



(12) 发明专利

(10) 授权公告号 CN 110392899 B

(45) 授权公告日 2023.09.15

(21) 申请号 201880015774.8

(74) 专利代理机构 中国贸促会专利商标事务所
有限公司 11038

(22) 申请日 2018.11.02

专利代理人 刘玉洁

(65) 同一申请的已公布的文献号

申请公布号 CN 110392899 A

(51) Int.CI.

G06Q 30/02 (2006.01)

(43) 申请公布日 2019.10.29

(56) 对比文件

(30) 优先权数据

15/844,991 2017.12.18 US

CN 103262069 A, 2013.08.21

(85) PCT国际申请进入国家阶段日

CN 107273454 A, 2017.10.20

2019.09.04

US 2014257924 A1, 2014.09.11

(86) PCT国际申请的申请数据

US 2017076299 A1, 2017.03.16

PCT/US2018/058919 2018.11.02

US 2014289167 A1, 2014.09.25

(87) PCT国际申请的公布数据

JP 2015076091 A, 2015.04.20

W02019/125612 EN 2019.06.27

JP 2009104408 A, 2009.05.14

(73) 专利权人 甲骨文国际公司

US 2015379427 A1, 2015.12.31

地址 美国加利福尼亚

US 2016055426 A1, 2016.02.25

(72) 发明人 M·雷 C·波佩斯库

US 2016162779 A1, 2016.06.09

US 2017330123 A1, 2017.11.16

WO 2016148107 A1, 2016.09.22

审查员 仵凡

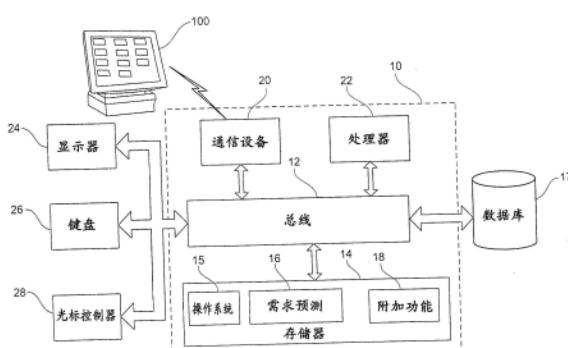
权利要求书3页 说明书13页 附图10页

(54) 发明名称

用于模型生成的动态特征选择

(57) 摘要

实施例生成包括优化的特征集的产品需求模型。实施例接收产品的销售历史并接收产品的相关特征集，并将相关特征的子集指定为强制特征。根据销售历史，实施例形成训练数据集和验证数据集，并从该相关特征集中随机选择一个或多个可选特征。实施例包括所选择的可选特征以及强制特征以创建特征测试集。实施例使用训练数据集和特征测试集来训练算法以生成经训练的算法并使用经训练的算法和验证数据集来计算提前停止度量。当提前停止度量低于预定义阈值时，特征测试集是优化的特征集。



1. 一种生成包括优化的特征集的产品需求模型的方法,该方法包括:
接收产品的销售历史;
接收产品的相关特征集,并将相关特征的子集指定为强制特征;
根据销售历史形成训练数据集和验证数据集;
从所述相关特征集中随机选择一个或多个可选特征以创建可选特征集合,并包括所述可选特征集合以及强制特征以创建第一特征测试集;
使用训练数据集和第一特征测试集来训练算法以生成经训练的算法;
使用经训练的算法和验证数据集来计算提前停止度量;
重复所述随机选择以创建修改的特征测试集,所述修改的特征测试集包括所述强制特征以及与第一特征测试集的可选特征集合不同的可选特征集合;
使用所述修改的特征测试集而不是第一特征测试集来重复所述训练和计算;以及
重复执行所述重复所述随机选择以及重复所述训练和计算直到所述提前停止度量低于预定义阈值;
其中当所述提前停止度量低于所述预定义阈值时,所述修改的特征测试集是所述优化的特征集。

2. 如权利要求1所述的方法,还包括重复所述方法以生成多个优化的特征集,其中每个优化的特征集被输入到预测算法中以生成经训练的模型。
3. 如权利要求1所述的方法,其中形成训练数据集和验证数据集包括随机选择销售历史的第一子集作为训练数据集,并使用销售历史的剩余部分作为验证数据集。
4. 如权利要求1所述的方法,其中所述算法包括机器学习算法,所述机器学习算法包括线性回归、支持向量机或人工神经网络之一。
5. 如权利要求1所述的方法,其中所述提前停止度量包括平均绝对百分比误差。
6. 如权利要求2所述的方法,还包括:

使用训练集训练与优化的特征集对应的多个模型,并使用对应的验证集来验证每个经训练的模型以及计算误差;
计算用于每个模型的模型权重;
输出模型组合,所述模型组合包括用于每个模型的预测和权重;以及
基于所述模型组合生成未来销售的预测。

7. 如权利要求6所述的方法,其中所述误差是均方根误差RMSE,并且对于每个训练集*i*的每个模型,计算模型权重w(*i*)包括:

$$w(i) = \frac{1}{1+RMSE(i)}.$$

8. 一种计算机可读介质,其上存储有指令,所述指令在由处理器执行时使所述处理器生成优化的特征集,所述生成包括:

接收产品的销售历史;
接收产品的相关特征集,并将相关特征的子集指定为强制特征;
根据销售历史形成训练数据集和验证数据集;
从所述相关特征集中随机选择一个或多个可选特征以创建可选特征集合,并包括所述可选特征集合以及强制特征以创建第一特征测试集;

使用训练数据集和第一特征测试集来训练算法以生成经训练的算法；

使用经训练的算法和验证数据集来计算提前停止度量；

重复所述随机选择以创建修改的特征测试集，所述修改的特征测试集包括所述强制特征以及与第一特征测试集的可选特征集合不同的可选特征集合；

使用所述修改的特征测试集而不是第一特征测试集来重复所述训练和计算；以及

重复执行所述重复所述随机选择以及重复所述训练和计算直到所述提前停止度量低于预定义阈值；

其中当所述提前停止度量低于所述预定义阈值时，所述修改的特征测试集是所述优化的特征集。

9. 如权利要求8所述的计算机可读介质，所述生成还包括重复所述生成以生成多个优化的特征集，其中每个优化的特征集被输入到预测算法中以生成经训练的模型。

10. 如权利要求8所述的计算机可读介质，其中形成训练数据集和验证数据集包括随机选择销售历史的第一子集作为训练数据集，并使用销售历史的剩余部分作为验证数据集。

11. 如权利要求8所述的计算机可读介质，其中所述算法包括机器学习算法，所述机器学习算法包括线性回归、支持向量机或人工神经网络之一。

12. 如权利要求8所述的计算机可读介质，其中所述提前停止度量包括平均绝对百分比误差。

13. 如权利要求9所述的计算机可读介质，还包括：

使用训练集训练与优化的特征集对应的多个模型，并使用对应的验证集来验证每个经训练的模型以及计算误差；

计算用于每个模型的模型权重；

输出模型组合，所述模型组合包括用于每个模型的预测和权重；以及

基于所述模型组合生成未来销售的预测。

14. 如权利要求13所述的计算机可读介质，其中所述误差是均方根误差RMSE，并且对于每个训练集*i*的每个模型，计算模型权重w(*i*)包括：

$$w(i) = \frac{1}{1+RMSE(i)}.$$

15. 一种零售销售预测系统，包括：

处理器，耦接到存储设备，所述处理器实现需求预测模块，包括：

接收产品的销售历史；

接收产品的相关特征集，并将相关特征的子集指定为强制特征；

根据销售历史形成训练数据集和验证数据集；

从所述相关特征集中随机选择一个或多个可选特征以创建可选特征集合，并包括所述可选特征集合以及强制特征以创建第一特征测试集；

使用训练数据集和第一特征测试集来训练算法以生成经训练的算法；

使用经训练的算法和验证数据集来计算提前停止度量；

重复所述随机选择以创建修改的特征测试集，所述修改的特征测试集包括所述强制特征以及与第一特征测试集的可选特征集合不同的可选特征集合；

使用所述修改的特征测试集而不是第一特征测试集来重复所述训练和计算；以及

重复执行所述重复所述随机选择以及重复所述训练和计算直到所述提前停止度量低于预定义阈值；

其中当所述提前停止度量低于所述预定义阈值时，所述修改的特征测试集是优化的特征集。

16. 如权利要求15所述的零售销售预测系统，还包括生成多个优化的特征集，其中每个优化的特征集被输入到预测算法中以生成经训练的模型。

17. 如权利要求15所述的零售销售预测系统，其中形成训练数据集和验证数据集包括随机选择销售历史的第一子集作为训练数据集，并使用销售历史的剩余部分作为验证数据集。

