Statistics and Application 统计学与应用, 2023, 12(6), 1683-1695
Published Online December 2023 in Hans. https://www.hanspub.org/journal/sa
https://doi.org/10.12677/sa.2023.126172

沃尔玛百货公司零售数据预测分析

肖　虎

云南财经大学统计与数学学院，云南　昆明

收稿日期：2023年9月25日；录用日期：2023年12月20日；发布日期：2023年12月28日

摘　要

本文从*Kaggle*网站上获取到沃尔玛在不同地区45家门店2010年、2011年、2012年的销售信息。首先，对数据进行纵向和横向统计分析，获取到门店的销售信息，并发现在2010-02-12、2010-09-10、2010-11-26、2010-12-31、2011-02-11、2011-09-09、2011-11-25、2011-12-30、2012-02-10、2012-09-07这些节假日的销售额会高于非节假日，对沃尔玛销售有益。接着，对解释变量进行相关性分析，发现燃油价的变动对居民消费指数影响较大，从而间接影响失业率，燃油价格升高在一定程度上会促进居民消费指数的上升，降低失业率，而燃油价格下降具有相反的变化趋势。然后，对数据进行标准化处理，使用多元线性回归模型在测试集上预测了销售额数据，并展示了部分预测结果及相关信息。最后，本文搭建了一个两层的深度神经网络模型并使用沃尔玛数据进行训练，该模型可用于预测其他门店的销售额。

关键词

相关性分析，标准化，多元线性回归模型

Wal-Mart Retail Sales Data Forecast Analysis

Hu Xiao

School of Statistics and Mathematics, Yunnan University of Finance and Economics, Kunming Yunnan

Received: Sep. 25th, 2023; accepted: Dec. 20th, 2023; published: Dec. 28th, 2023

Abstract

This article obtained the 2010, 2011 and 2012 sales information of 45 Wal-Mart stores in different regions from the *Kaggle* website. First, the data is analyzed vertically and horizontally to obtain the sales information of stores. Found 2010-02-12, 2010-09-10, 2010-11-26, 2010-12-31, 2011-02-11, 2011-09-09, 2011-11-25, 2011-12-30, 2012-02-10, 2012-09-07 these holidays are sold would be higher than a non-holiday, which would benefit Wal-Mart sales. Then, the correlation analysis of explanatory variables found that the change of fuel oil price has a great impact on the consumer index, thus indirectly affecting the unemployment rate. The increase of fuel price will promote the increase of consumer index and reduce the unemployment rate to some extent, while the decrease of fuel price

文章引用: 肖虎. 沃尔玛百货公司零售数据预测分析[J]. 统计学与应用, 2023, 12(6): 1683-1695.
DOI: 10.12677/sa.2023.126172

has the opposite trend. Then, the data were standardized, the multiple linear regression model was used to predict the sales data on the test set, and some of the predicted results and correlation information were displayed. Finally, in this paper, a two-layer deep neural network model is constructed and trained using Wal-Mart data. This model can be used to predict sales of other stores.

Keywords

Correlation Analysis, Standardization, Multiple Linear Regression Model

Copyright © 2023 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

1.1. 研究背景及意义

随着社会发展，人们生活水平提高，人们对于生活质量的要求也越来越高。购物便成为了人们满足这一需求的重要途径。沃尔玛百货有限公司是一家美国的世界性连锁企业，连续七年位居美国《财富》杂志首位，同时，也是世界上最大的私人雇主，在世界范围内拥有超过两百万的员工，是世界上最大的零售商。截至 2021 年，沃尔玛在全球 27 个国家开设了超过 10,000 家商场，而且随着该公司向新兴市场扩张，这一数字很可能会增加。这些门店的净销售额超过 5000 亿美元，而沃尔玛的毛利率一直稳定在 25%左右，这意味着沃尔玛经过几十年的发展，仍然具有较强的竞争力，在全球市场中仍然具有较大影响力。

在文献[1]-[15]中展示了沃尔玛在世界范围内能够取得这样的成功的关键，在很大程度上取决于尖端科技和物流系统进行巧妙搭配，使用计算机进行管理和全球领先的卫星定位系统控制公司物流，提高配送效率。文献[15] [16] [17]给出了沃尔玛一直采用最先进的系统进行运输安排，通过电脑系统和配送中心相互配合，为沃尔玛节省了相当多的成本，以速度和质量赢得用户的满意和忠诚。

这样大的一个百货公司要为各个门店合理配送商品并不容易，如果给销量差的门店配送较多的商品，会导致货物囤积，产生浪费；而给销量好的门店配送较少的商品，就会影响门店收益，甚至给消费者带来较差的消费体验，每天影响销量的因素有很多，如果能够预测销售额和准确的需求，将会很好地为沃尔玛节约成本和提高收益，为更多消费者带来更好的用户体验。

1.2. 国内外研究现状

沃尔玛是最早开始投资和部署大数据应用的零售商之一，在全球许多地区设立了沃尔玛大数据实验室，从中挖掘出消费者的消费习惯，商品之间的相关性，建立起消费者与商品之间便捷的联系，这为沃尔玛带来了巨大的利润，同时，也为消费者提供了更好的购物体验，参见文献[18] [19]。在 20 世纪 90 年代，沃尔玛超市管理人员对销售数据进行数据分析时就发现“啤酒”与“尿布”这两个毫无关系的商品会经常出现在同一个购物篮中，经过调查发现，这种现象发生在年轻父亲身上，由此，沃尔玛的超市管理人员们将“尿片”和“啤酒”经常性放在一起出售，大大提高了两者的销量，参见文献[20] [21]。对销售商品进行数据分析，将许多孤立的数据集成起来，能够实现智慧营销，帮助管理者更好地制定决策。在许多情况下，简单的数据分析只能提取到少量有价值的信息，更多有价值的信息需要使用数据挖掘技术，通过算法获取。实验研究表明，数据挖掘技术能够帮助沃尔玛挖掘更多客户信息，辅助日常管理，提高市场竞争力。

