

# 浅析数据挖掘技术在农业价格预测领域的应用

刘宇萌

重庆三峡学院计算机科学与工程学院 重庆万州 404020

**摘要:**近年来,农产品价格在农产品市场中扮演的角色越来越重要,而农产品价格的不稳定对农民的收入以及生活质量造成了很大的影响,同时农产品价格也是政府宏观调控的依据之一。因此,需要运用科学的手段来对农产品价格进行调整。数据挖掘技术是通过自动化工具来对大量数据进行挖掘和分析的过程。通过对数据进行分析,以便更好的帮助决策者调整市场策略,做出正确的决策。运用数据挖掘技术对农产品价格进行预测,以便更好的分析农产品市场,帮助农业生产者及时调整市场策略,做出正确的决策。

**关键词:**数据挖掘;价格预测;农产品价格

## Analysis on the application of data mining technology in the field of agricultural product price forecast

Yumeng Liu

School of Computer Science and Engineering, Chongqing Three Gorges University, Wanzhou 404020

**Abstract:** In recent years, the price of agricultural products plays an increasingly important role in the market of agricultural products, and the instability of the price of agricultural products has a great impact on the income and quality of life of farmers, and the price of agricultural products is also one of the basis of the government's macro-control. Therefore, it is necessary to use scientific means to adjust the price of agricultural products. Data mining technology is a process of mining and analyzing a large amount of data through automatic tools. By analyzing the data, it can better help decision makers adjust market strategies and make correct decisions. Using data mining technology to forecast the price of agricultural products, so as to better analyze the agricultural market, help agricultural producers timely adjust the market strategy, make the right decision.

**Keywords:** Data mining; Price forecasting; Agricultural product price

### 一、研究意义

我国是一个农业大国,农产品价格的波动对我国农业生产者以及农产品市场造成了很大的影响,因此需要对农产品价格进行有效预测,了解农产品价格变化规律,掌握农产品价格变化趋势,有利于正确引导农产品流通和农业生产,实现农产品供需平衡,为政府和农户提供结构调整的依据,有效提高农民效益。农产品价格的预测是对农产品历史数据进行分析,从中发现隐藏的数据变化规律,为农业生产者以及管理部门提供帮助,因此,对农产品价格进行预测一直是农业部门关注的重点,也是农民十分关注的重点问题。国内外学者针对农产品价格预测问题进行了不断地探索,取得许多实质性的研究成果。农产品价格的影响因素是农产品价格预测的前提,针对对农产品价格影响因素进行分析主要分为定性分析和定量分析两种。定性分析主要是农业研究者通过文献分析,访问以及调查来对影响农产品价格因素进行分析,从而对农产品价格进行预测,但是这样预测出的结果受人为因素的影响,导致偏差很大,

预测精度低。定量分析通常需要运用大量的数据,农产品价格受许多因素的影响,价格变化具有多样性和非线性,因此需要通过数据挖掘技术对数据进行提取分析建模,相对于定性分析方法,预测结果更加准确。

### 二、数据挖掘

#### 1.数据挖掘的概述

数据挖掘(Data Mining, DM)是指对大量的不完全或者杂乱无章的信息中具有隐性规律且有价值的数据进行有效挖掘,并将挖掘出的有效数据转化为知识,从而进一步转化为价值的过程<sup>[1]</sup>。数据挖掘是通过大量的收集数据,分析数据,从这些数据中提取所需要的信息,从而寻找出数据中存在的规律、规则、知识以及模式、关联、变化、异常和有意义的结构,数据挖掘是一类深层次的数据分析方法<sup>[2]</sup>。

#### 2.数据挖掘技术通常包括以下步骤:

数据清洗和预处理:农业数据通常比较杂乱,需要进行清洗和预处理,以保证数据质量。数据挖掘技术可以自

动化地完成这些工作。

**特征选择和提取：**从农业数据中选取和提取最具代表性和相关性的特征，是农业价格预测中非常关键的一步。数据挖掘技术可以自动化地完成这些工作，并且还可以对不同特征的重要性进行排序和评估。

**建立预测模型：**在选定好特征后，可以使用数据挖掘技术建立预测模型，比如回归模型、神经网络模型、支持向量机模型等等。这些模型可以根据历史数据和现有特征来预测未来的农业价格。

**模型评估和优化：**对建立好的模型进行评估和优化，是保证预测结果准确性和稳定性的关键。数据挖掘技术可以帮助农业从业者对模型进行分析和优化，以提高预测效果。

### 3.数据挖掘常用的方法

数据挖掘是一种通过分析和发现大量数据中的模式和规律，来提取有用信息的过程。下面列出了数据挖掘中常用的方法：

#### (1) 分类

分类是一种将数据分成预定义的类别的方法。该方法通常用于预测新数据所属的类别。常用的分类方法包括决策树、支持向量机、朴素贝叶斯等。

#### (2) 回归分析

回归分析方法是指对大量数据进行处理，确定数据集中因变量与自变量的相关关系，建立相应的回归方程。在回归分析中，可以根据因变量和自变量个数来进行分类，可以分为一元回归分析和多元回归分析。根据自变量与因变量的函数表达式进行分类，可以分为线性回归分析和非线性回归分析。回归分析的预测结果较为精确，一般用于预测模型的构建。