18. 如权利要求15所述的零售销售预测系统，其中所述算法包括机器学习算法，所述机器学习算法包括线性回归、支持向量机或人工神经网络之一。

19. 如权利要求15所述的零售销售预测系统，其中所述提前停止度量包括平均绝对百分比误差。

20. 如权利要求16所述的零售销售预测系统，还包括：

使用训练集训练与优化的特征集对应的多个模型，并使用对应的验证集来验证每个经训练的模型以及计算误差；

计算用于每个模型的模型权重；

输出模型组合，所述模型组合包括用于每个模型的预测和权重；以及

基于所述模型组合生成未来销售的预测。

21. 如权利要求20所述的零售销售预测系统，其中所述误差是均方根误差RMSE，并且对于每个训练集*i*的每个模型，计算模型权重w(*i*)包括：

$$w(i) = \frac{1}{1+RMSE(i)}.$$

22. 一种包括用于执行如权利要求1-7中任一项所述的方法的部件的装置。

用于模型生成的动态特征选择

技术领域

[0001] 一个实施例一般而言涉及计算机系统,具体而言涉及动态选择用于模型生成的特征的计算机系统。

背景技术

[0002] 产品通常通过制造商、分销商、运输商、零售商等的网络递送给消费者。一起向消费者递送产品的这种设施网络通常被称为“供应链”网络。

[0003] 产品的供应者(例如,制造商、供应商、零售商等)常常面临预测产品需求的任务,以便在存在不断变化的市场条件的情况下通过供应链网络提供平稳且高效的产品流。过高估计需求可能导致生产过剩和与持有库存相关联的成本增加(例如,存储成本、过时(obsolescence)等)。另一方面,低估需求会导致收入损失。

[0004] 另外,在零售行业,零售商需要预测他们未来的需求,以更好地管理他们的库存或促销/降价计划。零售商可以参与许多类型的促销以提高他们的销售额。为了生成准确的预测,零售商必须考虑可能影响需求的所有因素/特征,诸如促销、价格、季节性、天气等。

[0005] 用于预测产品需求的一种技术是主要基于该产品的历史需求信息(例如,基于过去的采购订单、过去的发货情况、过去的销售点数据等)来预测需求。但是,这种技术可能难以适应不断变化的市场条件,并且可能导致不准确的预测。另外,随着越来越多的因素需要考虑,传统的预测方法(诸如时间序列分析或回归)不能非常好地工作。

发明内容

[0006] 实施例生成包括优化的特征集的产品需求模型。实施例接收产品的销售历史并接收产品的相关特征集,并将相关特征的子集指定为强制特征。根据销售历史,实施例形成训练数据集和验证数据集,并从该相关特征集中随机选择一个或多个可选特征。实施例包括所选择的可选特征以及强制特征以创建特征测试集。实施例使用训练数据集和特征测试集来训练算法以生成经训练的算法,并使用经训练的算法和验证数据集来计算提前停止度量。当提前停止度量低于预定义阈值时,特征测试集是优化的特征集,并且当提前停止度量不低于预定义阈值时,实施例重复随机选择、训练和计算。

附图说明

[0007] 图1是根据本发明实施例的计算机服务器/系统的框图。

[0008] 图2是根据一个实施例的图1的需求预测模块在确定一个或多个优化的特征集时的功能的流程图,其中该一个或多个优化的特征集各自包括强制特征和一个或多个可选特征,并且可以用于需求预测。

[0009] 图3示出了根据本发明实施例的产品的特征集的简化示例。

[0010] 图4A和图4B示出了根据一个实施例的在执行三轮之后使用图3的特征集示例的图2的功能的输出的示例。

[0011] 图5是根据一个实施例的当使用多个经训练的模型确定聚合级别的促销效果时图1的需求预测模块的功能的流程图。

[0012] 图6示出了根据一个实施例的使用数据点的六轮模型估计。

[0013] 图7A和图7B示出了使用本发明实施例的预测与实际销售的比较。

[0014] 图8示出了根据一个实施例的包括如本文所公开的需求预测的集成的制造、库存和物流系统。

具体实施方式

[0015] 一个实施例确定物品的相关特征集，以便确定该物品的需求预测。特征集包括强制特征和可选特征。所确定的特征被用于训练随后用于生成需求预测的一个或多个模型。

[0016] 销售和需求预测方法大致可以被分为判断方法、外推方法和因果方法。外推方法仅使用活动本身的时间序列数据来生成预测。已知的特定算法/方法的范围从较简单的移动平均和指数平滑方法到更复杂的Box-Jenkins方法。虽然这些已知方法成功地识别和外推趋势、季节性和自相关的时间序列模式，但它们不考虑诸如价格改变和促销之类的外部因素。

[0017] 向量自回归(“VAR”)方法扩展Box-Jenkins方法以包括其它变量，但它们的复杂性使得估计困难。因果预测涉及使用表示被认为是结果的驱动因素的现象的输入来构建定量模型。这些方法可以像具有促销变量的线性回归算法一样简单。起点是具有促销变量(诸如降价、折扣或广告)的回归模型。想法是模型的简单性帮助管理者理解以及批准或指导对模型的修改，并且随着他们对决策辅助知识的了解越来越多，他们可以准备好实现更精细和更复杂的模型。

[0018] 为了改进需求预测，零售商已经开始转向现代机器学习技术，诸如支持向量机(“SVM”)、人工神经网络(“ANN”)、随机森林等。但是，通常零售商将只会为每个产品/位置挑选一个模型。如本文所使用的，零售商可以包括单个零售商店，或者可以包括全部由单个或多个物流运作集成和管理的大量零售商店。

[0019] 另外，对于许多机器学习算法(SVM、ANN、随机森林等)，零售商将使用特征集(即，物品的各种属性)来定义产品/位置/日历交叉点处的数据点。在这些算法中，零售商将使用与用于预测的相同的特征集来训练模型。另外，相同的特征集可以被若干不同的算法用于进行预测。“特征集”是影响物品的需求或销售以及描述物品属性的特征集合。特征的示例包括基本销售额、价格、季节性、品牌、促销、尺寸、颜色、包装尺寸、供应商、长度等。虽然诸如价格和季节性之类的特征可以对于所有类型的产品是相关的，但是一些其它特征是特定于物品的。例如，包装尺寸影响对酸奶的需求，但是包装的长度是无关紧要的。相反，品牌对于时尚产品非常重要，但对于诸如钉子或锤子之类的五金产品则不那么重要。

[0020] 为了生成准确的预测，零售商必须考虑可能影响需求/销售的所有因素/特征，诸如促销、价格变化、季节性、天气等。但是，可能的特征集可能会变得非常大(诸如20-50个不同的可能特征)，以至于传统的需求预测工具(诸如线性回归)无法处理如此多的独立变量。虽然为了获得更准确的需求预测而需要多少特征可以没有限制，但是对于一个物品/位置而言重要的特征中的一些特征对于另一个物品/位置可能不那么重要。

[0021] 已知的解决方案使用“经验法则”技术来生成用于需求预测的特征集。与已知方法

相比,实施例通过从必须包括在任何特征集中的“强制”特征集开始来自动确定特征集,然后确定最佳“可选”特征以完成特征集。

[0022] 图1是根据本发明实施例的计算机服务器/系统10的框图。虽然被示为单个系统,但是系统10的功能可以被实现为分布式系统。另外,本文公开的功能可以在可以经网络耦接在一起的单独的服务器或设备上实现。另外,可以不包括系统10的一个或多个组件。例如,对于服务器的功能,系统10可以需要包括处理器和存储器,但是可以不包括图1中所示的其它组件中的一个或多个,诸如键盘或显示器。

[0023] 系统10包括用于传送信息的总线12或其它通信机制,以及耦接到总线12用于处理信息的处理器22。处理器22可以是任何类型的通用或专用处理器。系统10还包括用于存储要由处理器22执行的信息和指令的存储器14。存储器14可以包括随机存取存储器(“RAM”)、只读存储器(“ROM”)、诸如磁盘或光盘之类的静态存储装置,或任何其它类型的计算机可读介质。系统10还包括通信设备20,诸如网络接口卡,以提供对网络的访问。因此,用户可以直接受地或通过网络远程地或以任何其它方法与系统10对接。

[0024] 计算机可读介质可以是由处理器22访问的任何可用介质,并且包括易失性和非易失性介质、可移动和不可移动介质、以及通信介质。通信介质可以包括计算机可读指令、数据结构、程序模块或调制数据信号(诸如载波或其它传输机制)中的其它数据,并且包括任何信息递送介质。

[0025] 处理器22还经由总线12耦接到诸如液晶显示器(“LCD”)之类的显示器24。键盘26和诸如计算机鼠标之类的光标控制设备28还耦接到总线12,以使用户能够与系统10对接。