本文将使用数据分析方法对沃尔玛部分地区数据进行分析，这有助于沃尔玛百货公司更好地执行商

业决策，对沃尔玛的发展具有十分重要的意义。

1.3. 研究思路

本研究主要分为三个部分：

第一部分：对数据进行预处理。本文所采用的原始数据未经过处理，数据中不乏存在缺失值、异常值等情况，需要在使用前通过 Python 进行填充和修改之后才能进行分析和预测。

第二部分：数据分析过程。经过预处理过后的数据就可以直接使用了，通过给定的数据分析不同指标对销售额的影响，分析不同年月季度和地区的销售额，节假日对销售的正负面影响。

第三部分：建模过程。本文将采用两种模型对销售额进行预测。为了保证数据研究的科学性和合理性，将数据分为训练集和测试集。使用训练数据建立模型，测试数据验证模型。通过实际结果和模型预测结果进行比较，验证模型的性能。

2. 模型简介

2.1. 多元线性回归模型

多元线性回归模型由多个解释变量预测或估计因变量，其非随机结构表达式如下：

$$Y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i} + \cdots + \beta_k x_{ki} + \mu_i i = 1,2,3,\cdots,n$$

其中，$x_{1i}, x_{2i}, x_{3i}, \cdots, x_{ki}$ 为 k 个解释变量，$\beta_j\left(j = 0,1,2,\cdots,k\right)$ 为线性回归模型的回归系数，其 $\mu_i \sim N(0,\delta^2)$。在正太分布假定下，如果 $X'X$ 是可逆的，则参数 β 的最小二乘估计为：

$$\hat{\beta} = \left(X'X\right)^{-1} X'y$$

从而可以得到因变量 y 的估计值：

$$\hat{y} = X\hat{\beta}$$

残差 $e = y - X\hat{\beta}$，则随机误差方差 δ^2 的最小二乘估计为：

$$\hat{\delta} = \frac{e'e}{n-p-1}$$

这样就可以求解出多元线性回归模型的参数。

2.2. 深度神经网络模型

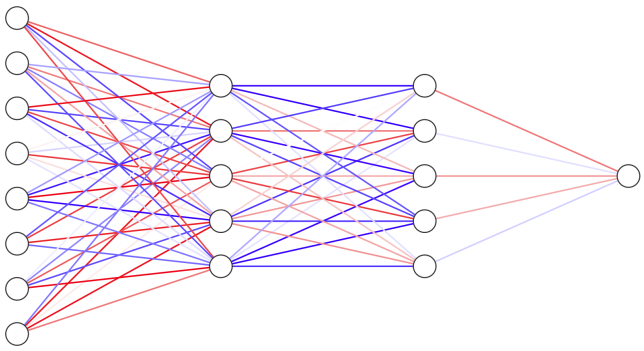

Figure 1. Neural network schema
图 1. 神经网络图示

深度神经网络(DNN)是一种多层无监督神经网络，它将上一层的输出特征作为下一层的输入，经过逐

层映射，最终将原始数据特征映射到另一个特征空间，经过这样的映射之后使得特征具有更好的表达。本文将构建一个具有两个隐藏层，单个输出神经元的深度神经网络用于销售额预测，其结构如图 1 所示。

网络中每一层到下一层之间的特征映射，通过如下线性关系得到中间结果：

$$z_j = \sum_{i=1}^{n} w_{ij} x_i + b$$

其中，w_{ij} 是下一层中第 j 个中间特征的第 i 个线性关系系数，b 为偏置项。

然后再接一个神经元激活函数：

$$f(x) = \begin{cases} 0, x < 0 \\ x, x \geq 0 \end{cases}$$

做非线性变换，就可以得到下一层神经元的输入。深度神经网络经过多个这样的特征映射可以实现对高度复杂函数的拟合，从而达到预测的目的。

为了提高深度神经网络的参数更新的效率，本文将采用 Adam 优化算法进行参数更新。Adam 优化算法结合了 AdaGrad 和 RMSProp 算法的最优性能，在深度神经网络训练过程中具有非常好的表现。

3. 数据分析

3.1. 数据来源及解释

本文所使用的沃尔玛商店数据来自于 *Kaggle* 竞赛网站，该数据本就是用于零售数据分析的，是沃尔玛在不同地区的门店的历史销售数据汇总而来，拥有 6000 多条数据，数据中包含 45 家门店的相关信息，从 2010 年到 2012 年中部分天数的销售额情况、节假日信息、对应华氏温度、燃料价格、居民消费指数 CPI 以及失业指数等。

3.2. 销售数据分析

在大多数情况下获取的数据都是不完整的，需要人为将获取到的数据的空值部分做预处理，本文所使用的沃尔玛数据也不例外，在此使用均值填充的方式来填补空值。使用 Python 经过预处理后，获取到数据的一些描述性统计信息，如下表 1 所示。

