實施例提供的一種煤礦銷量數據的預測系統的结构示意圖；

[0051] 圖6為本發明實施例提供的一種電子設備的结构示意圖。

[0052] 圖標：

[0053] 501-數據获取模块；502-第一數據處理模块；503-第二數據處理模块；504-第三數據處理模块；505-銷量預測模块；

[0054] 101-处理器；102-存储器；103-总线；104-通信接口。

具体实施方式

[0055] 为使本發明實施例的目的、技術方案和優點更加清楚，下面將結合附图对本发明的技术方案进行清楚、完整地描述，显然，所描述的实施例是本发明一部分实施例，而不是全部的实施例。基于本发明中的实施例，本领域普通技术人员在没有做出创造性劳动前提下所获得的所有其他实施例，都属于本发明保护的范围。

[0056] 目前在煤炭产运销储管理中，数据的采集和填报存在人为篡改数据的可能，从而影响数据真实性，导致煤炭监管的准确性和效率都比较差。

[0057] 基于此，本發明實施例提供的一種煤礦銷量數據的預測方法、系統及電子設備，該方法從AI識別、電子圍欄和物聯網三個維度對煤炭運輸銷售流程中的數據進行全程跟蹤采集，能夠提高運銷數據的真實性和準確性，並通過數據挖掘和分析，能夠有效地監管煤炭產運銷的全流程，有效防止企業瞞報或漏報車次信息，提高管的效率。

[0058] 為便于對本實施例進行理解，首先對本發明實施例所公開的一種煤礦銷量數據的預測方法進行詳細介紹，該方法應用於煤礦運輸銷售的場景中。如圖1所示，煤礦銷量數據的預測方法包括以下步驟：

[0059] 步驟S101，分別获取在預設時間段內識別車輛車牌產生的車牌識別數據、識別車輛出入電子圍欄產生的電子圍欄識別數據以及車輛過磅時產生的磅單數據。

[0060] 具體的說，可在預設時間段通過相關車牌識別裝置來获取車牌識別數據；電子圍欄識別數據可包括：進入區域時間和離開區域時間；磅單數據可包括：車輛的淨重和核定載重。第一時間周期諸如為12小時、一天、一周等。

[0061] 針對單一煤礦，一般設置有電子圍欄覆蓋企業礦場，在進出礦場的位置處設置有車牌識別設備，且基於物聯網(Internet of Things, 簡稱IoT)對礦場進行管理。一輛符合運輸銷售規範的車輛進出礦場運輸煤礦的流程一般是：進入電子圍欄合法區域、車輛入礦、空車過磅、車輛拉煤、重車過磅、車輛出礦、離開電子圍欄合法區域。

[0062] 根據以上情況，本實施例可以從電子圍欄、車牌識別和物聯網三個維度获取煤礦數據。針對電子圍欄維度，可以获取如下電子圍欄識別數據：車輛的車牌信息、進入電子圍欄合法區域時的進入區域時間a1、離開電子圍欄合法區域時的離開區域時間a2。針對AI識別維度，可以获取如下AI識別數據：車牌信息、車輛入礦和出礦時的入礦時間b1、出礦時間b2、AI識別下空車過磅的第一空車時間c1和重車過磅的第一重車時間c2。針對物聯網維度，可以获取如下物聯網識別數據：車牌信息、車輛的淨重和核定載重、物聯網識別下空車過磅

的第二空车时间d1和重车过磅的第二重车时间d2。

[0063] 为便于数据记录,可以采用数据表记录上述各维度的煤矿数据,数据表分别包括:车牌识别数据表、电子围栏识别数据表和磅单数据表。

[0064] 步骤S102,基于车牌识别数据确定车辆的第一理论出入次数,并根据第一理论出入次数生成车辆的第一销量信息。

[0065] 基于车牌识别数据确定车辆的第一理论出入次数,然后根据第一理论出入次数和车辆的核定载重确定车辆的第一销量信息。在本实施例中,可以确定车牌识别数据中的第一空车时间和第一重车时间之间的第一消耗时长,将第一消耗时长与预设的第一正常耗时范围进行比较,根据比较结果确定车辆的第一理论出入次数。示例性的:若第一消耗时长大于第一正常耗时范围的上限值,则记录第一理论出入次数为2;若第一消耗时长小于或等于上述上限值,则记录第一理论出入次数为1。根据以上方式记录第一时间周期内第一理论出入次数的总次数。

[0066] 根据车牌识别数据中的车牌信息,从物联网识别数据中查询是否存在与该车牌信息匹配的核定载重;若存在则使用该核定载重,若不存在则使用默认载重。而后,根据第一理论出入次数的总次数和车辆的核定载重之间的乘积,计算车辆在AI识别下的第一销量信息。

[0067] 步骤S103,基于电子围栏识别数据确定车辆的第二理论出入次数,并根据第二理论出入次数生成车辆的第二销量信息。

[0068] 基于电子围栏识别数据确定车辆的第二理论出入次数,并根据第二理论出入次数和车辆的核定载重生车辆在电子围栏识别下的第二销量信息。在本实施例中,可以确定电子围栏识别数据中离开区域时间和进入区域时间之间的第二消耗时长,将第二消耗时长与预设的第二正常耗时范围进行比较,根据比较结果确定第二理论出入次数。记录第一时间周期内第二理论出入次数的总次数。根据第二理论出入次数的总次数和车辆的核定载重之间的乘积,生成车辆在电子围栏识别下的第二销量信息。

[0069] 步骤S104,基于磅单数据生成第三销量信息。

[0070] 基于各车辆的净重确定第一时间周期内的车辆净重总和,在本实施例中,可以基于物联网识别车辆,每识别到一次车辆时均可以在物联网磅单信息表中进行一次车辆记录,由此可以得到第一时间周期内的车辆出入次数的总和,根据出入次数的总和与车辆的净重之间的乘积,计算车辆净重总和,最终生成第三销量信息。