[0026] 在一个实施例中,存储器14存储当由处理器22执行时提供功能的软件模块。模块包括为系统10提供操作系统功能的操作系统15。模块还包括需求预测模块16,需求预测模块16确定用于需求预测的最佳特征集,以及本文公开的所有其它功能。系统10可以是更大系统的一部分。因此,系统10可以包括一个或多个附加的功能模块18,以包括附加功能,诸如零售管理系统(例如,来自Oracle公司的“Oracle零售需求预测系统”或“Oracle零售高级科学引擎”(“ORASE”))或企业资源计划(“ERP”)系统。数据库17耦接到总线12,以便为模块16和18提供集中式存储并且存储客户数据、产品数据、交易数据等等。在一个实施例中,数据库17是可以使用结构化查询语言(“SQL”)来管理所存储的数据的关系数据库管理系统(“RDBMS”)。在一个实施例中,专用销售点(“POS”)终端100生成用于预测需求的交易数据和历史销售数据(例如,关于每个零售商店处的每个物品/SKU的交易的数据)。根据一个实施例,POS终端100本身可以包括用于预测需求的附加处理功能。

[0027] 在一个实施例中,特别是当存在大量零售商店、大量商品和大量历史数据时,数据库17被实现为存储器内数据库(“IMDB”)。IMDB是一种数据库管理系统,该数据库管理系统主要依赖于主存储器来存储计算机数据。它与采用盘存储机制的数据库管理系统形成对比。因为盘存取比存储器存取慢,所以主存储器数据库比盘优化的数据库更快,内部优化算法更简单,执行更少的CPU指令。存取存储器中的数据消除了查询数据时的查找时间,这提供比盘更快、更可预测的性能。

[0028] 在一个实施例中,当被实现为IMDB时,数据库17基于分布式数据网格来实现。分布式数据网格是其中计算机服务器的集合在一个或多个集群中一起工作以在分布式或集群环境内管理信息和相关操作(诸如计算)的系统。分布式数据网格可以被用于管理跨服务器

共享的应用对象和数据。分布式数据网格提供低响应时间、高吞吐量、可预测的可扩展性、连续可用性和信息可靠性。在特定示例中，分布式数据网格（诸如，例如来自Oracle公司的“Oracle Coherence”数据网格）将信息存储在存储器内以实现更高的性能，并且采用冗余来保持该信息的副本跨多个服务器同步，从而确保系统的弹性以及在服务器发生故障时数据的持续可用性。

[0029] 在一个实施例中，系统10是计算/数据处理系统，该计算/数据处理系统包括用于企业组织的分布式应用的集合或应用，并且系统10还可以实现物流、制造和库存管理功能。应用和计算系统10可以被配置为与基于云的网络系统、软件即服务（“SaaS”）体系架构或其它类型的计算解决方案一起操作，或者被实现为基于云的网络系统、软件即服务（“SaaS”）体系架构或其它类型的计算解决方案。

[0030] 实施例使用自动确定的一个或多个特征集来生成根据一个或多个不同算法生成的一个或多个经训练的模型，以便确定销售预测或需求预测。预测是供应链的重要驱动因素。如果预测不准确，那么分配和补货执行不佳，从而导致零售商的财务损失。通过本文公开的实施例可以实现对促销或非促销物品的预测准确性的改进。另外，可以更好地理解促销对需求的影响。例如，这有助于零售商更有效地计划有关渠道、定价和客户群的促销活动。

[0031] 从以下角度公开实施例：对于在某个地点（例如，零售场所）销售的物品（即，诸如酸奶或男士衬衫之类的一类物品），可以在各个时间（即，预定义的零售时段，诸如一天、一周、一个月、一年等）以各种方式促销物品。零售日历具有许多零售时段（例如，许多周），这些零售时段在典型的日历年中以特定方式（例如，四（4）个十三（13）周的季度）组织。零售时段可以在过去或未来发生。历史销售/业绩数据可以包括例如在多个过去零售时段的每一个中销售的物品的多个单元以及相关联的促销数据（即，对于每个零售时段，哪些促销对于该时段是有效的）。

[0032] 如下面所公开的，实施例使用根据一个或多个不同算法和一个或多个特征集生成的一个或多个经训练的模型，并且可以最终组合来自多个经训练的模型的预测以得到最终需求预测。在一些实施例中使用的经训练的模型可以包括经训练的线性回归模型或机器学习技术，诸如决策树或回归树、支持向量机（“SVM”）或神经网络。

[0033] 结合根据利用不同输入和特征集训练的线性回归算法生成的经训练的线性回归模型，对输出变量与多个输入变量之间的线性关系的搜索导致在回归设置中逐步选择输入变量。在一些实施例中，目标是构建将输出变量表达为输入变量的线性函数加上常数的函数。逐步回归中的两种一般方法是前向选择和后向选择。

[0034] 在前向选择中，根据预定的标准，基于变量对模型的贡献，一次一个地引入变量。在后向选择中，起初所有输入变量都被构建到模型中，然后如果基于预定标准判断输入变量对模型没有贡献，那么从回归公式中移除它们。

[0035] 在机器学习中，SVM是具有相关联的学习算法的受监督的学习模型，该学习算法分析用于分类和回归分析的数据。给定训练示例集，每个训练示例被标记为属于两个类别中的一个或另一个，SVM训练算法构建将新的示例指派给一个类别或另一个类别的模型，从而使它成为非概率的二元线性分类器。SVM模型是作为空间中的点的示例的表示，这些示例被映射为使得单独类别的示例被尽可能宽的明确间隙划分。然后将新的示例映射到同一个空

间中，并基于它们落在间隙的哪一侧来预测新的示例属于一个类别。

[0036] 除了分类之外，SVM还成功应用于销售或需求预测、能够处理常见的度量，诸如销售额，以及价格、促销、外部因素(诸如天气和人口统计信息)。

[0037] SVM及其支持向量回归 (“SVR”) 的回归版本使用内核函数隐式地将实例映射到更高维度的特征空间中。在它的最基本的形式中，SVR理想地寻求在这个空间中识别与被映射的输出点在一定距离内的线性函数。这种“软边际公式”允许并惩罚超出预定距离的偏差，并最小化违规的总和以及识别线性关系的向量的范数。

[0038] 回归树技术以决策树格式将数据划分为较小的子集，并在用于预测结果的每个叶子处拟合线性回归模型。替代模型树方法的彼此不同之处主要在于要被分支的输入变量的选择标准、所使用的拆分标准以及在树的每个叶子处构建的模型。虽然在针对特定情况的预测可以追溯到树中的条件和适用于满足那些条件的情况的回归函数的意义上树是透明的，但是具有许多层的树不容易以一般化的方式解释。

[0039] 人工神经网络 (“ANN”) 是一种信息处理范式，其灵感来自生物神经系统(诸如大脑)处理信息的方式。这种模型的关键要素是信息处理系统的新颖结构。它由大量高度互连的处理元件(即，神经元)组成，这些处理元件协同工作以解决具体问题。ANN通过示例进行学习。通过学习过程为具体应用(诸如模式识别或数据分类)配置ANN。在生物系统中的学习涉及调整神经元之间存在的突触连接。对于ANN也是如此。由于神经网络最擅长识别数据的模式或趋势，因此它们非常适合预计或预测需求。

[0040] 图2是根据一个实施例的图1的需求预测模块16在确定一个或多个优化的特征集时的功能的流程图，该一个或多个优化的特征集各自包括强制特征和一个或多个可选特征，并且可以用于需求预测。在一个实施例中，图2(以及下面的图5)的流程图的功能由存储在存储器或其它计算机可读或有形介质中的软件实现，并由处理器执行。在其它实施例中，功能可以由硬件(例如，通过使用专用集成电路 (“ASIC”)、可编程门阵列 (“PGA”)、现场可编程门阵列 (“FPGA”) 等)执行，或由硬件和软件的任意组合执行。

[0041] 在202处，针对特定类/类别的产品针对所有商店的所有物品或者仅针对单个感兴趣的物品接收历史物品销售数据。例如，类/类别可以是“酸奶”、“咖啡”或“牛奶”。每个类具有一个或多个子类，一直到SKU或通用产品代码 (“UPC”) 级别，这将是每个单独的待售物品，例如，对于酸奶类，子类可以是每个品牌的酸奶，并且进一步的子类可以是风味、尺寸、类型(例如，希腊或常规)，直到与所售出的每种单独的不同类型的酸奶物品对应的SKU。在一个实施例中，从图2的功能生成的确定的特征集针对给定的产品(即，给定地点的类别，诸如Baltimore, MD区域中的酸奶)。

[0042] 历史销售和业绩数据可以包括例如表示跨多个过去零售时段的物品的过去销售和促销的数据。历史业绩数据可以被分段为过去几周的零售时段，其中过去的每周具有指派给它的数值以指示该周销售的物品的数量。根据一个实施例，历史业绩数据还可以包括表示跨零售时段的价格折扣和其它促销组成部分的值的数值。根据一个实施例，可以经由网络通信访问物品的历史性能数据，包括从每个零售商店的每个POS终端100访问和/或从数据库17访问物品的历史性能数据。

[0043] 历史业绩数据包括跨多个时间段(例如，多个周)与多个促销组成部分相关联的销售数据。促销组成部分的示例包括但不限于价格折扣组成部分、电视广告组成部分、收音机