Table 1. Descriptive statistical analysis
表 1. 描述性统计分析

	Store	Weekly_Sales	Holiday_Flag	Temperature	Fule_Price	CPI	Unemployment
Count	6435	6435	6435	6435	6435	6435	6435
mean		1.046965e+06		60.66	3.35	171.57	7.99
std		5.643666e+05		18.44	0.45	39.35	1.87
min		2.099862e+05	0	−2.06	2.47	126.06	3.87
25%		5.533501e+05		47.46	2.93	131.73	6.89
50%		9.607460e+05		62.67	3.44	182.61	7.87
75%		1.420159e+06		74.94	3.73	212.74	8.62
max		3.818686e+06	1	100.14	4.46	227.23	14.31

从给出的数据描述来看，所使用的数据具有八个属性，6435 条记录，提供了 45 个门店的销售额、节假日、温度、油价、居民消费指数、失业率等相关信息。

在所提供的数据中，包含 2010 年、2011 年以及 2012 年的销售信息，下面分析 45 家门店三年的销售额情况，绘制出如下图 2 的柱状图。

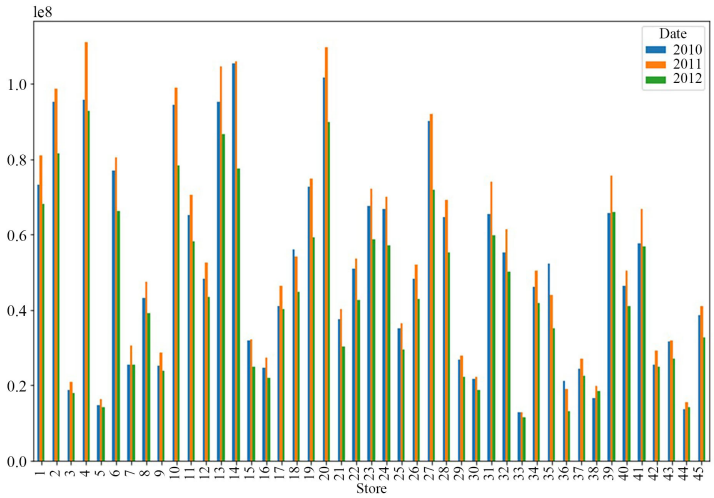

Figure 2. Bar chart of sales statistics of each store
图 2. 各门店销售统计数据柱状图

所绘制的柱状图反映出在所提供的 45 家门店中普遍 2011 年的销量最好，2010 年次之，2012 年销量最差。为了能看出这三年 45 家门店销售额的分布情况，将三年的销售额堆叠起来得到如下图 3 的柱状图。

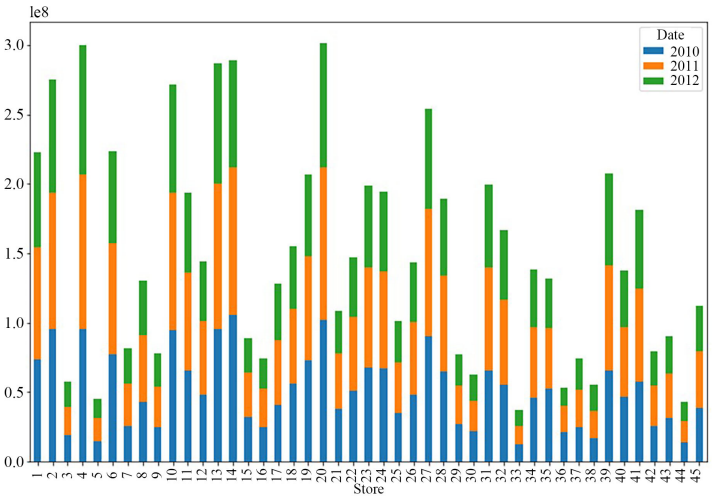

Figure 3. Stacked bar chart of sales statistics of each store
图 3. 各门店销售统计数据堆叠柱状图

从图中可以看出，第 20 家门店的销售额略高于其他门店，与第 4 家门店的差距较小，第 33 家门店的销售额最少。下面再分析 45 家门店三年销售额的标准差，绘制出如下图 4 的折线图。

从图中可以看出，第 14 家门店的标准差最大，第 33 家门店最小，说明第 14 家门店的销售额变化较大，第 33 家门店由于本身销售额较小，因此销售额变化不太大。

前面纵向比较 45 个门店 2010 年、2011 年及 2012 年的销售额情况，下面横向比较 2010 年、2011 年及 2012 年不同月份的销售额情况，绘制出三年各个月 45 家门店销售的折线图，如图 5、图 6、图 7 所示。

