目前整体论文的思路已经出来了，接下来的工作就是对论文的整体进行一个语言的润色，增删改查

题目

作者

摘要：以筛选的16个基本面特征变量作为甲醇期货价格预测变量，利用随机森林构建预测模型。依据模型的预测精度来检验基本面特征对甲醇期货价格的预测能力。研究结果发现：经过归一化处理后的预测变量并没能让模型的预测精度提高，反而较大程度地增大了模型预测的不稳定性；特征变量甲醇毛利、甲醇现货价格、甲醇产量对预测模型的贡献度一直处于前列（对模型贡献度最高的），这表明这三个特征变量对甲醇期货价格具有预测意义；利用随机森林和XGBoost能够建立有效的甲醇期货价格预测模型。（考虑一下是不是要进行一下跟踪交易）（摘要中目前还缺少一个与时间序列分析之间进行的比较）

关键字：期货价格；价格预测；随机森林；Aberration交易策略

引言

期货交易是市场经济发展到一定阶段的产物，成熟的期货交易市场在一定程度上可以认为是完全竞争的市场，是经济学中最理想的市场形式。我国第一个商品期货交易所的出现要追溯到上世纪90年代，而世界上最早出现期货交易场所的时间是19世纪50年代，相比于一些国外期货市场而言，我国的期货交易市场还不太成熟，期货理论对现实期货交易的支撑还不够。国内的甲醇期货于2011年10月28日在郑州商品交易所上市，到今年已有十年之久，相关企业和投资者对甲醇期货的关注度也在不断上升，对甲醇期货价格进行分析对国内甲醇期现货市场的健康发展以及相关企业和投资者进行甲醇期现货交易具有重要意义。

在现有期货市场的研究中，研究方法主要分为数理统计和机器学习方法。其中数理统计模型包括自回归模型、差分自回归移动平均模型、广义自回归条件异常差模型等。机器学习方法包括决策树、支持向量机、随机森林、人工神经网络等【4】。随着计算机技术的发展，机器学习也逐渐被应用到金融领域的研究中，并表现出很好的研究性能【5】。Einav和Levin（2014）指出在经济学研究领域中，经济模型关注的重点主要放在能够代表工序变化的因素上面，研究人员需要对数据的异质性进行校正，并试图获得估计参数的异方差一致的标准误差和协方差。此外还需要通过不同模型来评估模型的鲁棒性[6]。与传统计量模型相比，机器学习算法对数据的前提假设较少，是以数据为驱动，将重点放在模型的输出结果上，进而揭示出特征对模型的解释能力[7]。机器学习被认为是计算机通过自身的不断观察和学习来提升自己与真实世界交互能力的一门学科[8]。Breiman指出随机森林算法模型不需要考虑多元回归模型中的多重共线性问题[9]，Pérez-Rave等（2019）指出随机森林算法可以有效避免模型的过拟合，是一个能够做出精确预测的强鲁棒性算法[11]。Chen和Guestrin（2016）提出的XGBoost算法已被成功应用于多个领域。黄卿等（2018）将XGBoost应用到股指期货的变动方向的预测上，验证了该方法可以有效应用于我国金融领域的研究【黄卿，谢合亮. 机器学习方法在股指期货预测中的应用研究——— 基于 BP 神经网络、SVM 和 XGBoost 的比较分析 [J].】。

本文基于ARIMA、随机森林和XGBoost回归模型构建对照模型，以甲醇期货价格作为ARIMA的输入变量，以甲醇产业链上下游产品的基本面数据作为随机森林和XGBoost模型的输入，将其应用于甲醇期货价格序列的预测分析，分析产业链上下游产品的基本面特征变量对甲醇期货价格预测的有效性。

预测变量的筛选方法

期货价格预测效果取决于预测变量的选择和预测模型的构建。在预测变量的选择上，众多实证研究表明产业链上下游产品的基本面数据对金融资产价格预测具有重要意义【13】。

1.甲醇产业链结构分析

甲醇是有合成气生产的重要化学品之一，它既是一种燃料，也是一种化工原料，目前市场上交易的甲醇主要以工业甲醇为主，其生产所需的原材料主要为煤、焦油、天然气等，其在生产过程中不会产生任何副产品，并且消耗过程中没有污染物的排放。甲醇自工业化生产以来，生产工艺也随着技术的进步以及能源结构的变化发生了较大变化，由原先的以煤和焦炭为原料进行生产，转变为以天然气为主要原材料进行生产。由于我国独特的能源结构的分布呈现出“富煤贫油少气”的特征，国内甲醇的生产依旧以煤为主。目前国内以煤作为原料甲醇生产产能占总产能的65%，以天然气为原料的产能占总产能的19%，以焦炉气为原料的产能占总产能的16%。