广告组成部分、报纸广告组成部分、电子邮件广告组成部分、互联网广告组成部分和店内广告组成部分。

[0044] 汇集所有有效数据点以形成具有给定聚合级别的N个数据点的训练数据集D。聚合级别是高于汇集数据的SKU/商店/周的交叉点。聚合级别的示例是子类/商店。这个级别可用的数据由特定子类中的所有SKU确定。实施例中的聚合级别通常被选择为足够低以捕获商品的低级细节,但也足够高以使数据池足够丰富以用于对促销效果的健壮估计。另一个示例聚合级别是产品线/地理区域,诸如在Baltimore, MD区域销售的所有酸奶。

[0045] 在204处,接收产品的相关特征集。每个特征都描述产品并且可以影响产品的销售。在所有特征当中,识别出强制特征集(“M”)。在一个实施例中,强制特征集可以由零售商自己基于过去的知识或个人偏好来手动识别。例如,对于T恤衫,零售商可能认为“颜色”是强制特征,即使过去的数据指示它对需求/销售的影响相对小。在另一个实施例中,可以基于先前的销售历史自动生成强制特征的列表。然后,特征集的其余特征被视为可选特征集(“0”)。

[0046] 图3示出了根据本发明实施例的产品的特征集的简化示例。在图3的示例中,产品是酸奶,并且特征集包括10个特征:季节性、价格、品牌、包装尺寸、物品尺寸、风味、颜色、店内展示、端架(end cap)和买一送一(“Bogo”)促销(即,特征编号1-10)。带阴影的特征(季节性、价格、品牌、风味和Bogo)形成强制特征集(1、2、3、6、10)。其余特征(包装尺寸、物品尺寸、颜色、店内展示和端架)形成可选特征集(4、5、7、8、9)。虽然图3中所示的简化示例仅包括总共10个特征,但实际上,产品的特征集可以包括50-100个特征。

[0047] 在206处,提取产品的销售历史数据的全部或子集(例如,Baltimore地区的酸奶的两年销售历史)。销售历史数据的随机选择的部分被用作训练数据集,并且其余部分被用作验证数据集。在一个实施例中,随机选择80%的销售历史以形成训练数据集,并且剩余的20%用于形成验证数据集。

[0048] 在208处,已经被接收/选择的提前停止度量确定何时将完成迭代过程并且确定优化的特征集。在一个实施例中,使用平均绝对百分比误差(“MAPE”)并将MAPE与误差阈值“e”进行比较。另外,如果未确定优化的特征集,那么最大迭代次数被接收并且还确定何时完成迭代过程。在208处,确定是否已达到提前停止度量或最大迭代次数。

[0049] 如果在208处为否,那么在210处可选特征集0中的特征被随机选择并被放置在特征测试集“S”中。添加到集合S的可选特征的数量也是随机的。210的功能可以重复多次,并且可选特征以及它们中的多少被包括可以每次都不同,直到找到强制特征和可选特征的令人满意的组合。

[0050] 在212处,将集合M的所有强制特征添加到集合S中。

[0051] 在214处,使用来自206的训练数据集并使用特征测试集S的特征(即,强制特征和可选特征两者)来训练算法以生成经训练的算法(即,模型)。要训练的算法可以是任何期望的算法,诸如上面公开的算法(例如,线性回归、ANN等)。

[0052] 在216处,通过对验证数据集应用经训练的算法(即,模型)来计算提前停止度量,并且功能在208处继续,在208处确定提前停止度量是否低于阈值e或达到最大迭代。因此,重复步骤210、212、214和216。

[0053] 216的示例(其中提前停止度量是MAPE)如下:

$$[0054] \quad MAPE = \frac{100}{N} \sum_{i=1}^N \frac{|predicted\ value(i) - actual\ value(i)|}{|actual\ value(i)|}$$

[0055] 假定有三个时间段,这将使N=3。假设这三个时段的预测值是:

[0056] Predicted value (i=1) =3

[0057] Predicted value (i=2) =2

[0058] Predicted value (i=3) =1

[0059] 假设实际值是:

[0060] Actual value (i=1) =1

[0061] Actual value (i=2) =2

[0062] Actual value (i=3) =3

[0063] 于是公式变为:

$$[0064] \quad MAPE = \frac{100}{3} \left(\frac{|3-1|}{1} + \frac{|2-2|}{2} + \frac{|1-3|}{3} \right) \% = \frac{8 * 100}{9} \% = 88.89\%$$

[0065] 如果在208处达到提前停止度量,那么在220处,优化的特征集是212的特征测试集S.

[0066] 如果在没有得到优化的特征集的情况下达到最大迭代次数,那么功能在222处结束。在一个实施例中,然后可以增加阈值e,并且功能可以在步骤206处继续。

[0067] 在一个实施例中,执行多轮图2的功能以便产生多个优化的特征集。每个特征集可以被用作预测算法的输入以生成预测训练模型。然后可以聚合多个经训练的模型以生成需求预测,如下面结合图5详细公开的。图2的功能的输出是一个或多个优化的特征集。

[0068] 图4A和图4B示出了根据一个实施例在执行三轮之后使用图3的特征集示例的图2的功能的输出的示例。在图4A和图4B的示例中,商品是酸奶,地点是Baltimore, MD区域,并且提前停止度量是MAPE。在第1轮中,提前停止阈值为10%,并且最大迭代次数为15。如图所示,在第1轮中,在第4次迭代时达到提前停止度量(即,9.5),并且用于训练算法以生成第一模型的输出的优化的特征集是(1, 2, 3, 6, 10, 5, 9, 7)。在第2轮中,提前停止阈值为10%,并且最大迭代次数为12。如图所示,在第2轮中,在第6次迭代时达到提前停止度量(即,8.9),并且用于训练算法以生成第一模型的输出的优化的特征集是(1, 2, 3, 6, 10, 4, 8, 9)。在第3轮中,提前停止阈值为8%,并且最大迭代次数为15。如图所示,在第2轮中,在达到最大迭代次数15之后,尚未达到提前停止阈值,因此在第3轮中没有找到优化的特征集。虽然图4A和图4B的每个示例包括相同数量的可选特征,但是实际数量在不同的迭代内可以有所不同。

[0069] 如所公开的,本发明实施例的功能利用提前停止标准自动选择用于给定产品/地点的最佳特征集,该提前停止标准允许机器学习工具预测准确的需求。特征选择可以自动地允许零售商在较高级别设置数据特征,但针对每个产品/地点用不同的特征子集来训练/预测需求。自动化的参数调节有助于实现比猜测/尝试已知方法更准确的预测,因为实施例对于每个产品/地点探索比人类可以手动处理的更多的搜索空间。另外,实施例通过需要比任何已知解决方案更少的处理周期和存储器存储来改进计算机的性能。

[0070] 在上面公开的使用图2的功能生成一个或多个优化的特征集的实施例中,实施例

使用优化的特征集作为预测算法的输入来生成预测模型。图5是根据一个实施例的当使用多个经训练的模型确定聚合级别的促销效果时图1的促销效果模块16的功能的流程图。可以使用图2的功能生成多个经训练的模型。

[0071] 在502处,针对产品的特定类/类别针对所有商店的所有物品接收历史物品销售数据。

[0072] 历史业绩数据包括跨多个时间段(例如,多个周)的与多个促销组成部分相关联的销售数据。促销组成部分的示例包括但不限于价格折扣组成部分、电视广告组成部分、收音机广告组成部分、报纸广告组成部分、电子邮件广告组成部分、互联网广告组成部分和店内广告组成部分。

[0073] 汇集所有有效数据点以形成具有给定聚合级别的N个数据点的训练数据集D。聚合级别是高于汇集数据的SKU/商店/周的交叉点。聚合级别的示例是子类/商店。这个级别可用的数据由特定子类中的所有SKU确定。实施例中的聚合级别通常被选择得足够低以捕获商品的低级细节,但也足够高以使数据池足够丰富以用于对促销效果的健壮估计。

[0074] 例如,如果子类中有50个物品平均销售约一年(即,52周),并且链中有50个零售商店,那么:

[0075] $N = 50 * 52 * 50 = 130000$ 个数据点作为502的结果,训练数据集D被形成有N个数据点。在这个示例中,给定的聚合级别是子类/商店。

[0076] 在504处,数据集D被多次采样以形成多个不同的训练集D(i)。实施例通过从D均匀地并且利用替换随机采样来生成m个新的训练集D(i),每个训练集D(i)的尺寸为n'(例如,N的80%)。利用替换的采样被用于利用替换找到概率。具体而言,确定存在多个数据点、球、卡或其它对象的某个事件的概率,并且每当选择每个物品时替换该物品。例如,当利用替换采样两个时,两个样本值是独立的。因此,第一个被替换的不影响第二个被替换的。从数学上讲,这意味着两者之间的协方差为零。

[0077] 没有用于训练(即,不形成采样集的一部分)的数据点(N-n')作为验证/测试集T(i)被用于验证。例如,在一个实施例中,生成五个训练集。每个训练集具有 $(130,000) * (0.8) = 104000$ 个数据点,并且每个测试/验证集包括26000个剩余数据点。由于随机采样,每个训练集都不同。