[0071] 步骤S105,通过预先训练好的机器学习模型根据第一销量信息、第二销量信息和第三销量信息进行销量预测,得到预测销量信息。

[0072] 通过预先训练好的机器学习模型根据第一销量信息、第二销量信息和第三销量信息进行销量预测,得到目标销量信息。在实现过程中,通过预先训练好的机器学习模型根据预测公式,对第一销量信息、第二销量信息和第三销量信息进行销量预测,得到目标销量信息;其中,预测公式包括:

[0073] $T = \alpha S1 + \beta S2 + \gamma S3 + d$

[0074] 其中,T表示预测销量信息,S1表示第一销量信息,S2表示第二销量信息,S3表示第三销量信息, α 、 β 和 γ 分别为预设的权重,d表示预设的误差值。

[0075] 为便于理解,以下分别从车牌识别、电子围栏和物联网三个维度,对上述实施例提

供的煤炭数据处理方法展开详细描述。

[0076] 在一些实施方式中,针对车牌识别维度下的步骤S102,首先将第一重车时间和对应的第一空车时间之间的时长确定为第一消耗时长;然后将第一消耗时长与预设的第一正常耗时范围进行比较,并根据比较结果确定第一理论出入次数。具体的如图2所示,包括:

[0077] (1)从多组AI识别数据中的车牌信息中,选择包含有效车牌的目标车牌信息,并将目标车牌信息所属的AI识别数据确定为目标AI识别数据;

[0078] 对于每辆车,在完成一次拉煤流程中,会出现在某个卡口未成功识别车牌或未记录时间的现象。基于此,本实施例针对采集到的K1组AI识别数据,可以选择第一空车时间c1和第一重车时间c2中有至少一项为非空值的AI识别数据,换言之:将仅有车牌信息而第一空车时间c1和第一重车时间c2均为空值的一组AI识别数据进行滤除。通过上述基于时间的筛选可以得到K2($K2 \leq K1$)组AI识别数据。

[0079] K2组AI识别数据可能对应于不同的车牌信息,那么,对同一车牌信息对应的多条AI识别数据可以进行聚合。具体的,对第一时间周期(如单日)内某个车牌信息的多条AI识别数据,如果当前AI识别数据中空车过磅的第一空车时间为空值,重车过磅的第一重车时间为非空值,则根据AI识别车辆信息表从第一空车时间为空值的行,向上遍历,直至遍历到首个第一空车时间为非空值的行数据中的第一重车时间,并将该遍历到的第一重车时间与当前的第一重车时间按最大时间聚合成一条。根据以上聚合方式,能够缩减同一车牌信息对应的AI识别数据的组数。相应地,将K2组AI识别数据聚合为K3($K3 \leq K2$)组AI识别数据。

[0080] 针对K3组AI识别数据,可以根据车牌信息作进一步筛选。筛选方式为,判断AI识别数据的车牌信息是否为有效车牌;其中,有效车牌是指车牌内容清晰完整。若是有效车牌,则将该车牌信息确定为有效的目标车牌信息,并将目标车牌信息所属的AI识别数据确定为目标AI识别数据。若不是有效车牌,比如车牌空为空、未挂载车牌、车牌号遮挡等,则将该车牌信息确定为无效车牌信息,并删除该无效车牌信息所属的AI识别数据。通过上述基于车牌信息的筛选方式可以得到K4($K4 \leq K3$)组AI识别数据。

[0081] 将以上经过多次筛选后最终的K4组AI识别数据确定为目标AI识别数据。

[0082] (2)判断同一组目标AI识别数据中的第一空车时间和第一重车时间是否为非空值。

[0083] 本实施例可以基于判断结果确定车辆的第一理论出入次数,具体可参照如下表1以及如下步骤(3)-(6)。

[0084] 表1

[0085]	第一空车时间 C1	第一重车时间 C2	比较结果	第一理论 出入次数	类型
	空值	非空值		1	异常
	非空值	非空值	$c2-c1 \leq U_{\pm 1}$	1	正常
	非空值	非空值	$c2-c1 > U_{\pm 1}$	2	异常
	非空值	空值	$c2-c1 \leq U_{\pm 1}$	1	正常
	非空值	空值	$b2-c1 > U_{\pm 1}$	2	异常
	非空值	空值, 且出矿时间 b2 为空值		0	异常

[0086] (3) 如果第一空车时间为空值, 第一重车时间为非空值, 则确定车辆的第一理论出入次数为1。以及, 将第一理论出入次数和与其匹配的车牌信息进行关联。本实施例还可以标记车辆出入次数的类型为异常, 在一些描述中, 可以将车辆出入次数简称为车次。

[0087] (4) 如果第一空车时间和第一重车时间均为非空值, 则将第一重车时间和第一空车时间之间的时长确定为车辆在AI识别下的第一消耗时长。

[0088] 在目标AI识别数据中的第一空车时间和第一重车时间记录完整, 即都为非空值的情况下, 将第一重车时间c2和第一空车时间c1之间的时长确定为车辆在AI识别下的第一消耗时长t1, $t1 = c2 - c1$ 。

[0089] 而后, 将第一消耗时长与预设的第一正常耗时范围进行比较, 并根据比较结果确定第一理论出入次数。其中, 考虑真实情况会存在很多异常数据, 于是, 可以预先获取指定时间段(如一个月)内大量的历史第一消耗时长, 依据正态分布 3σ 原则, 计算正常耗时的上下限值 $Lower = \mu - 3\sigma$, $Upper = \mu + 3\sigma$; 根据以上历史第一消耗时长和正态分布 3σ 原则确定第一正常耗时范围。或者, 在历史第一消耗时长的数据量不足的情况下, 可以采用经验值确定第一正常耗时范围。

[0090] 将第一消耗时长与预设的第一正常耗时范围进行比较, 当第一消耗时长小于或等于第一正常耗时范围的上限值 $U_{\pm 1}$ 时, 确定第一理论出入次数为1; 以及, 标记车辆出入次数的类型为正常。