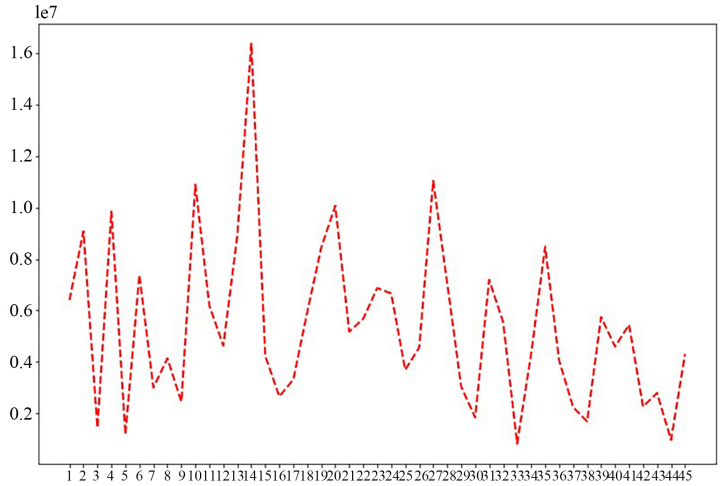

Figure 4. Line chart of sales statistics of each store
图 4. 各门店销售统计数据折线图

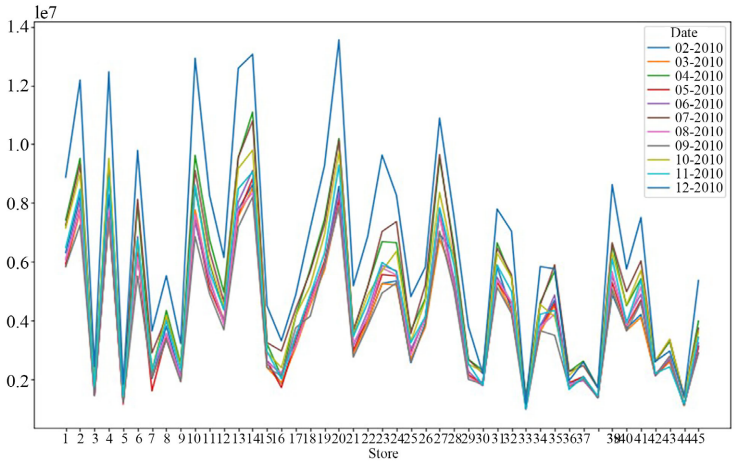

Figure 5. Line chart of monthly sales data of each store in the first year
图 5. 各家门店第一年各月销售数据折线图

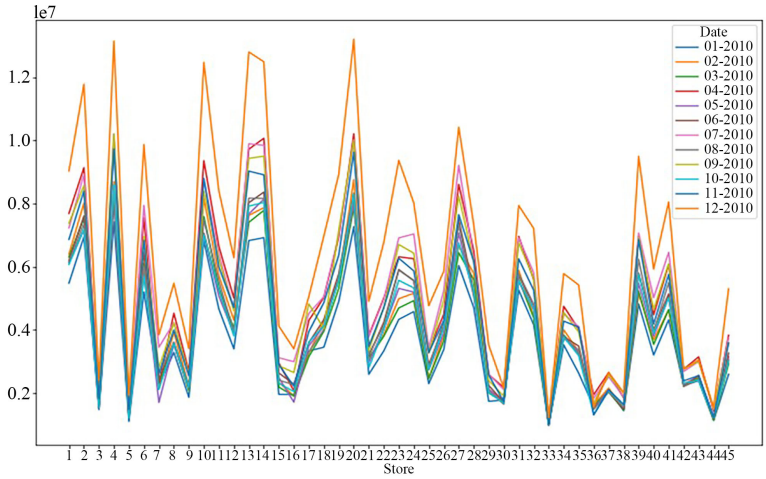

Figure 6. Line chart of monthly sales data of each store in the second year
图 6. 各家门店第二年各月销售数据折线图

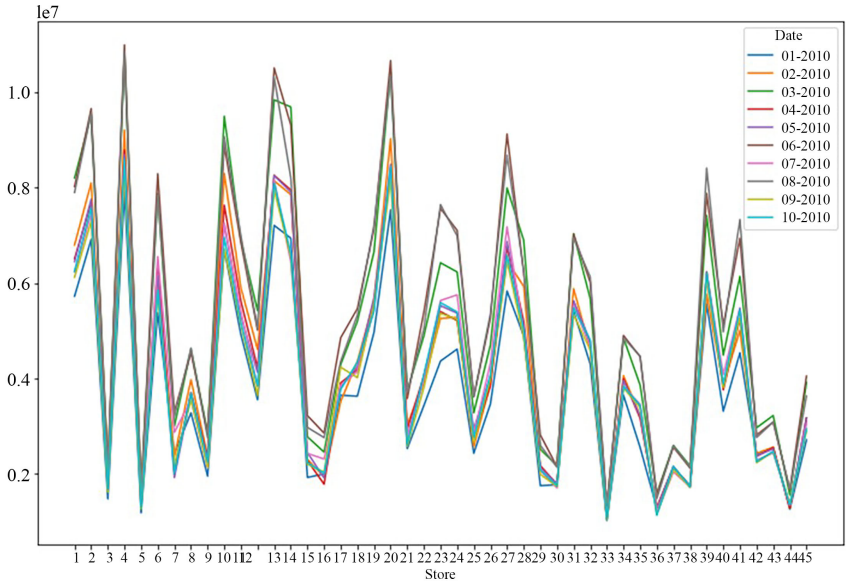

Figure 7. Line chart of monthly sales data of each store in the third year
图 7. 各家门店第三年各月销售数据折线图

从折线图中可以看出，45 家门店在 2010 年和 2011 年 2 月中的销售额普遍较高，2012 年较复杂，没有明显的区分。从折线图的走势来看，2010 年、2011 年及 2012 年各门店销售额的情况大致相同，没有太大变化。下面给出三年中 45 家门店各个月的销售总额，得到如下图 8 的柱状图。