[0078] 在506处,对于204处的每个训练集D(i),运行多个可能的不同机器算法中的一个机器算法以产生/训练模型。在一个实施例中,对于每个训练集D(i),使用以下机器学习算法之一来产生模型M(i):线性回归、支持向量机("SVM")和人工神经网络("ANN")。一般而言,机器学习算法可以从数据中学习并对数据做出预测。机器学习算法通过从输入观察结果的示例训练集构建模型来操作,以便将数据驱动的预测或决策表达为输出,而不是遵循严格的静态程序指令。

[0079] 一般而言,使用机器学习算法训练模型是描述如何基于输入特征集计算模型的输出的方式。例如,对于线性回归模型,预测可以建模如下:预测=基本需求*季节性*促销1*促销2*...促销效果10。对于不同的训练方法,输出将是不同的。例如:(1)对于线性回归,训练将产生针对季节性、促销效果1...促销效果10的估计;(2)对于SVM,训练将产生“支持向量”,“支持向量”是与某个权重相关联的输入数据点的集合;(3)对于ANN,训练输出将是针对每个节点的最终激活函数和对应的权重。

[0080] 在508处,验证每个模型并使用测试集确定误差。对于每个模型M(i),实施例应用测试集T(i)来预测结果并计算均方根误差RMSE(i)。例如,对于其中存在10个数据点x1,...x10的测试数据集i,实施例基于经训练的模型预测这10个点的输出。如果输出为P1,...P10,那么RMSE计算如下:

$$[0081] rmse = \sqrt{\left(\sum_{n=1}^{10} (xi - pi)^2 \right) / 10}$$

[0082] 在510处,针对每个模型,计算模型权重。在一个实施例中,对于每个模型M(i),其权重w(i)确定如下:

$$[0083] w(i) = \frac{1}{1 + RMSE(i)}$$

[0084] 然后,实施例确定w(i)的总和如下:

[0085] S=Sum(w(i))最后,实施例将用于每个w(i)的权重归一化如下:

$$[0086] w'(i) = \frac{w(i)}{S}$$

[0087] 在512处,输出模型组合。为了预测未来需求,对于每个数据点x,将M(i)迭代地应用于输入以产生最终结果y,如下所示:

$$[0088] y = \text{sum}(f(M(i), x) * w'(i))$$

[0089] 其中y是预测的需求,并且f是创建预测的函数,与模型对应。例如,考虑三个模型。对于给定的点x,模型产生预测,并在下表中给出权重:

模型	预测	权重
模型1	4	0.5
模型2	4.5	0.3
模型3	3.9	0.2

[0090] 最终需求预测计算为:

$$[0091] y = 4 * 0.5 + 4.5 * 0.3 + 3.9 * 0.2 = 4.13.$$

[0092] 在514处,最终需求预测被用于制造生产、装运物流和库存控制。在一个实施例中,最终需求预测被发送到其它专用计算机,诸如库存控制系统、制造系统、装运和物流系统以及销售辅助系统。在一个实施例中最终需求预测是以单独的数据位的形式,这些单独的数据位已经从需求预测被转换并且被存储和发送到其它专用计算机系统,在那里它们由这些系统存储和使用。因此,附加物品可以被制造、存储、装运等,并且可以最佳地给物品定价。

[0093] 如上所述,在一个实施例中,图2的功能的输出可以被用作图5的功能的输入以生成需求预测。例如,在图5的一个实施例中,考虑三种算法(例如,线性回归、SVM、ANN),并假设图2生成三个优化的特征集。如以上所讨论的,在图5中,使用优化的预测算法训练算法,并且结果将是9(三乘三)个训练模型。这九个模型根据它们的权重被组合以创建最终预测。

[0094] 图6和图7A-图7B示出了根据一个实施例的在聚合级别确定促销效果的示例。在图6和图7A-图7B的示例中,假设对于零售商“A”,在亚特兰大,GA区域中有2年的酸奶类别的历史。假设亚特兰大地区有20家零售商店,并且每家商店包括大约100种不同的酸奶UPC/SKU。

[0095] 根据上面的502,在这个简化示例中,对于形成训练数据集D的物品/商店/周销售

聚合级别,总共有 $20*100*104=2080000$ 个数据点,其中20是零售商店的数量,100是SKU的数量,104是两年历史销售期的周数。

[0097] 还假设零售商提供10种不同类型的促销。促销被称为“促销1”、“促销2”、“促销3”…“促销10”。在这个示例中,需求模型如下:

[0098] 销售=(基本需求)*(季节性)*(促销1效果)*(促销2效果)*…(促销10效果)

[0099] 可以使用已知方法在物品/商店级别计算基本需求,诸如移动平均、简单指数平滑等。可以使用已知方法在类别/区域级别计算季节性,诸如加法和乘法冬季指数平滑模型。挑战在于估计十个促销效果(即,估计每个促销对在该促销生效的每个销售时段期间的销售预测的影响)。在这个示例中,因为仅存在两年的销售历史,因此使用已知的估计方法难以估计物品/商店级别的促销效果。

[0100] 图6示出了根据一个实施例的使用数据点的六轮模型估计。对于每轮,使用线性回归确定每个促销1-10的促销效果。每轮使用相同类型的算法。例如,每轮可以使用线性回归、SVM、神经网络等。在每轮之后,生成描述所使用的训练集的参数集。该参数集被称为“模型”。因此,在图6的示例中,基于六轮获得六个模型。

[0101] 在A轮(行601)中,所有可用数据点被用于与本发明的确定进行比较。对于第1-5轮(行602-606),采样数据被用于进行估计(根据图5的504),并且剩余的测试数据被用于测试/验证模型(根据图5的508)。在一个实施例中,采样数据是数据点的80%,并且测试数据是数据的剩余20%。

[0102] 在图6所示的示例中,线性回归被用于训练。由于每轮使用不同的训练数据集,因此每轮的估计效果将不同。促销效果是特定于产品/地点的,但不是特定于时间段的。每轮使用相同的模型和方法。

[0103] 对于训练/测试中的每一个,基于测试数据计算列611中的RMSE(根据图5的508),并且如上面所公开的,在列612、613中计算每轮的对应权重 $w(i)$ 和归一化权重 $w'(i)$ 。

[0104] 图7A和图7B示出了使用本发明实施例的预测与实际销售的比较。

[0105] 对于13周销售时段期间的每一周,以及对于给定商店/SKU(例如,在特定零售商店销售的特定类型的酸奶),行701提供基线需求,行702提供季节性,以及行702-712为每次促销提供该促销是否在对应的一周期间有效的指示(如“X”所示)。行713指示对应时间段期间的实际销售额。

[0106] 对于促销效果的预测,行714指示来自A轮的每周销售额的预测,其中使用利用所有可用数据的已知方法来使用所有数据点。行715-719指示(使用本发明的实施例)针对每个时间段使用轮1-5中的每一轮的预测/估计,并且行720是来自轮1-5的平均预测。列721使用RMSE来示出使用本发明实施例的方法实现最佳性能(即,根据本发明实施例的行720具有比使用利用整个数据集而不采样的已知方法的行720更小的RMSE)。

[0107] 代替利用引入偏差的削减的数据来估计促销影响,实施例利用整个数据集的丰富性,但是使用采样来降低必要的处理能力。实施例是完全自动化的并且可以被调整以平衡性能和准确性。另外,实施例提供了对促销物品的预测准确性的改进。预测是供应链最重要的驱动因素之一。如果预测不准确,那么分配和补货表现不佳,从而导致公司财务损失。

[0108] 一般而言,购物者会特别关注促销物品。如果促销计划得不当,而且预测太高,那么物品将保持未售出状态,并且它们需要以折扣价出售,或者浪费增加。在这两种情况下,

盈利能力下降。如果预测低,那么不满足需求,零售商会遇到销售额损失和客户满意度低的情况。两者都对收入产生负面影响。实施例通过平衡促销/销售预测的准确性和可靠性来避免销售额损失或不必要的降价。

[0109] 如所公开的,实施例的一个目标是为物品选择相关特征以最终最大化预测准确性。良好的预测一般不会带来任何借贷(credit)。物品总是可以获得并且它们以全价而不是折扣价出售。库存水平不应当太高,因此零售商的资金不会被库存占用。零售商和供应商应当能够可靠地计划劳动力和生产能力。

[0110] 但是,如果预测错误(即,不准确),那么情况会发生戏剧性变化。这些影响可能对许多业务领域产生负面影响。例如,如果预测太低,那么比所需的少的产品到达零售商并且销售一空。缺货情况给零售商带来收入损失并且客户满意度降低的影响。低的预测还会影响供应商,供应商必须缩减产量,并审查他们对当前劳动力的需求。

[0111] 如果预测过高,那么也存在负面影响。零售商将订购超过他们可以卖出的产品。如果产品易腐烂,那么它们可能会变质,从而增加浪费。即使它们不易腐烂,零售商也可能以折扣价出售额外的物品,这会对收入产生负面影响。否则,零售商可能将商品退还给供应商。这会影响供应商,因为他们拥有没有需求的额外的产品。另外,制造商可能浪费时间和金钱来生产错误的产品,这会对供应商的收入产生负面影响。