甲醇的下游产业链也十分丰富，涉及到化工、建材、能源、医药、农药等众多行业，在我国经济中扮演着十分重要角色。它不仅是一种重要的化工原料，可以用来生产甲醛、二甲醚、醋酸等，还是一种重要的燃料，用于加入汽油掺烧或者代替汽油作为动力燃料，同时还是农药生产的重要原料等。

2.备选预测变量选取与预处理

经过上述甲醇产业链结构分析可知，甲醇产业链的上游产品主要为煤、天然气、焦炉气，下游产品主要有甲醛、二甲醚、醋酸等。理想情况下，根据生产理论，上游产品作为下游产品的生产原料，其销售价格将作为下游产品生产的直接费用，并直接反映到下游产品的销售价格；从采购成本角度出发，上游产品的库存量也将直接影响到下游产品的采购成本，并呈现出负向相关关系；从供需关系的角度出发，当产品的供给充足时，价格会出现下降的趋势，反之则会出现上升的趋势，从而呈现出负相关关系，而产品的产能将直接影响产品的库存量，并表现出正相关性；从商品期货现实交易过程来看，期货交易最终将以现货的方式进行交割，在众多期现货价格研究中也已证实两者之间的强相关性。

根据上述分析，结合我国甲醇产业链的实际情况，以及数据的可获得性，本文选取的样本时间区间为2016年1月2日至2021年5月31日，选择的上下游产品分别为煤和醋酸，备选的基本面特征为：煤现货价格、产量、进口量、库存量；甲醇现货价格、产量、进口量、出口量、库存量、毛利；醋酸现货价格、进口量、出口量、库存量、消费量、毛利，共16个特征变量。其中三种产品的现货价格来自生意社网站，甲醇期货价格取自大智慧甲醇主力合约收盘价，其余基本面特征变量样本来源于钢联客户端。

3.备选预测变量预处理

由于本文所选取的备选特征变量中，有些数据是日频，而有些数据是周频和月频，需要对数据进行预处理。结合现实交易过程，样本数据基本是在交易日的收盘时刻才能出现，因而为了做出具有实用价值的模型，也需要对样本进行预处理。本文的处理方式是将备选预测变量数据向后偏移一个交易日，既当日的甲醇期货价格与前一交易日的备选特征变量值构成一个样本。对于较大口径特征变量，本文采用的方式是将两条相邻样本日期之间内所有交易日的数据均取值为离交易日较近的样本取值。将经过上述处理的各个特征变量与甲醇期货收盘价进行匹配以形成最终的样本。

3.备选预测变量相关性检验及预测变量筛选

对上述经过预处理后获得样本数据进行相关性检验。本文选用SPSS统计软件对样本数据进行Pearson、Spearman、Kendall相关性检验，检验结果如下所示。

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | | 甲醇产量 | 甲醇出口量 | 甲醇进口量 | 甲醇库存量 | 甲醇现货价格 | 甲醇毛利 | 天然气产量 | 天然气进口量 |
| 期货价格 | Pearson 相关性 | -.164\*\* | .270\*\* | -.506\*\* | -.457\*\* | .948\*\* | .959\*\* | -.122\*\* | -.009 |
| 显著性（双侧） | .000 | .000 | .000 | .000 | .000 | .000 | .000 | .754 |
| Kendall相关系数 | -.136\*\* | .143\*\* | -.344\*\* | -.322\*\* | .814\*\* | .849\*\* | -.122\*\* | -.010 |
| Sig.（双侧） | .000 | .000 | .000 | .000 | .000 | .000 | .000 | .618 |
| Spearman相关系数 | -.105\*\* | .195\*\* | -.506\*\* | -.458\*\* | .956\*\* | .968\*\* | -.173\*\* | -.039 |
| Sig.（双侧） | .000 | .000 | .000 | .000 | .000 | .000 | .000 | .170 |
|  | | 天然气库存量 | 天然气现货价格 | 冰醋酸出口量 | 冰醋酸进口量 | 冰醋酸库存量 | 冰醋酸消费量 | 冰醋酸现货价格 | 冰醋酸毛利 |
| 期货价格 | Pearson 相关性 | -.091\*\* | .425\*\* | .344\*\* | -.483\*\* | -.248\*\* | -.052 | .582\*\* | .379\*\* |
| 显著性（双侧） | .002 | .000 | .000 | .000 | .000 | .070 | .000 | .000 |
| Kendall相关系数 | -.090\*\* | .307\*\* | .250\*\* | -.301\*\* | -.175\*\* | .002 | .536\*\* | .324\*\* |
| Sig.（双侧） | .000 | .000 | .000 | .000 | .000 | .915 | .000 | .000 |
| Spearman相关系数 | -.123\*\* | .455\*\* | .347\*\* | -.430\*\* | -.252\*\* | .009 | .751\*\* | .480\*\* |
| Sig.（双侧） | .000 | .000 | .000 | .000 | .000 | .745 | .000 | .000 |
| \*\*. 在置信度（双测）为 0.01 时，相关性是显著的。 | | | | |  |  |  |  |  |