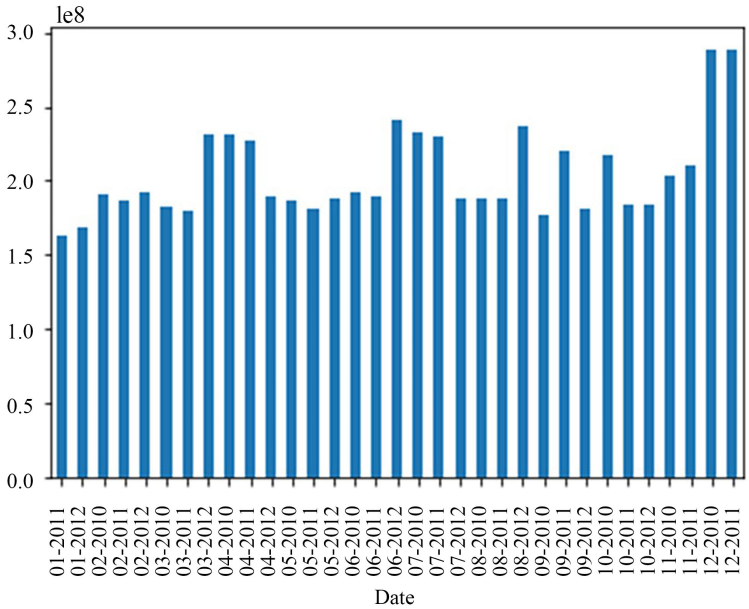

Figure 8. Bar chart of monthly sales data of each store for three years
图 8. 各门店三年各月销售数据柱状图

该统计柱状图表明，在十二月时的销售额会比较好，其他月份虽然偶尔销售额会有所增加，但基本比较稳定。将 2010 年和 2011 年各月的销售额数据以季度为单位，得到如下两年中 45 家门店各季度的柱状图分布，如下图 9 所示。

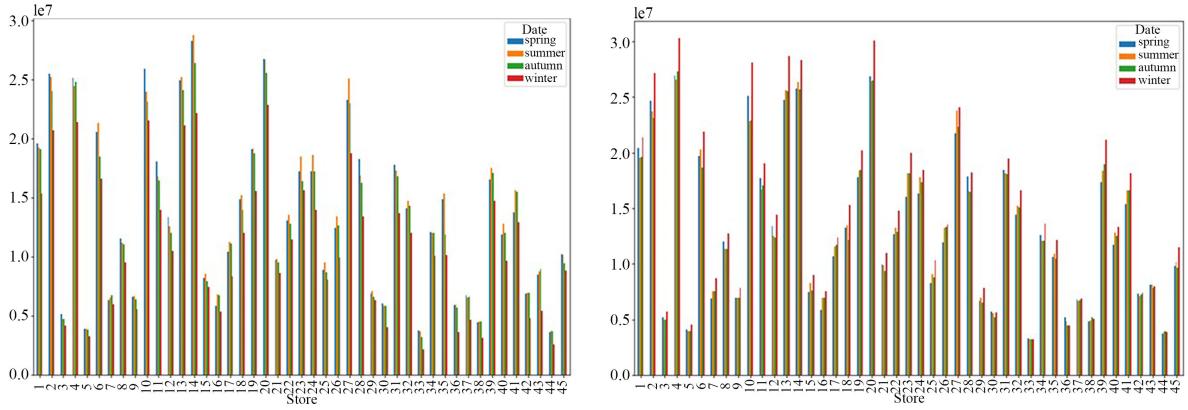

Figure 9. Bar chart of quarterly sales data of each store for three years in 2010 and 2011
图 9. 2010 年和 2011 年各门店三年各季度销售数据柱状图

从柱状图中可以看出，在 2010 年冬季的销售额最差，对于其他三个季节，不同门店的销售额各不相同，没有明显的规律，但普遍是春夏季的销售额较好。但在 2011 年冬季的销售额普遍最好，其他三个季节没有特别的规律。

所给定的沃尔玛数据提供了节假日标记信息，数据中 1 表示对应的节假日，0 表示一般工作日。节假日包括劳动节、感恩节、圣诞节等，为了了解节日对销售额产生的影响，对节假日销售额数据进行统计分析，将一般工作日的平均销售额作为正面和负面的影响指标，得到节假日销售额高于所有商店在非节假日期间的平均销售额的统计图，如图 10 所示。

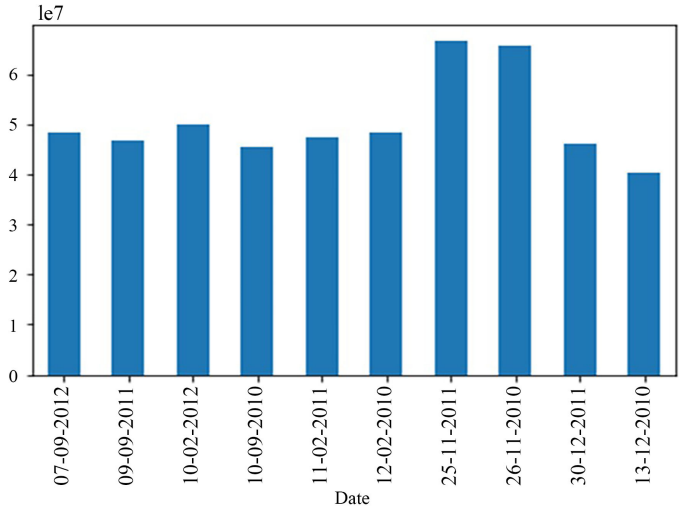

Figure 10. Statistical chart of holiday sales are higher than the average sales of all stores during non-holiday periods
图 10. 节假日销售额高于所有商店在非节假日期间的平均销售额的统计图

该柱状图统计了在这些节假日中，销售额高于一般工作日的销售额，节假日在沃尔玛销售中产生了有益的影响，提高了收益。

3.3. 本章小结

本章介绍了所使用的沃尔玛零售数据来源，介绍了数据的基本信息。通过对数据进行纵向分析发现，45 家门店在 2011 年的销售额最好，2012 年销售额最差。其中，第 20 家门店销售额最高，第 33 家门店