[0112] 图8示出了根据一个实施例的包括如本文所公开的需求预测的集成的制造、库存和物流系统800。如图8中所示,系统800可以包括产品预测系统870,该产品预测系统870预测未来的产品需求并且在一些情况下预测和/或考虑对数十万种产品的未来需求,或者在一些应用中预测在一个或多个零售商店801-804处的数千万或更多的产品。预测系统870通过云网络850或其它类型的通信网络与一个或多个库存系统820和一个或多个制造系统880进行通信。

[0113] 预测系统870通过实现结合上面的图2和图5公开的功能来生成需求预测。库存系统820存储库存并提供运输物流以使用卡车810-813或一些其它运输机制将物品递送到商店801-804。在一个实施例中,库存系统820实现企业资源计划(“ERP”)专用计算机系统或专用库存控制系统,企业资源计划(“ERP”)专用计算机系统或专用库存控制系统使用来自预测系统810的输入来确定库存水平以及将物品递送到商店801-804的数量和时间。

[0114] 制造系统880制造要发送到库存系统820的物品,并提供运输物流以使用卡车881或一些其它运输机制将物品递送到库存系统820。在一个实施例中,制造系统880实现ERP专用计算机系统或专用制造系统,ERP专用计算机系统或专用制造系统使用来自预测系统870的输入来确定要制造的物品的数量、用于制造的资源的库存,以及将物品递送到库存系统820的数量和定时。

[0115] 预测系统870可以利用来自库存系统820、销售跟踪系统(未示出)和/或数据库的信息来预测对产品的需求。在预测需求时,预测系统870试图预测由事件、天气、社会需求、经济因素和其它因素导致的一种或多种产品的非典型需求。可以跟踪数十、数百到数千个不同的变量,这些变量可以对一个或多个产品的需求产生影响。这些变量的变化可以导致非典型需求。例如,可以跟踪预测的天气的变化,并且可以使用与预测的天气相关联的一个或多个变量来确定这种天气变化是否可能对需求产生影响,并且是否可以进一步预测需求的变化。

[0116] 一般而言,图8的元素执行库存的销售、制造或消费。由于随机性和影响销售的外部因素,用于直接消费者销售的零售地点/商店801-804表现出最不稳定的库存模式。但是,消耗库存的制造设施和场所(诸如产品集成商、互联网装运商等在本地设施中使用的产品)也受益于本文所公开的需求预测。如所公开的,每个零售地点801-804将销售数据和历史预测数据发送到预测系统870。销售数据包括先前销售周期(即,几周)(通常是库存周期的4-7周)中每个销售时段(通常是几天)的每个物品或SKU/UPC的库存消耗统计信息。

[0117] 预测系统870将销售数据存储在储存库872中,并采用销售数据来生成补充库存的订单。订单包括物品的集合和用于维持商店801-804的库存水平的每种物品的数量。

[0118] 许多零售订购计划依赖于销售时段和销售周期的一周中的天。在一种配置中,在具有库存统计信息的库存管理环境中,其中库存统计信息特定于一周中的每一天,库存系统820通过针对一周中的每一天搜集来自先前销售额的库存水平统计信息来确定目标库存水平。实施例基于库存水平统计信息来计算一周中的每一天的库存水平,使得安全存货适应一周中不同天之间的库存变化。对于多种物品中的每一种物品,实施例呈现指示目标库存水平的存货水平,包括一周中每天的安全存货。实施例基于订货交付时间计算订购数量,使得订购的数量达到关于一周中的所确定的一天的所呈现的存货水平。识别实际存货水平包括从历史数据中识别前几周的一周中的一天的存货水平,从而随着时间关注一周中的同一天,而不是一周中所有天的平均值。

[0119] 在特定配置中,所公开的实施例可以与专用和/或特别大批量的零售环境结合使用。在大型物流和配送操作中,尽可能装满卡车,并且如果需要将物品推迟到后续的行程,那么选择将最不可能中断销售活动的那些物品是有益的。因而,实施例可与POS系统100结合操作,以识别倾向于比其它物品更快地售出和补货的高周转率或高流通率物品。物品上的UPC条形码符号或射频标识(“RFID”)包括单独地或与数据库查找相结合将物品指定为适合如本文所定义的安全存货处理的高周转率物品的字段、标志或值。

[0120] 通过为库存数据库中表示的多个物品中的每一个识别产品标识符的字段和表示该物品的安全存货的字段,并基于指示由销售量导致的产品补充需求增加的产品周转率为每个产品标识符确定产品细分字段,可以提供高周转率物品。所公开的实施例基于周转率字段来确定是否计算安全存货,即,在给定产品吞吐量的情况下,根据安全存货进行重新供应的开销和负担是否值得。

[0121] 在其它实施例中,供应物流可以调用每天高于一卡车的递送频率,因此触发具有更高粒度的重新供应窗口。在这种情况下,安全存货可以比单独的一天更具体,诸如周一上午和周一下午,或者指定一周中的特定一天内的多个递送或时间窗口,诸如上午7:00、上午11:00和下午4:00。

[0122] 包括所生成的需求预测的实施例可以用于根据所运输物品的需求和利润率来实现供应物流并指定递送工具(即,卡车)和清单(即,包含的物品)。高周转率物品可能被视为在特定递送时具有优先级空间,但还可以基于所包括的物品的利润率或增高标价(markup)而被选择,并且选择包括具有最大创收潜力的物品。

[0123] 在使用本文公开的需求预测并且具有多个运输车辆的这种产品库存装运环境中,每个车辆(例如,卡车)被配置用于接收固定有效载荷的物品以便递送到销售地点以供库存补充。通过对于包括第一物品和第二物品在内的多个物品中的每种物品计算安全存货并基

于计算出的第一物品和第二物品的安全存货来确定要装入递送车辆的第一物品和第二物品中的每种物品的量,实施例可以为装载递送车辆提供指导。如果在交付车辆中没有足够的空间用于第一物品和第二物品的确定的量,那么实施例基于安全存货重新计算卡车装载量,这意味着某些物品需要被省略并被推迟到后续的递送。

[0124] 本文具体说明和/或描述了若干实施例。但是,应该认识到的是,在不脱离本发明的精神和预期范围的情况下,所公开的实施例的修改和变化由上述教导覆盖并且在所附权利要求的范围内。