[0091] 当第一消耗时长大于第一正常耗时范围的上限值 $U_{\pm 1}$ 时, 确定第一理论出入次数为2, 以及标记车辆出入次数的类型为异常。

[0092] (5) 如果第一空车时间为非空值, 第一重车时间为空值, 则将非空值的出矿时间和第一空车时间之间的时长确定为第一消耗时长。与第一重车时间最为接近的时间为出矿时间b2, 从而, 在第一重车时间为空值的情况下, 可以依据出矿时间b2辅助判断。具体包括: 判断出矿时间b2是否为非空值, 在出矿时间b2确定为非空值后, 将出矿时间b2和第一空车时间c1之间的时长确定为第一消耗时长t1, $t1 = b2 - c1$ 。

[0093] 而后, 将第一消耗时长与预设的第一正常耗时范围进行比较, 并根据比较结果确定第一理论出入次数。该过程与前述实施例相同, 在此不再赘述, 相应的第一理论出入次数可参照表1的第五、六行。

[0094] (6) 如果第一空车时间为非空值,第一重车时间为空值,且出矿时间为空值,则确定车辆的第一理论出入次数为0。

[0095] 根据以上实施例,基于AI识别数据确定车辆在第一时间周期内的第一理论出入次数之和,而后便可以根据第一理论出入次数之和与车辆的核定载重,生成车辆在AI识别下的第一销量信息S1。

[0096] 本实施例还可以将第一时间周期内的第一销量信息S1,与匹配的企业和日期进行关联保存。

[0097] 基于电子围栏识别数据确定车辆的第二理论出入次数的过程,首先将离开区域时间和进入区域时间之间的时长确定为车辆在电子围栏识别下的第二消耗时长;再将第二消耗时长与预设的第二正常耗时范围进行比较;当第二消耗时长在第二正常耗时范围内时,确定车辆的第二理论出入次数为1;当第二消耗时长大于或等于第二正常耗时范围的上限值时,确定车辆的第二理论出入次数为2;当第二消耗时长小于或等于第二正常耗时范围的下限值时,确定车辆的第二理论出入次数为0。具体参照图3,步骤如下:

[0098] 将离开区域时间a2和进入区域时间a1之间的时长确定为车辆在电子围栏识别下的第二消耗时长t2;将第二消耗时长t2与预设的第二正常耗时范围进行比较。第二正常耗时范围的确定方式可参照前述实施例中第一正常耗时范围的确定方式,在此不再具体说明。

[0099] 当第二消耗时长在第二正常耗时范围内时,确定车辆的第二理论出入次数为1。具体的,第二正常耗时范围的下限值为 L_{F2} ,上限值为 U_{L2} 。在 $L_{F2} \leq a2 - a1 \leq U_{L2}$ 的情况下,记录电子围栏维度下车辆的第二理论出入次数为1,标记车次类型为正常。

[0100] 当第二消耗时长大于或等于第二正常耗时范围的上限值时,确定车辆的第二理论出入次数为2。在 $a2 - a1 \geq U_{L2}$ 的情况下,记录电子围栏维度下车辆的第二理论出入次数为2,标记车次类型为超时。

[0101] 当第二消耗时长小于或等于第二正常耗时范围的下限值时,确定车辆的第二理论出入次数为0。在 $a2 - a1 \leq L_{F2}$ 的情况下,记录电子围栏维度下车辆的第二理论出入次数为0,标记车次类型为异常。

[0102] 根据以上实施例,基于电子围栏识别数据确定车辆在第一时间周期内的第二理论出入次数之和,而后便可以根据第二理论出入次数之和与车辆的核定载重,生成车辆在电子围栏识别下的第二销量信息S2。

[0103] 本实施例还可以将第一时间周期内的第二销量信息S2,与匹配的企业和日期进行关联保存。

[0104] 针对电子围栏识别数据,获取物联网磅单信息表中第一时间周期内每辆车辆的净重并求和,得到车辆净重总和S3。以及,将车辆净重总和S3,与匹配的企业和日期进行关联保存。

[0105] 根据以上实施例,获取多个第一时间周期内的第一销量信息S1、第二销量信息S2和车辆净重总和S3。例如,以“日”为第一时间周期为例,可以获取一个月内真实的第一销量信息S1、第二销量信息S2和车辆净重总和S3,对上述一个月的煤炭数据进行数据分析。在一些方式中,针对任一单维度的煤炭数据,如第一销量信息S1,可以分析其在不同日期的差异趋势,计算一个月内的差异均值、标准差;分析过程中采用的置信区间为基于 3σ 原则确定

的,分析结果在上下限之外的数据,可以提示数据异常,输出包括企业、异常销量和日期在内的数据异常提示信息。

[0106] 关于置信区间,若历史数据不足,可以计算当日与前一日的变化率 $k = \text{当日销量} / \text{前日销量} * 100\%$,设置 k 在 $50\% \sim 200\%$ 为正常区间,超过该正常区间的提示数据异常。

[0107] 针对三个维度或任意两个维度的煤炭数据,如第一销量信息 $S1$ 、第二销量信息 $S2$ 和车辆净重总和 $S3$,可以分析三者之间的销量差值,计算差值的均值、标准差。

[0108] 在本实施例中,通过预先训练好的机器学习模型根据第一销量信息、第二销量信息和第三销量信息进行销量预测,得到目标销量信息。其中,以AI识别下的第一销量信息、电子围栏识别下的第二销量信息、物联网识别下的第三销量信息作为特征数据,以理论销量信息作为训练目标,对机器学习模型进行训练。

[0109] 上述第一销量信息、第二销量信息、第三销量信息的获取方式,可参照前述关于第一销量信息、第二销量信息和第三销量信息的实施例,这些特征数据为真实的历史数据。