销售额最少且变化最小，第 14 家门店的销售额变化最大；横向比较发现，其中，45 家门店在 2010 年和 2011 年 2 月中的销售额普遍较高，三年各门店销售额的情况大致相同，没有太大变化；在对三年各月份的统计分析中，12 月的销售额最好，在 2010 年冬季的销售额最差，2011 年冬季的销售额普遍最好；最后，统计发现，在 2010-02-12、2010-09-10、2010-11-26、2010-12-31、2011-02-11、2011-09-09、2011-11-25、2011-12-30、2012-02-10、2012-09-07 这些节假日的销售额高于非节假日，对沃尔玛销售有帮助。因此，需要为各门店合理分配商品，为销售业绩比较好的门店派送充足商品，在 12 月之前需要囤积足够多的商品。

4. 数据销售预测

所提供的沃尔玛数据包含温度、燃油价、居民消费指数、失业率等属性，在使用这些属性预测沃尔玛销售额之前，需要考虑属性之间的相关性，找到解释变量之间具有较强关联的特征，能够观察属性之间存在的线性关系，有利于进行数据分析。在属性特征非常多的情况下，利用属性之间的相关关系往往可以达到变量选择的目的，减少多重共线问题，有利于提高模型稳健性，从而提升模型性能。

本文分析了数据中温度、燃油价、居民消费指数、失业率四个解释变量的相关性，得到如下图 11 所示的热力图。

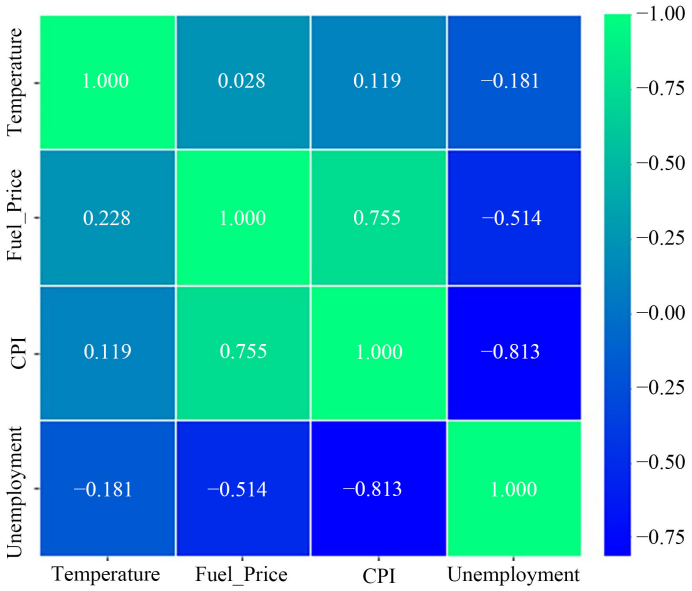

Figure 11. Correlated variable heat map
图 11. 相关变量热力图

从上面得到的热力图可以看出，燃油价和居民消费指数的相关性达到 0.755，具有较强的正相关性，另一方面，居民消费指数和失业率具有较强的负相关性，达到−0.813，从而，燃油价和失业率都与居民消费指数具有较强的线性关系，从图中也能看出燃油价和失业率仍然有着−0.514 的负相关关系；以上从沃尔玛数据得到的结果说明，燃油价的变动对居民消费指数影响较大，从而间接影响失业率，燃油价格升高在一定程度上会促进居民消费指数的上升，降低失业率，而燃油价格下降具有相反的变化趋势。

4.1. 数据预处理

在进行建模之前，需要将沃尔玛数据进行标准化处理，使得模型的输入数据控制在合理范围之内，这样可以消除量纲对数据预测的影响，使模型更稳定。

标准化处理的函数：

$$\overline{x_i} = \frac{x_i - \mu}{\delta}$$

其中，μ 表示样本 x_i 的均值，δ 表示样本 x_i 的方差。

在此，本文使用 2012 年的数据来训练模型，经过标准化处理的序列为如下表 2 所示。

Table 2. Normalized variable value
表 2. 标准化变量数值

Weekly_Sales	Temperature	Fuel_Price	CPI	Unemployment
0.5688	−1.8306	−1.521	−1.1302	1.2959
0.5577	−2.0982	−1.5774	−1.0966	1.2959
0.3648	−1.9982	−1.6572	−1.0858	1.2959
……	……	……	……	……

4.2. 基于多元线性回归模型预测拟合

本节将使用温度、燃油价、居民消费指数以及失业率等特征，建立多元线性回归模型预测销售额。先将原始数据划分为训练集和测试集，使测试集占所有数据的 20%。使用 Python 的 sklearn 模块中的 Linear Regression 方法，构建了一个多元线性回归模型，得到温度、燃油价、居民消费指数以及失业率等特征的系数为−0.2144、−0.0089、0.349、0.1448，由此，可以看出燃油价格对销售额的影响占比最小，居民消费指数对销售额的影响最大。将在训练集上训练好后的模型运用到测试集上，得到预测值及相关信息，如下表 3 所示。

Table 3. Comparison of actual values and predicted values
表 3. 实际值与预测值比较

实际值	预测值	残差	相对误差(%)
1428218.27	1552420.12	124201.85	8.70
1802477.43	1637125.28	−165352.14	−9.17
1635078.41	1511257.68	−123820.72	−7.57
1380020.27	1559190.34	179170.07	12.98
1494251.5	1478971.7	−15279.79	−1.02
1649614.93	1597187.95	−52426.97	−3.18
2270188.99	1657370.96	−612818.02	−26.99
1464693.46	1522419.03	57725.57	3.94
1661767.33	1553002.63	−108764.69	−6.55
……	……	……	……