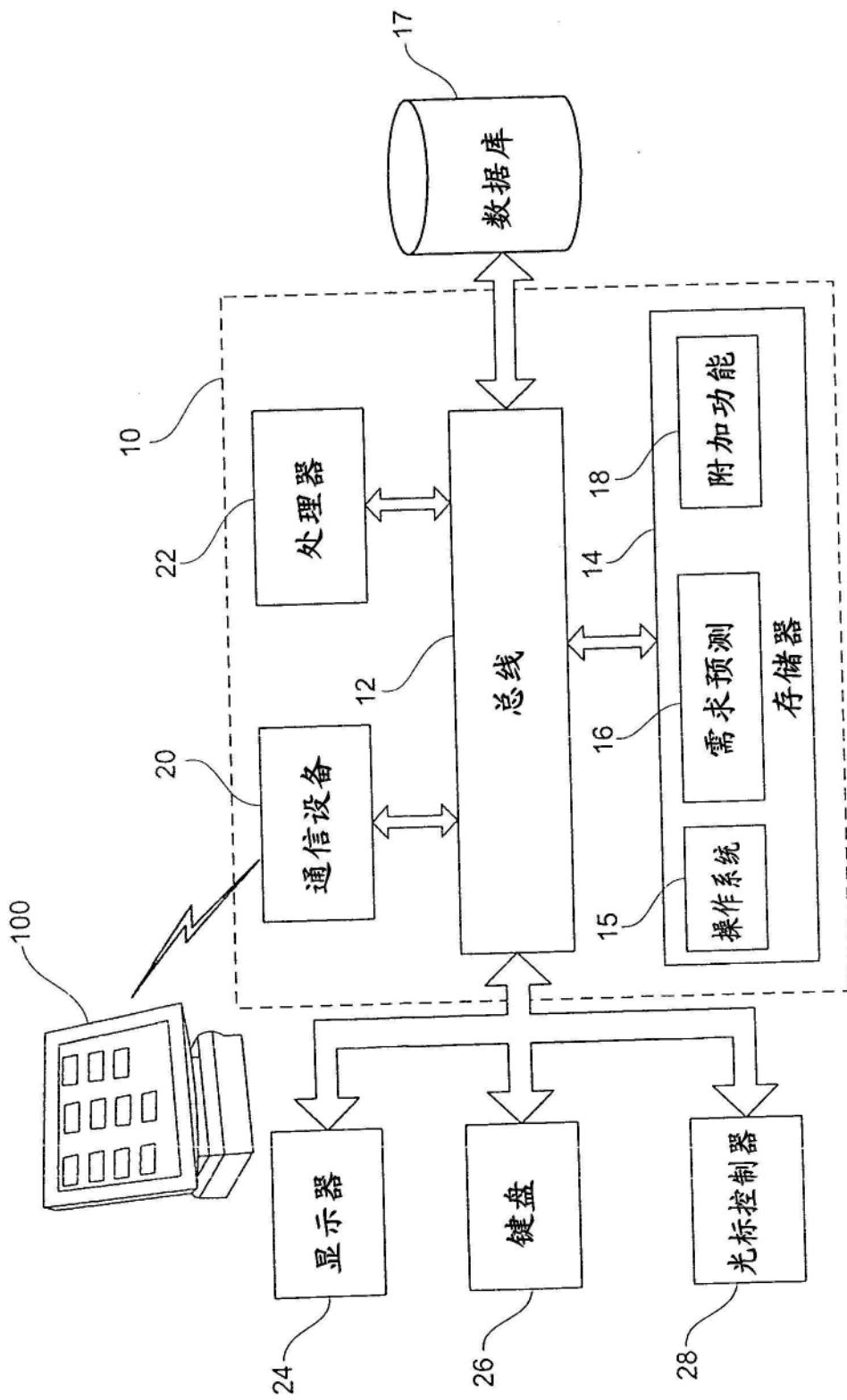


图1

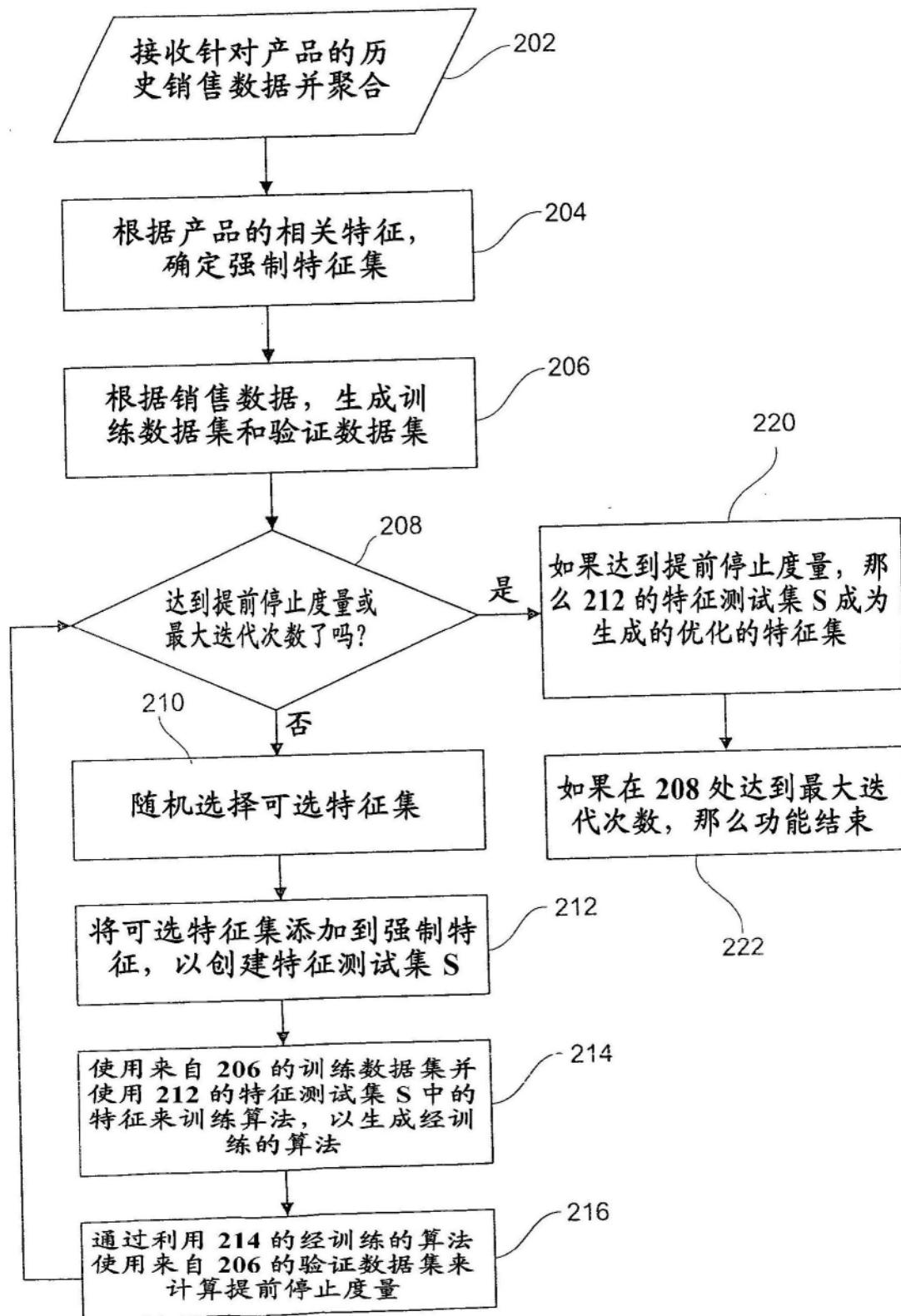


图2

特征#	1	2	3	4	5	6	7	8	9	10
名称	季节性	价格	品牌	包装尺寸	物品尺寸	风味	颜色	店内显示	端架	买一送一

图3

A

商品	地点	提前停止	轮#	提前停 止阈值	最大迭代	迭代#	所选择的特征#	提前停止量		停止
								度	是否	
酸奶 Balti- more, MD	1	10	15		1	1	1,2,3,6,10,4,8	20	否	停止
						2	1,2,3,6,10,4,9	15	否	
						3	1,2,3,6,10,5,9	11	否	
						4	1,2,3,6,10,5,7	9,5	是	
						1	1,2,3,6,10,4,7	11	否	
	2	10	12		2	2	1,2,3,6,10,4,8	15	否	停止
						3	1,2,3,6,10,5,7,8,9	18	否	
						4	1,2,3,6,10,4,5,7,9	16	否	
						5	1,2,3,6,10,4,5,7	22	否	
						6	1,2,3,6,10,4,8,9	8,9	是	
8	3	8	15		1	1	1,2,3,6,10,4,7	20	否	停止
						2	1,2,3,6,10,4,8	32	否	
						否	
						14	1,2,3,6,10,4,5,7,9	16	否	
						15	1,2,3,6,10,4,5,7	22	是	