[0110] 通过预先训练好的机器学习模型根据第一销量信息、第二销量信息和第三销量信息进行销量预测,得到预测销量信息的过程中,首先将第一销量信息、第二销量信息和第三销量信息代入预测公式,得到预测销量信息;预测公式包括:

$$[0111] \quad T = \alpha S1 + \beta S2 + \gamma S3 + d$$

[0112] 其中, T 表示预测销量信息, $S1$ 表示第一销量信息, $S2$ 表示第二销量信息, $S3$ 表示第三销量信息, α 、 β 和 γ 分别为预设的权重, d 表示预设的误差值。具体的如图4所示,对于作为训练目标的预测销量信息的获取方法,可以参照如下实施例。

[0113] 步骤1,将进入区域时间 $a1$ 、入矿时间 $b1$ 、第一空车时间 $c1$ 和第二空车时间 $d1$ 中的最小时间确定为开始时间;将离开区域时间 $a2$ 、出矿时间 $b2$ 、第一重车时间 $c2$ 和第二重车时间 $d2$ 中的最大时间确定为离开时间;将离开时间和开始时间之间的时长确定车辆在拉煤销量流程中的第三消耗时长;

[0114] 根据车辆进入电子围栏合法区域、车辆入矿、空车过磅、车辆拉煤、重车过磅、车辆出矿、离开电子围栏合法区域这一流程,获取车辆在各个维度下的时间;通常情况下,时间先后应满足: $a1 < b1 < c1 < d1 < c2 < d2 < b2 < a2$,若不满足该关系,可以直接记录车辆的第三理论出入次数为0,并标记车次类型为异常。在满足上述关系的情况下,由于实际场景设备问题可能导致某个流程缺少时间记录,从而,理论真实车次的具体计算过程中,以第二空车时间 $d1$ 和第一重车时间 $c2$ 为时间节点划分开始时间和离开时间。进入区域时间 $a1$ 、入矿时间 $b1$ 、第一空车时间 $c1$ 和第二空车时间 $d1$ 不都为空,车次的开始时间 $T_{\text{开始}} = \min(a1, b1, c1, d1)$ 。离开区域时间 $a2$ 、出矿时间 $b2$ 、第一重车时间 $c2$ 和第二重车时间 $d2$ 不都为空,车次的离开时间 $T_{\text{离开}} = \max(c2, d2, b2, a2)$ 。其中,第三理论出入次数还用于机器学习模型的训练过程,用于训练模型时的理论真实标签。

[0115] 将离开时间和开始时间之间的时长 $T_{\text{离开}} - T_{\text{开始}}$,确定为车辆在拉煤销量流程中的第三消耗时长。

[0116] 步骤2,将第三消耗时长与预设的第三正常耗时范围进行比较,并根据比较结果确定车辆的第三理论出入次数。以及,还可以根据第三消耗时长与第三正常耗时范围之间的比较结果,标记车辆出入次数的类型为正常、异常或超时。

[0117] 如果比较结果为 $L_{\text{下}3} \leq T_{\text{离开}} - T_{\text{开始}} \leq U_{\text{上}3}$,记录车辆的第三理论出入次数为1,标车次类

型为正常。如果 $T_{离开} - T_{开始} \geq U_{上3}$, 记录车辆的第三理论出入次数为2, 标记车次类型为超时。如果 $T_{离开} - T_{开始} \leq L_{下3}$, 记录车辆的第三理论出入次数为0, 标记车次类型为异常。

[0118] 步骤3, 根据第三理论出入次数和车辆的核定载重, 生成车辆在第一时间周期内的理论销量信息。具体的, 确定车辆在第一时间周期内的第三理论出入次数之和, 而后便可以根据第三理论出入次数之和与车辆的核定载重, 生成车辆的预测销量信息。

[0119] 综上, 本发明实施例提供的煤矿销量数据的预测方法, 首先在第一时间周期内, 获取车辆的电子围栏识别数据、车牌识别数据和磅单数据, 以上多维度的煤炭数据提供了可靠的数据基础, 保证数据的真实性和全面性。其中, 基于磅单数据能够较好地实现对煤炭的重量等数据的实时监测和录入, 确保煤炭销售数据的真实性。基于车牌识别数据确定车辆的第一理论出入次数, 并根据第一理论出入次数和车辆的核定载重, 生成车辆在车牌识别下的第一销量信息; 该过程利用AI识别对车辆出入矿和过磅时的车牌进行识别, 获取车辆信息, 对每辆车进行跟踪, 从而实现对车辆的全程跟踪和监管, 有效防止企业瞒报或漏报车次信息, 保障数据真实性。基于电子围栏识别数据确定车辆的第二理论出入次数, 并根据第二理论出入次数和车辆的核定载重, 生成车辆在电子围栏识别下的第二销量信息; 以上内容利用电子围栏技术实现电子煤矿管控: 在各个关键节点设置电子围栏, 加强对进出矿山的车辆流量和货物流向的监管和管控, 避免非法运输及数据造假等问题的产生, 提高运销数据的真实性和准确性。基于各车辆的净重确定第一时间周期内的车辆净重总和。

[0120] 最后, 利用以上煤炭数据局, 通过预先训练好的机器学习模型根据第一销量信息、第二销量信息和车辆净重总和进行销量预测, 得到目标销量信息; 以及, 对多个第一时间周期之间的第一销量信息、第二销量信息和车辆净重总和进行数据分析, 数据分析包括: 差异趋势、均值、标准差。本方案基于以上多维度数据推算企业真实销量, 各维度之间相互比对, 检验数据的准确性和有效性, 并通过对多个关键指标进行数据挖掘和分析, 能够有效地监管煤炭产运销的全流程, 提高管的效率和精度, 避免监管中存在的数据真实性问题和企业瞒报漏报问题发生。因此, 本发明能够提高运销流程中煤炭数据的真实性、准确性, 进而提升煤炭监管效率。

[0121] 对应于上述方法实施例, 本发明实施例提供了一种煤矿销量数据的预测系统, 如图5所示, 该系统包括以下模块:

[0122] 数据获取模块501, 用于分别获取在预设时间段内识别车辆车牌产生的车牌识别数据、识别车辆出入电子围栏产生的电子围栏识别数据以及车辆过磅时产生的磅单数据;