#### (3) 聚类

聚类是把一组个体按照相似性归纳成若干类别即“物以类聚”其目的是使得属于同一类别的个体之间的距离尽可能小，而不同类别上的个体间的距离尽可能大<sup>[1]</sup>。组内相似性越大，组间差距越大，说明聚类效果越好。常用的聚类方法包括 K-Means 聚类、层次聚类等。

#### (4) 关联规则挖掘

关联规则挖掘又叫关联分析，关联规则挖掘是一种从数据集中发现有趣关联关系的方法<sup>[4]</sup>。反映了事物与事务

之间的相关性，如果两个或多个事务之间存在相关性，其中一个事务就能通过其他事务预测到。常用于实体商店或线上电商推荐系统。关联规则挖掘是一种从数据集中发现有趣关联关系的方法。该方法通常用于发现数据中的隐藏规律和关系。常用的关联规则挖掘算法包括 Apriori 算法、FP-growth 算法等。

#### (5) 特征

特征选择是从数据样本集的所有属性（特征）中选择出具有较强表达优势能够达到某一目标的若干属性，即初始属性集的一个子集，从而达到降低维度的目的<sup>[5]</sup>，通过特征选择可以减少原始数据获取的时间，避免数据冗余，缩减数据存储空间，提高模型运行效率和性能。

#### (6) 时间序列分析

时间序列分析是一种基于统计学方法的数据挖掘技术，适用于对时间序列数据进行建模和预测。该方法通常用于预测未来的趋势和模式。常用的时间序列分析方法包括 ARIMA 模型、SARIMA 模型等。

## 三、数据挖掘技术在农业价格预测领域的发展方向

**时间序列分析：**数据挖掘技术可以用于分析历史数据，以确定趋势和周期性，这对于农业价格预测非常重要。时间序列分析可以帮助预测未来价格趋势，并提供决策支持。

**机器学习：**机器学习是一种数据挖掘技术，可以用于构建预测模型。机器学习算法可以处理大量的农业数据，包括市场供求情况、气象数据、物流数据等等，从而构建出可靠的价格预测模型。

**模型集成：**模型集成是将多个模型结合起来，以提高预测精度和鲁棒性。在农业价格预测中，多个模型可以根据数据特点和需求进行组合，以提高预测准确性。

**大数据分析：**随着农业领域数据的不断增加，数据挖掘技术可以帮助挖掘这些数据的潜在价值。大数据分析可以帮助农业企业更好地了解市场情况，优化供应链，提高决策效率。

**人工智能：**人工智能是数据挖掘技术的一种进化形式，可以帮助农业企业更好地理解市场和消费者的需求。人工智能可以自动化价格预测过程，提高决策效率和准确性。

总之，随着数据挖掘技术的不断发展和应用，农业价格预测的准确性和效率将会不断提高，为农业企业的发展

提供有力支持。

#### 四、数据挖掘技术在农业价格预测领域应用优势

**处理大量数据：**农产品市场数据往往涉及到大量的数据，例如气象数据、土地利用数据、作物生长周期数据、交易量数据等等。数据挖掘技术可以处理这些大量数据，从中提取出关键的特征，帮助预测农产品的价格走势。

**发现隐藏模式：**数据挖掘技术可以帮助发现数据中的隐藏模式和趋势。例如，可以通过对历史数据的分析，发现某些天气模式或政策变化对农产品价格的影响。

**提高预测准确度：**数据挖掘技术可以通过对数据进行分析 and 建模，来预测农产品的价格走势。相比传统的方法，数据挖掘技术能够更加准确地预测价格的波动，提高预测的准确度。

**实时监测价格：**数据挖掘技术可以实时监测农产品市场的价格变化，及时发现价格的波动，从而提醒农民或交易者采取相应的行动，例如调整销售策略或采购策略。

**辅助决策：**数据挖掘技术可以为农业企业或政府部门提供决策支持，帮助他们更好地制定策略和政策，提高农业产业的效益。

总之，数据挖掘技术在农产品价格预测中的应用，可以帮助农业产业更好地应对市场变化，提高效益和竞争力。

#### 五、数据挖掘在农业价格预测中的具体应用

##### 1. 数据收集

进行农产品价格预测之前，首先要对农产品价格的影响因素数据进行收集，用户可以通过农业信息综合网站，国家统计局等收集相关农产品信息，包括供需，产量，价格，替代品价格等相关信息，这些收集到的信息称为原始数据，这些原始数据可能会存在缺失值，异常值，会包含许多噪声和其他干扰信息，直接对原始数据进行挖掘可能会得到错误的结果，所以需要运用一些数据处理工具对原始数据进行筛选和预处理，确保原始数据的准确性，使处理好的数据更加符合挖掘前的要求。