图4A

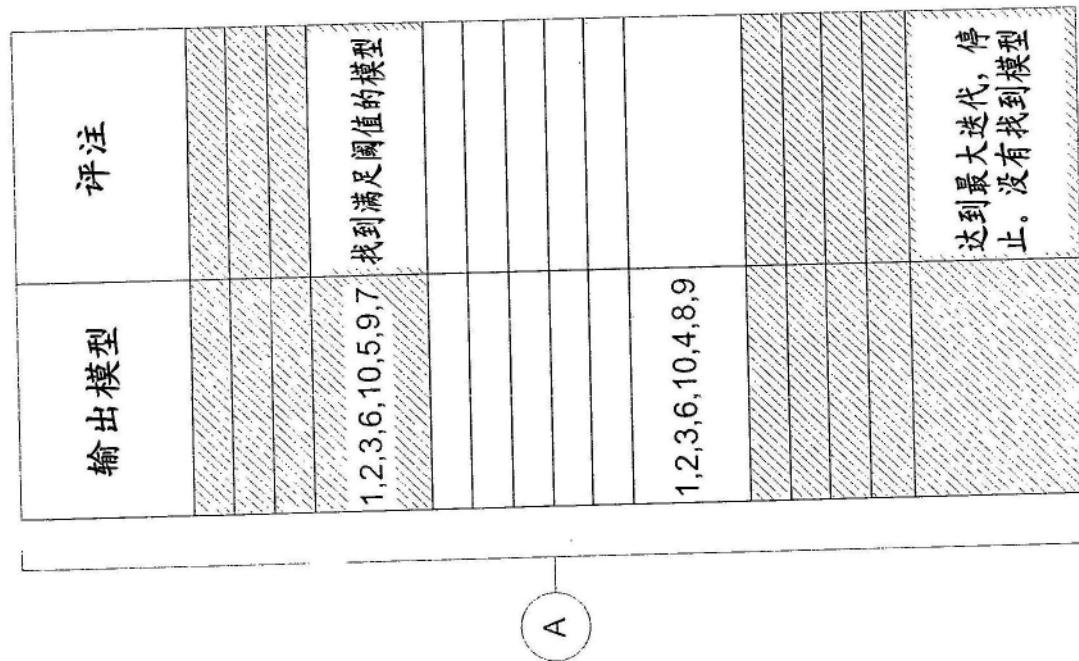


图4B

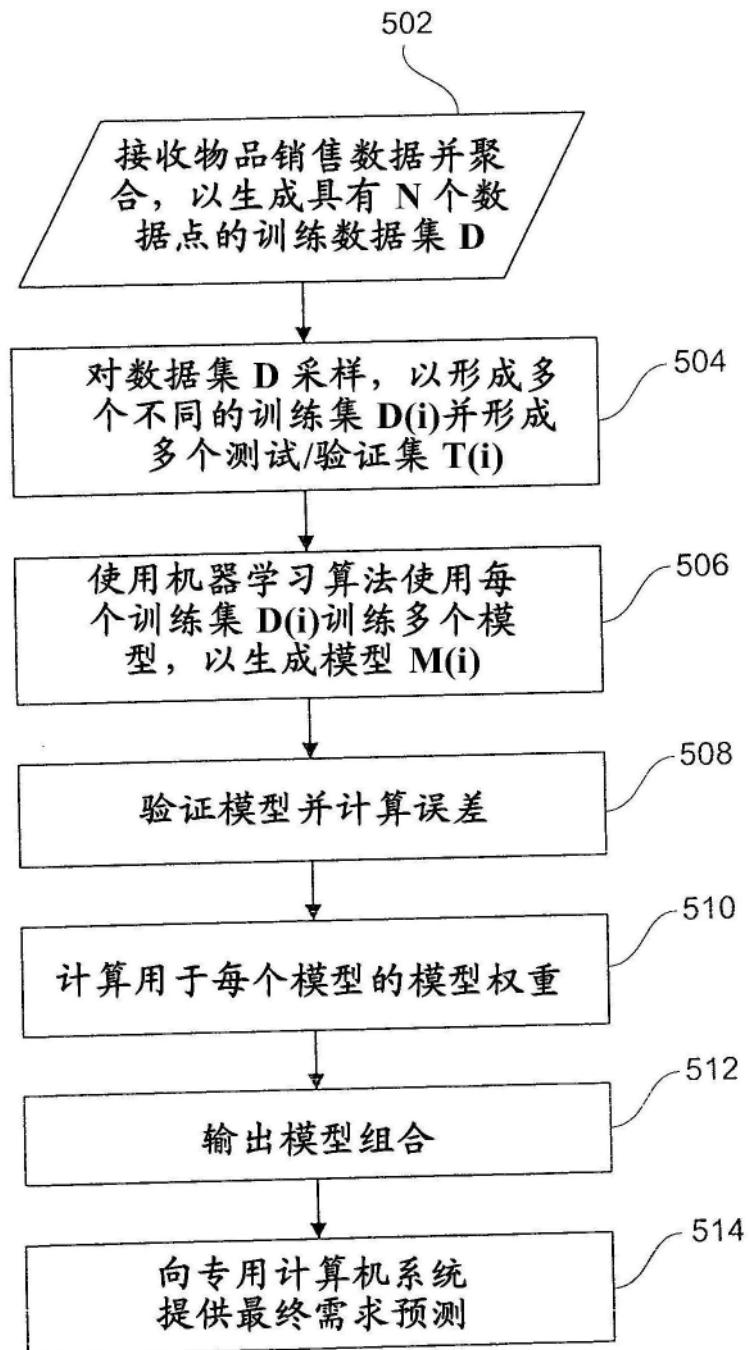


图5

类型	轮	促销效果估计										RMSE (i)	w(i)	w'(i)	
		促销 ₁	促销 ₂	促销 ₃	促销 ₄	促销 ₅	促销 ₆	促销 ₇	促销 ₈	促销 ₉	促销 ₁₀				
包括所有数据点	A	1.09	1.21	1.56	3.2	2.12	1.12	2.12	3.15	1.6	1.78	NA	NA	NA	
	1	1.16	1.14	1.57	3.14	2.03	1.14	2.06	3.19	1.54	1.79	9.75	0.093023	0.144835	
	2	1.12	1.25	1.58	3.11	2.13	1.02	2.15	3.17	1.63	1.70	4.77	0.17331	0.269841	
	3	1.06	1.13	1.49	3.14	2.06	1.06	2.13	3.15	1.60	1.81	6.39	0.135318	0.210687	
利用替换进行采样， 80%用于训练并且 20%用于测试		4	1.06	1.23	1.48	3.10	2.04	1.07	2.15	3.08	1.54	1.80	5.94	0.144092	0.224349
		5	1.04	1.14	1.56	3.25	2.02	1.04	2.02	3.19	1.51	1.72	9.36	0.096525	0.150288

图6

预测估计

在一个给定的商店 /sku		周 1	周 2	周 3	周 4	周 5
701	基线	200	200	200	200	200
702	季节性	1.1	1.2	1.5	1.6	1.7
703	促销 1	x				
704	促销 2		x			
705	促销 3			x		
706	促销 4					
707	促销 5					
708	促销 6					
709	促销 7					
710	促销 8			x		
711	促销 9				x	
712	促销 10					
713	实际销售	242	286	465	480	340
714	预测	轮 A	239.8	290.4	480	499.2
715		轮 1	255.241	272.5578	461.1735	501.3791
716		轮 2	246.9645	301.0822	488.2887	504.2234
717		轮 3	233.5692	271.8177	481.0587	478.1017
718		轮 4	233.1904	295.6165	461.8454	474.4456
719		轮 5	228.9424	273.5752	452.5576	499.5444
720	最终	239.5422	285.4249	471.5356	490.9239	340

图 7A

721

8

图7B

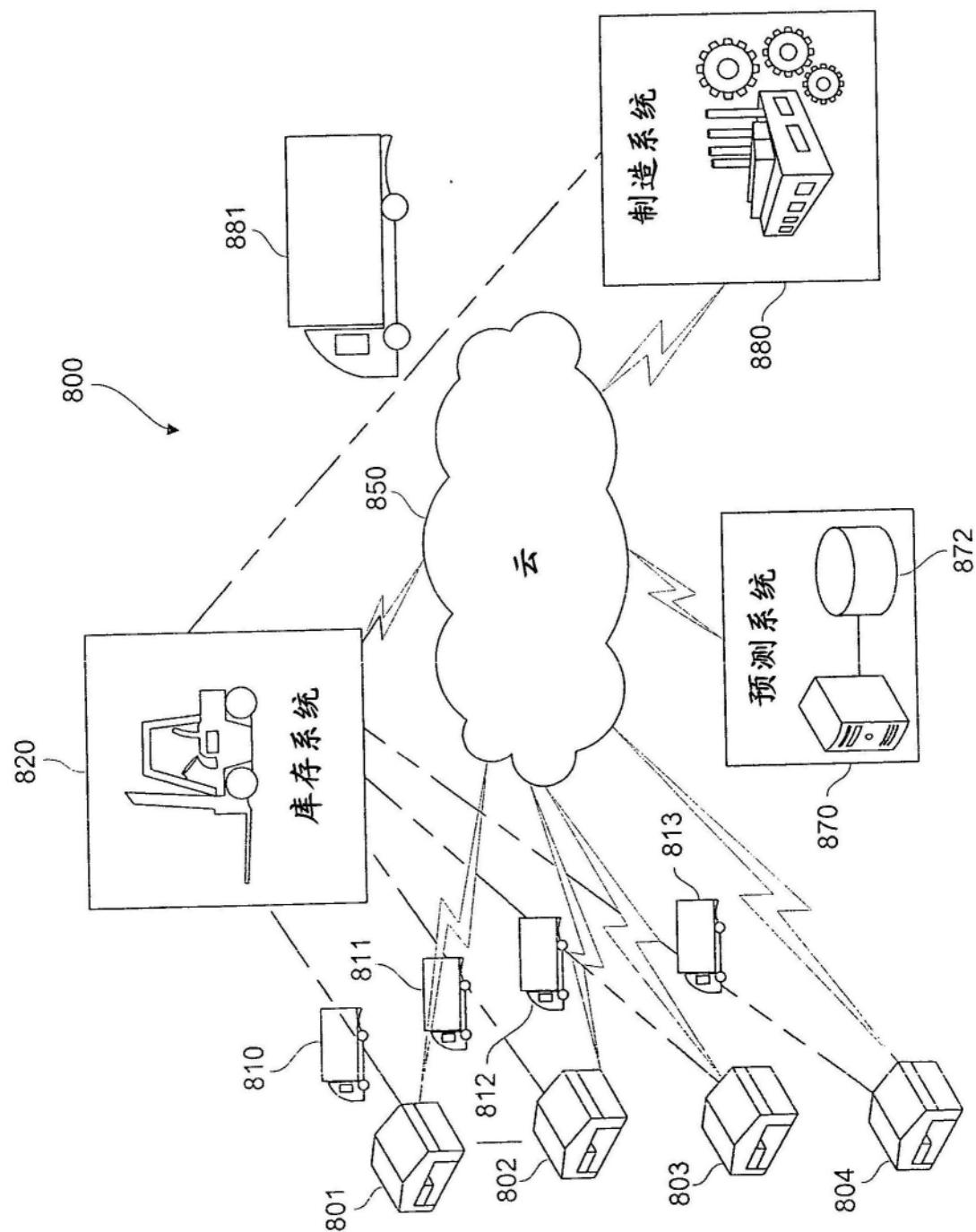


图8