[0123] 第一数据处理模块502, 用于基于车牌识别数据确定车辆的第一理论出入次数, 并根据第一理论出入次数生成车辆的第一销量信息;

[0124] 第二数据处理模块503, 用于基于电子围栏识别数据确定车辆的第二理论出入次数, 并根据第二理论出入次数生成车辆的第二销量信息;

[0125] 第三数据处理模块504, 用于基于磅单数据生成第三销量信息;

[0126] 销量预测模块505, 用于通过预先训练好的机器学习模型根据第一销量信息、第二销量信息和第三销量信息进行销量预测, 得到预测销量信息。

[0127] 在一些实施方式中, 车牌识别数据包括: 车牌信息、出矿时记录的出矿时间、空车过磅时记录的第一空车时间和重车过磅时记录的第一重车时间; 第一数据处理模块502具体用于:

- [0128] 将第一重车时间和对应的第一空车时间之间的时长确定为第一消耗时长；
- [0129] 将第一消耗时长与预设的第一正常耗时范围进行比较,并根据比较结果确定第一理论出入次数。
- [0130] 在一些实施方式中,第一数据处理模块502具体用于:
- [0131] 当第一消耗时长小于或等于第一正常耗时范围的上限值时,确定第一理论出入次数为1;
- [0132] 当第一消耗时长大于第一正常耗时范围的上限值时,确定第一理论出入次数为2。
- [0133] 在一些实施方式中,电子围栏识别数据至少包括:进入区域时间和离开区域时间;第二数据处理模块503具体用于:
- [0134] 将离开区域时间和进入区域时间之间的时长确定为车辆在电子围栏识别下的第二消耗时长;
- [0135] 将第二消耗时长与预设的第二正常耗时范围进行比较;
- [0136] 当第二消耗时长在第二正常耗时范围内时,确定车辆的第二理论出入次数为1;
- [0137] 当第二消耗时长大于或等于第二正常耗时范围的上限值时,确定车辆的第二理论出入次数为2;
- [0138] 当第二消耗时长小于或等于第二正常耗时范围的下限值时,确定车辆的第二理论出入次数为0。
- [0139] 在一些实施方式中,销量预测模块505具体用于:
- [0140] 将第一销量信息、第二销量信息和第三销量信息代入预测公式,得到预测销量信息;预测公式包括:
- [0141] $T = \alpha S1 + \beta S2 + \gamma S3 + d$
- [0142] 其中,T表示预测销量信息,S1表示第一销量信息,S2表示第二销量信息,S3表示第三销量信息, α 、 β 和 γ 分别为预设的权重,d表示预设的误差值。
- [0143] 在一些实施方式中,该煤矿销量数据的预测系统还包括模型训练模块,其用于:
- [0144] 以第一销量信息、第二销量信息、第三销量信息作为特征数据,以理论销量信息作为训练目标,对机器学习模型进行训练。
- [0145] 在一些实施方式中,车牌识别数据还包括:入矿时间,磅单数据还包括:空车过磅的第二空车时间和重车过磅的第二重车时间,磅单数据包括:车辆的净重和核定载重;模型训练模块在理论销量信息的获取过程中还用于:
- [0146] 将进入区域时间、入矿时间、第一空车时间和第二空车时间中的最小时间确定为开始时间;
- [0147] 将离开区域时间、出矿时间、第一重车时间和第二重车时间中的最大时间确定为离开时间;
- [0148] 将离开时间和开始时间之间的时长确定车辆在拉煤销量流程中的第三消耗时长;
- [0149] 将第三消耗时长与预设的第三正常耗时范围进行比较,并根据比较结果确定车辆的第三理论出入次数;其中,第三理论出入次数还用于机器学习模型的训练过程;
- [0150] 根据第三理论出入次数和车辆的核定载重,生成车辆在预设时间段内的理论销量信息。
- [0151] 在一些实施方式中,系统还包括标记模块,其用于:根据第三消耗时长与第三正常

耗时范围之间的比较结果,标记车辆出入次数的类型为正常、异常或超时。

[0152] 本实施例提供的煤矿销量数据的预测系统,与上述实施例提供的煤矿销量数据的预测方法具有相同的技术特征,所以也能解决相同的技术问题,达到相同的技术效果。为简要描述,实施例部分未提及之处,可参考前述煤矿销量数据的预测方法实施例中相应内容。

[0153] 本实施例还提供一种电子设备,该电子设备的结构示意图如图6所示,该设备包括处理器101和存储器102;其中,存储器102用于存储一条或多条计算机指令,一条或多条计算机指令被处理器执行,以实现上述煤矿销量数据的预测方法。

[0154] 图6所示的电子设备还包括总线103和通信接口104,处理器101、通信接口104和存储器102通过总线103连接。

[0155] 其中,存储器102可能包含高速随机存取存储器(RAM,Random Access Memory),也可能还包括非不稳定的存储器(non-volatile memory),例如至少一个磁盘存储器。总线103可以是ISA总线、PCI总线或EISA总线等。所述总线可以分为地址总线、数据总线、控制总线等。为便于表示,图6中仅用一个双向箭头表示,但并不表示仅有一根总线或一种类型的总线。

[0156] 通信接口104用于通过网络接口与至少一个用户终端及其它网络单元连接,将封装好的IPv4报文或IPv4报文通过网络接口发送至用户终端。

[0157] 处理器101可能是一种集成电路芯片,具有信号的处理能力。在实现过程中,上述方法的各步骤可以通过处理器101中的硬件的集成逻辑电路或者软件形式的指令完成。上述的处理器101可以是通用处理器,包括中央处理器(Central Processing Unit,简称CPU)、网络处理器(Network Processor,简称NP)等;还可以是数字信号处理器(Digital Signal Processor,简称DSP)、专用集成电路(Application Specific Integrated Circuit,简称ASIC)、现场可编程门阵列(Field-Programmable Gate Array,简称FPGA)或者其他可编程逻辑器件、分立门或者晶体管逻辑器件、分立硬件组件。可以实现或者执行本公开实施例中的公开的各方法、步骤及逻辑框图。通用处理器可以是微处理器或者该处理器也可以是任何常规的处理器等。结合本公开实施例所公开的方法的步骤可以直接体现为硬件译码处理器执行完成,或者用译码处理器中的硬件及软件模块组合执行完成。软件模块可以位于随机存储器,闪存、只读存储器,可编程只读存储器或者电可擦写可编程存储器、寄存器等本领域成熟的存储介质中。该存储介质位于存储器102,处理器101读取存储器102中的信息,结合其硬件完成前述实施例的方法的步骤。