##### 2. 农产品价格影响因素

首先我们需要了解影响农产品价格的因素，通常农产品价格会受到很多自然因素和社会因素的影响，例如：替代品价格，农产品的产量，种植面积，居民消费水平，气

候因素，受灾面积等，这些因素相互之间都具备着一定的关联性。通过运用一定的方法分析影响因素之间的相关性，提取关联程度较高的价格影响因素输入模型中，得出较为准确的预测结果。

##### 3. 价格预测的研究方法

数据挖掘对于价格指数的预测方法大体可以分为两类：定性分析法和定量分析法。

定性分析主要是农业研究者通过文献分析，访问以及调查来对影响农产品价格因素进行分析，从而对农产品价格进行预测，但是这样预测出的结果受人为因素的影响，导致偏差很大，预测精度低。

定量分析通常需要运用大量的数据，农产品价格受许多因素的影响，价格变化具有多样性和非线性，因此需要通过数据挖掘技术对数据进行提取分析建模，相对于定性分析方法，预测结果更加准确。常用的定量分析方法有两种：

###### (1) 时间序列分析预测法

时间序列分析法就是根据商品过去一段时期的价格数据，寻找时间变化对价格的影响关系，依此搭建时间序列模型，预测出未来一段时间的价格指数数据。其主要目的就是发现价格过去的规律，并将过去的规律延伸到未来，依此对未来价格进行预测。

###### (2) 回归分析预测法

回归的主要思想就是，多种因素都可以对商品价格造成不可见的影响，价格和影响因素之间存在着因果关系。回归分析就是分析商品价格影响因素，收集大量商品价格数据和其影响因素数据，再利用数学方法对数据进行深度分析和处理，进一步研究价格和其影响因素之间的关系，并依此对未来价格进行预测。

##### 4. 数据挖掘预测模型

**线性回归模型：**这是一种基本的预测模型，它可以使用历史数据来预测未来价格。这个模型假设价格和其他因素之间存在线性关系，并使用回归分析来找出这些关系。

**决策树模型：**这是一种基于树形结构的预测模型，它可以根据历史数据中的特征值来预测未来价格。在决策树模型中，树的每个节点都代表一个特征值，每个分支代表一个可能的取值，最终叶子节点给出预测的价格。

**随机森林模型：**这是一种基于集成学习的预测模型，

它将多个决策树组合起来,以提高预测的准确性<sup>[6]</sup>。随机森林模型对于处理大量数据和具有复杂关系的数据集非常有效。

**支持向量机模型:**这是一种基于机器学习的预测模型,它使用一个高维空间中的超平面来区分不同的价格。支持向量机模型适用于小数据集和高维度数据集<sup>[7]</sup>。

**神经网络模型:**这是一种基于深度学习的预测模型,它使用多个神经元来学习输入数据中的模式,以便预测未来价格。神经网络模型适用于大量数据和具有复杂关系的数据集。

## 六、结束语

数据挖掘技术在农业领域的应用已经十分广泛,本文介绍了数据挖掘技术在农产品价格预测中的应用,数据挖掘的相关步骤及方法以及数据挖掘相关的价格预测模型。数据挖掘技术在农业价格预测领域的研究意义非常大,运用数据挖掘技术对农产品未来价格进行预测,可以更加准确地预测未来的农业价格,为农业生产和市场决策提供更加可靠的支撑,可以有效帮助农业生产者规避风险,及时监测市场情况,维持市场稳定有序发展,有利于政府宏观调控。随着社会的进步和发展,数据挖掘技术将会日趋成

熟,将会在越来越多的领域中发挥重要作用。我们处在信息化时代,采用信息化技术处理数据是必然趋势,将农业与信息技术相结合,加快推进农业现代化,推动农业高质量发展。

## 参考文献:

[1]叶婷,马宏娟,卢锐等.人工智能在智慧农业中的应用——以数据挖掘与机器学习为例 [J].智慧农业导刊,2022,2(18):27-29+32.

[2]陈章良.基于数据挖掘的预测决策模型应用与研究[J].中国管理信息化,2009,12(01):57-59.

[3]刘春玲,崔凌云,贾冬青等.数据挖掘技术在农业领域的应用[J].农机化研究,2010,32(07):201-204.

[4]李英,汤庸.基于关联规则与相似度的数据挖掘算法研究[J].华南师范大学学报(自然科学版),2021,53(05):121-127.

[5]蒋盛益,郑琪,张倩生.基于聚类的特征选择方法[J].电子学报,2008,36(增刊1):158-160.

[6]孙悦,袁健.基于 Spark 的改进随机森林算法[J].电子科技,2019,32(4):60-64.

[7]任靓.基于 RNN-SVM 模型的期货价格预测[D].辽宁师范大学,2020.