4.3. 构建沃尔玛数据深度神经网络模型

深度神经网络模型被广泛应用于数据预测，本文将深度神经网络模型运用于沃尔玛数据的预测分析中，搭建一个具有两个隐藏层的神经网络。第一层隐藏层具有 10 个神经元，后跟一个 ReLU 激活函数；

第二层有 5 个神经元，后跟一个 ReLU 激活函数，最后是一个全连接层，其结构如下图 12 所示。

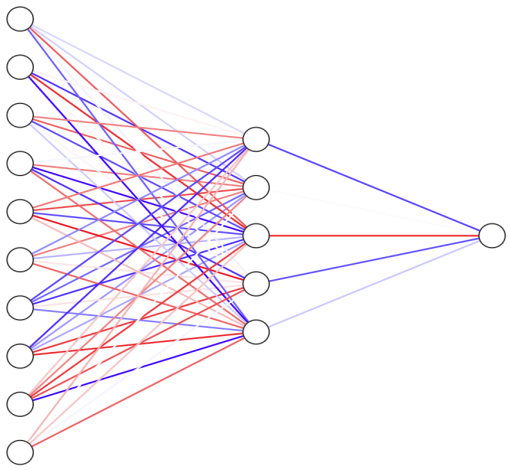

Figure 12. Neural network model schema
图 12. 神经网络模型图示

在模型训练之前，将数据划分成训练集、验证集和测试集，训练集和验证集用于模型训练过程中，将训练的批次大小设置为 120，训练 50 轮次，学习率为 0.001，训练得到该模型在训练集和验证集上损失函数的变化曲线，如下图 13 所示。

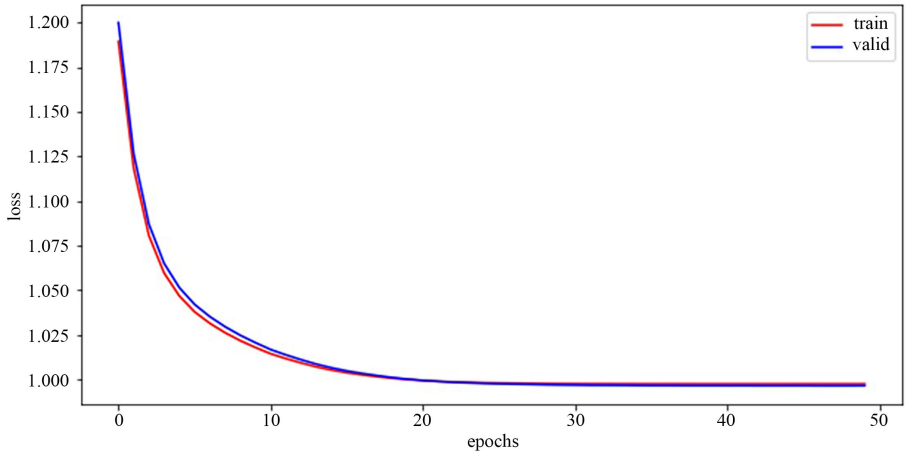

Figure 13. Loss function change curve
图 13. 损失函数变化曲线

这样就构建了一个深度神经网络模型，该模型可以用来预测其他门店的销售额。

4.4. 本章小结

本节对解释变量进行相关性分析发现，燃油价的变动对居民消费指数影响较大，从而间接影响失业率，燃油价格升高在一定程度上会促进居民消费指数的上升，降低失业率，而燃油价格下降具有相反的变化趋势。然后，对数据进行标准化处理，使用多元线性回归模型在测试集上预测了销售额数据，并展示了部分预测结果及相关信息。在本节最后，搭建了一个两层的深度神经网络模型并使用沃尔玛数据进行训练，该模型可用于预测其他门店的销售额。

5. 结论与建议

本文对从 *Kaggle* 竞赛网站上获取的沃尔玛数据进行数据分析并建立预测模型，得到如下结论：

1) 对数据中的 45 家门店进行纵向和横向分析，得出不同地区不同门店销售情况存在差异，其中，各门店在 2 月和 12 月的销售情况最好，建议这两个月前可以囤积一定量的货物；在 2010-02-12、2010-09-10、2010-11-26、2010-12-31、2011-02-11、2011-09-09、2011-11-25、2011-12-30、2012-02-10、2012-09-07 这些节假日的销售额普遍高于非节假日，也需要充分准备商品，为顾客提供满意的服务。

2) 在对响应变量温度、燃油价、居民消费指数、失业率等属性进行相关性分析中发现，燃油价格与居民消费指数呈正相关，失业率呈负相关；使用多元线性回归模型在测试集上预测了销售额数据，得出燃油价格对销售额的影响占比最小，居民消费指数对销售额的影响最大。

根据以上数据分析和建模的结果，为了帮助沃尔玛增加销售额，提高收益率，在此提出如下建议：

在日常销售过程中，根据以上分析结果显示，在 2 月和 12 月之前应该增加库存，在以下日期的节假日到来之前也应该增加库存：2010-02-12、2010-09-10、2010-11-26、2010-12-31、2011-02-11、2011-09-09、2011-11-25、2011-12-30、2012-02-10、2012-09-07。同时，沃尔玛的销售人员应该关注燃油价。燃油价对居民消费指数的影响较大，在建立多元线性回归模型进行预测中，居民消费指数对销售额影响较大，从而燃油价间接影响了销售额。

参考文献

[1] 王福, 刘俊华, 长青, 陈玉宏, 崔莹, 韩丽萍. 场景链如何赋能新零售商业模式生态化创新?——海尔智家案例研究[J/OL]. 南开管理评论, 2023: 1-22. http://kns.cnki.net/kcms/detail/12.1288.F.20230510.1615.004.html

[2] 靳江燕. 供应链集中度对零售企业价值创造的影响——基于批发零售上市企业数据的经验分析[J]. 商业经济研究, 2023(5): 54-57.