[0158] 本发明实施例还提供了一种可读存储介质,该可读存储介质上存储有计算机程序,该计算机程序被处理器运行时执行前述实施例的煤矿销量数据的预测方法的步骤。

[0159] 在本申请所提供的几个实施例中,应该理解到,所揭露的系统、设备和方法,可以通过其它的方式实现。以上所描述的装置实施例仅仅是示意性的,例如,所述单元的划分,仅仅为一种逻辑功能划分,实际实现时可以有另外的划分方式,又例如,多个单元或组件可以结合或者可以集成到另一个系统,或一些特征可以忽略,或不执行。另一点,所显示或讨论的相互之间的耦合或直接耦合或通信连接可以是通过一些通信接口,设备或单元的间接耦合或通信连接,可以是电性,机械或其它的形式。

[0160] 所述作为分离部件说明的单元可以是或者也可以不是物理上分开的,作为单元显示的部件可以是或者也可以不是物理单元,即可以位于一个地方,或者也可以分布到多个

网络单元上。可以根据实际的需要选择其中的部分或者全部单元来实现本实施例方案的目的。

[0161] 另外,在本发明各个实施例中的各功能单元可以集成在一个处理单元中,也可以是各个单元单独物理存在,也可以两个或两个以上单元集成在一个单元中。

[0162] 所述功能如果以软件功能单元的形式实现并作为独立的产品销售或使用,可以存储在一个处理器可执行的非易失的计算机可读取存储介质中。基于这样的理解,本发明的技术方案本质上或者说对现有技术做出贡献的部分或者该技术方案的部分可以用软件产品的形式体现出来,该计算机软件产品存储在一个存储介质中,包括若干指令用以使得一台计算机设备(可以是个人计算机,服务器,或者网络设备等)执行本发明各个实施例所述方法的全部或部分步骤。而前述的存储介质包括:U盘、移动硬盘、只读存储器(ROM,Read-Only Memory)、随机存取存储器(RAM,Random Access Memory)、磁碟或者光盘等各种可以存储程序代码的介质。

[0163] 最后应说明的是:以上所述实施例,仅为本发明的具体实施方式,用以说明本发明的技术方案,而非对其限制,本发明的保护范围并不局限于此,尽管参照前述实施例对本发明进行了详细的说明,本领域的普通技术人员应当理解:任何熟悉本技术领域的技术人员在本发明揭露的技术范围内,其依然可以对前述实施例所记载的技术方案进行修改或可轻易想到变化,或者对其中部分技术特征进行等同替换;而这些修改、变化或者替换,并不使相应技术方案的本质脱离本发明实施例技术方案的精神和范围,都应涵盖在本发明的保护范围之内。因此,本发明的保护范围应所述以权利要求的保护范围为准。

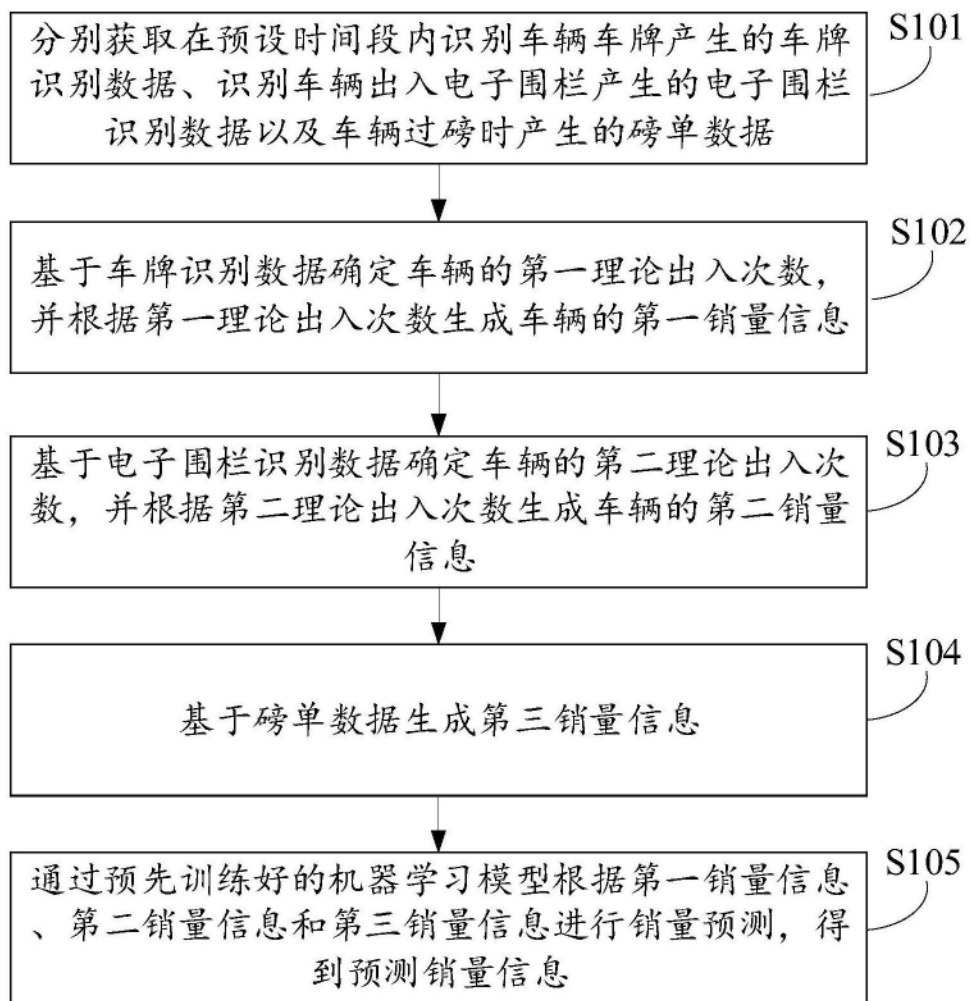


图1

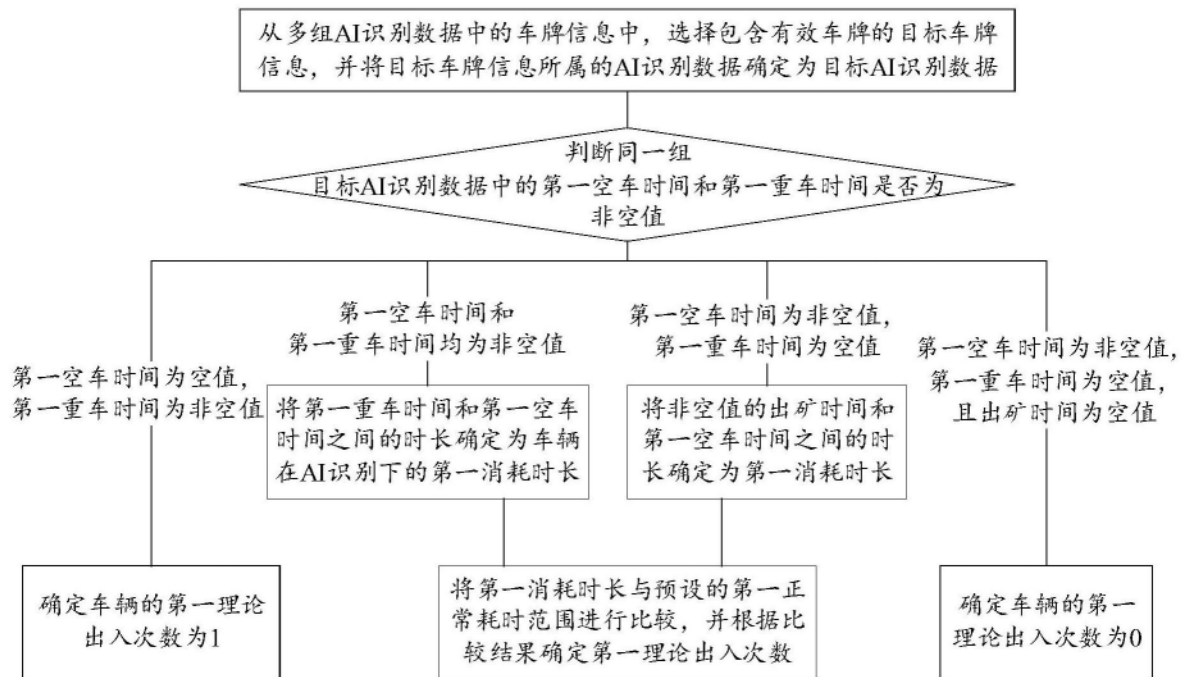


图2

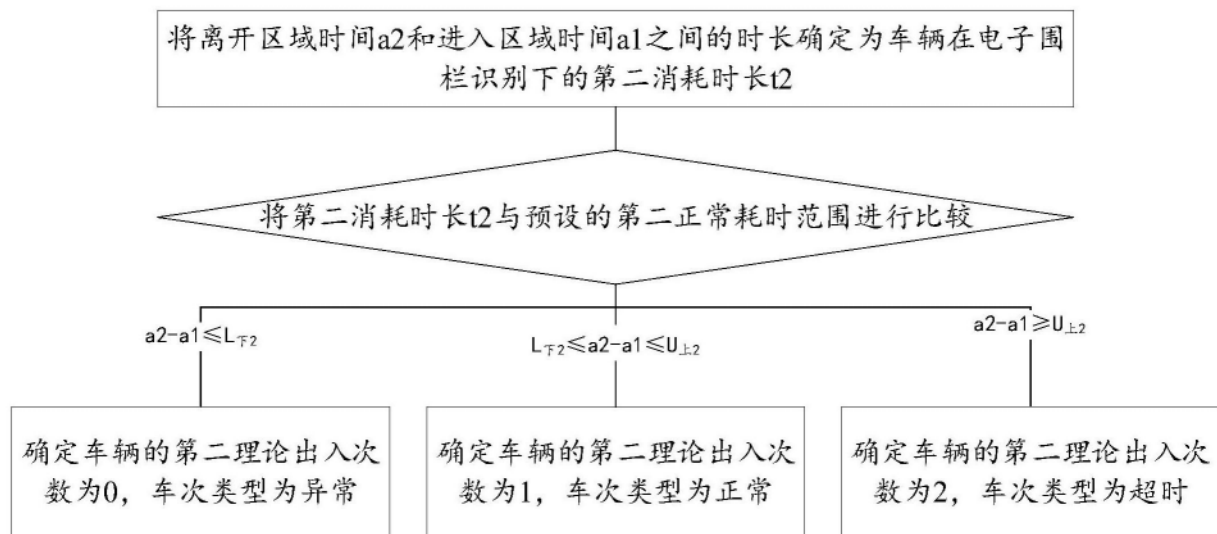


图3