[3] 吴佳晨, 陈梦琳, 童彤, 王怡, 戴梦苇, 吕海萍. 生鲜电商新零售模式的消费者满意度影响因素分析——以盒马鲜生为例[J]. 中国市场, 2023(7): 118-120. https://doi.org/10.13939/j.cnki.zgsc.2023.07.118

[4] 王烽权, 江积海, 蔡春花. 相得益彰: 数据驱动新零售商业模式闭环的构建机理——盒马案例研究[J/OL]. 南开管理评论, 2022: 1-23. http://kns.cnki.net/kcms/detail/12.1288.F.20221027.1130.004.html

[5] 寇军, 赵泽洪. 大数据背景下产业链内部间隙弥补与企业运营效率提升[J]. 河北工程大学学报(社会科学版), 2023, 40(1): 18-26.

[6] 胡旭升, 叶本领, 张伟, 乐强勇. 提升基层烟草专卖管理数据分析能力[J]. 活力, 2022(19): 163-165.

[7] 朱萸, 何治呈. 基于知识流动热点数据的成品油销售企业零售客户流失建模分析[J]. 资源与产业, 2022, 24(5): 117-123. https://doi.org/10.13776/j.cnki.resourcesindustries.20221012.003

[8] 盛思诗. 基于数据驱动的 T 零售企业库存成本控制方案研究[D]: [硕士学位论文]. 长沙: 中南林业科技大学, 2022. https://doi.org/10.27662/d.cnki.gznlc.2022.000812

[9] 贺菁伟, 杨东谕. 长短期记忆模型在低频数据预测中的应用——以新冠肺炎疫情对北京市社会消费品零售总额的影响测算为例[J]. 统计理论与实践, 2021(1): 24-28.

[10] 李小星, 徐永利. 基于函数性数据分析的中国社会消费品零售总额数据的预测[J]. 北京化工大学学报(自然科学版), 2018, 45(3): 107-112. https://doi.org/10.13543/j.bhxbzr.2018.03.017

[11] 罗倩, 陈峻焜. 新零售业态下顾客消费意愿的影响因素研究——基于顾客体验视角[J]. 经济研究导刊, 2019(8): 116-121, 151.

[12] 王洪鑫, 刘玉慧. 网络购买生鲜农产品的消费者满意度影响因素实证研究[J]. 消费经济, 2015, 31(6): 81-86.

[13] 祝燕萍. 数据支持下的盒马鲜生新零售模式研究[J]. 现代商业, 2022(2): 12-14.

[14] 郭晓玲, 李凯. 供应链集中度、市场地位与企业研发投入: 横向与纵向的二维视角[J]. 产经评论, 2019, 10(2): 5-19.

[15] 江积海, 王若瑾. 新零售业态商业模式中的价值倍增动因及创造机理——永辉超级物种的案例研究[J]. 管理评论, 2020, 32(8): 325-336.

[16] 谢美娥. 大数据在物流管理中的应用研究[J]. 物流工程与管理, 2020, 42(3): 49-51.

[17] 田宸宇, 鲁逸飞, 谢恒多, 刘永建, 简思春, 康杭辉, 杨兴有, 鲁黎明. 多元回归与 BP 神经网络在烟草叶片 SPAD 反演时的建模精度比较[J]. 智慧农业导刊, 2023, 3(9): 9-11, 15. https://doi.org/10.20028/j.zhnydk.2023.09.003

[18] Farias-Basulto, G.A., Ulbrich, C., Schlatmann, R. and Klenk, R. (2023) Periodical Evaluation of Photovoltaic Modules and Diode Parameter Extraction Method Using Multiple Linear Regression Models. *Japanese Journal of Applied Physics*, **62**, SK1023. https://doi.org/10.35848/1347-4065/acc668

[19] Brix K.V., Tear, L., DeForest, D.K. and Adams, W.J. (2023) Development of Multiple Linear Regression Models for Predicting Chronic Iron Toxicity to Aquatic Organisms. *Environmental Toxicology and Chemistry*, **42**, 1386-1400. https://doi.org/10.1002/etc.5623

[20] Çetintürk, S., Tarakçioğlu, N., Acarer, M., Çunkaş, M. and Salur, E. (2023) Using Quadratic Multiple Linear Regression Models to Investigate the Effect of Inoculant Type and T6 Heat Treatment on Microstructural, Mechanical and Corrosion Properties of Al-Cu Alloy Produced by Casting. *Materials Today Communications*, **35**, Article ID: 105549. https://doi.org/10.1016/j.mtcomm.2023.105549

[21] Pan, C.H. (2022) Analysing the Characteristics of Precipitation Variation in Hefei in Recent Years Based on Multiple Linear Regression. *Academic Journal of Environment & Earth Science*, **4**, 35-40. https://doi.org/10.25236/AJEE.2022.040807