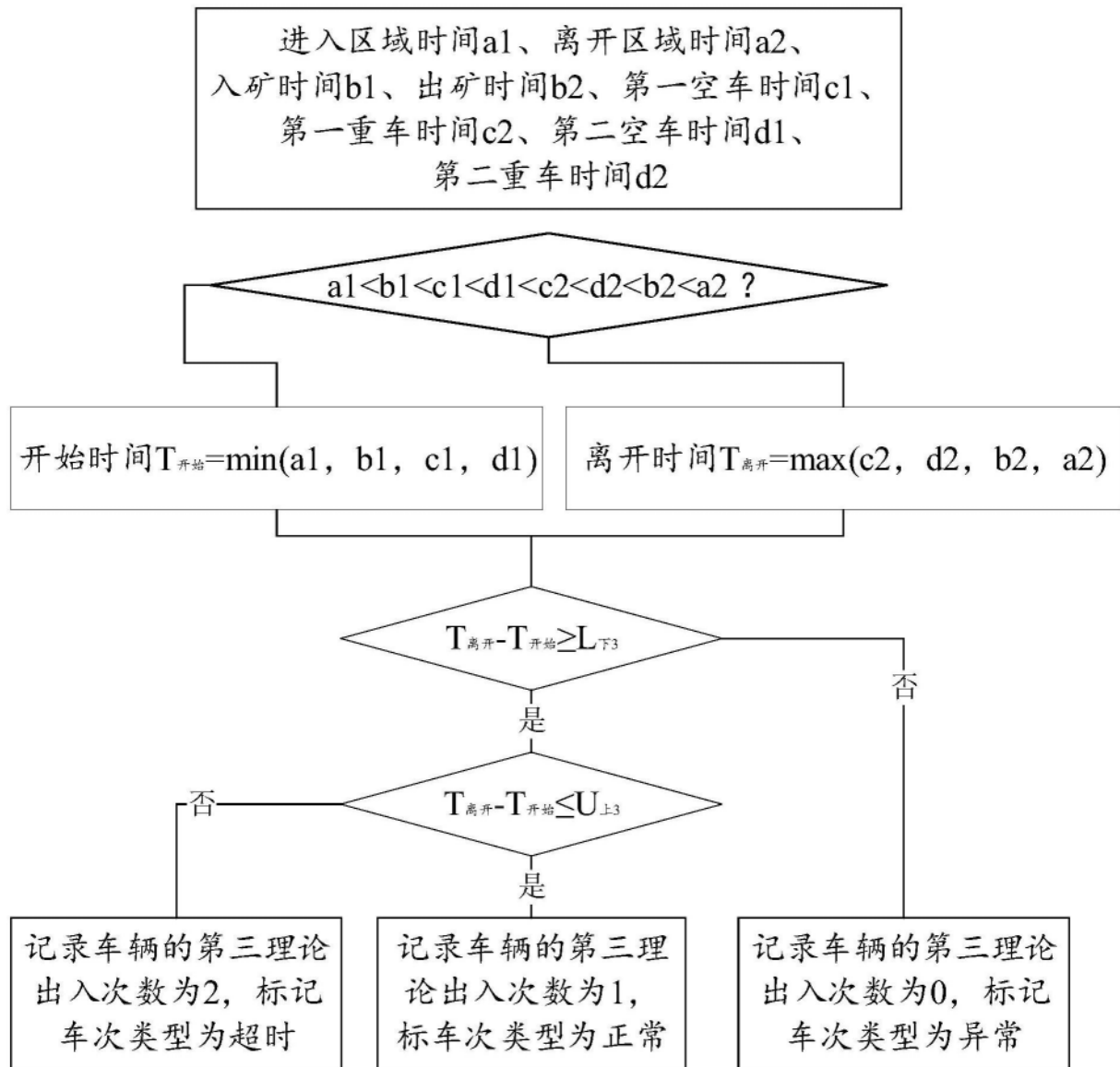


图4

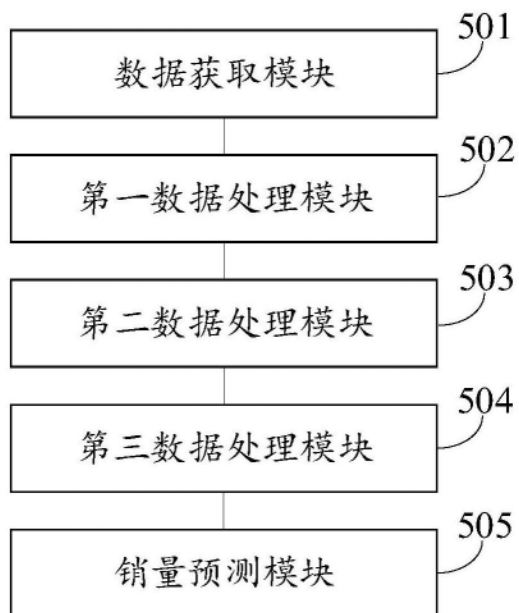


图5

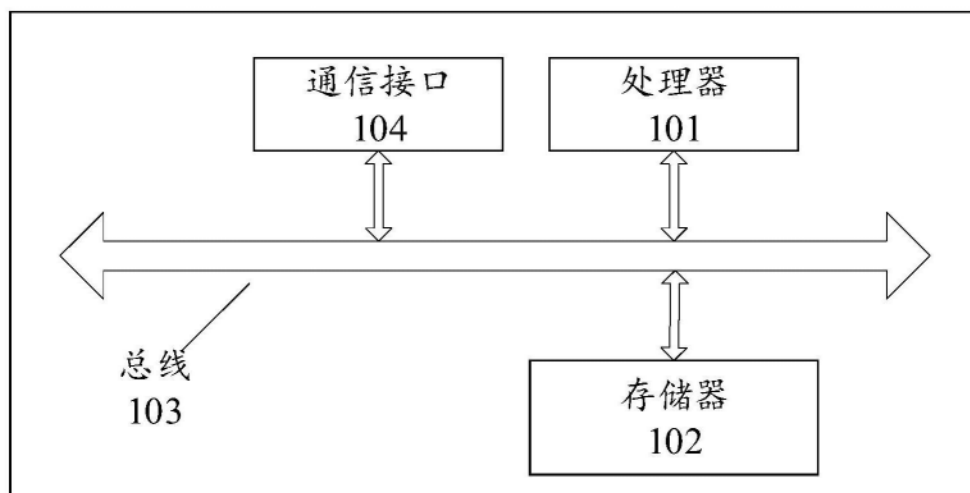


图6