根据上表可以看出，在0.01显著水平（双侧检验）下，除了煤进口量、醋酸消费量之外的所有特征变量均与甲醇期货价格显著相关。根据相关性检验结果，本文将煤进口量和醋酸消费量之外的共14个备选特征变量选为预测变量。

本文仅将被预测变量与各备选预测变量之间的相关性检验结果呈现于上，根据Pearson相关性检验结果，各备选预测变量两两之间，共44对的相关性检验结果中，有33对相关性检验结果是显著的，表明预测变量之间是存在共线性的。（是不是放在这里不是很合适）

（以上变量的筛选是没有问题的，接下来就是处理实验的部分）

研究方法

根据被预测变量和预测变量之间的相关性检验结果，与被预测变量呈现出强相关性的预测变量仅为甲醇现货价格和甲醇毛利，其余变量呈现出中、弱相关性。根据各预测变量相互之间的相关性检验结果，在44对预测变量之间相关性检验结果中，有33对检验结果是显著的，表明预测变量之间是存在多重共线性。方匡南等（2011）指出随机森林不需要考虑共线性问题，对异常值和噪声有很好的容忍度，且不易出现过拟合，能够很好地处理高维度数据，且预测准确率高[14]。结合甲醇期货价格所呈现出的特点（是否要考虑一下和比特币所呈现出的动态性进行比较）（这里可能需要画一张图），本文选用适应能力较强的随机森林回归模型对甲醇期货价格进行预测，考虑到选取的预测变量量纲不一致可能对模型精度产生影响，因此对预测变量进行归一化处理构建对照模型，排除量纲对模型造成负面影响，并采用MAPE和R2作为模型的评价指标。（思考一下回测检验的描述放在哪里是比较合适的）为了检验模型的实用性，本文将设计回测检验方法来检验模型的实际价值。

甲醇期货价格预测模型构建

随机森林是决策树模型的集成算法，它的强实用性以及高预测精度已在大量实证分析中得到论证。它将n棵决策树的预测结果结合起来以做出比单棵决策树更为精确的预测，（这里要说一下什么是随机森林，可以参考一下其他的论文是如何进行论述的），随机森林在构造单棵决策树时除了在选择样本是使用随机抽样，在构建单棵决策树的输入特征时也采用随机抽样的方式从原有特征空间中选择子特征空间，此做法有效提高了模型的稳定性。随机森林回归模型的预测结果是由众多弱评估器的预测结果进行平均而得，即众多决策树预测结果的平均值作为整体模型的回归值。（这里可能还是要扩展一下，可以借助原始的paper来找公式）

随机森林回归模型的构建

随机森林回归模型的构建方法是：首先，从N个交易日时间序列样本集中，按7:3比例有序导出70%的样本量作为模型的训练集D，30%样本量作为模型的测试集；其次，使用bootstrap重抽样方法从训练集D中抽取同样数量的子样本集，重复此步骤n次生成子样本集Di（i=1，2，3，。。。n）；接着，以随机抽样的方式对上述生成的各个子样本集进行特征抽样，抽样的特征数量以随机方式获取并保持在70%以上，各个样本集根据特征抽样的结果进行训练子集数据的重新生成；再次，根据抽样的各子样本集构建决策树，并对n棵决策树的结果取算术平均值作为模型的模型的预测结果；最后，对上述模型进行调优，选出预测效果较高精度的回归模型。

考虑到本文所选取的预测变量的量纲不同，可能会对模型的预测结果产生影响。为了避免预测变量的不同量纲对预测模型造成负面影响，本文采取对被预测变量和筛选出的预测变量进行标准化处理，使用上述相同的方式构建随机森林回归模型，形成对照模型，对比模型回归结果选取甲醇期货价格的预测模型。标准化处理的方法为：

其中，为第i个特征变量的第j个样本值，为第i个特征变量的样本均值，为第i个特征变量的样本标准差，为第i个特征变量的第j个经过标准化后的样本值；为被预测变量的第j个样本值，为被预测变量的样本均值，为被预测变量的样本标准差，为被预测变量的第j个经过标准化后的样本值。

模型评估指标

本文采用的模型评估指标为平均绝对百分比误差（MAPE）和R2。设n为样本数，和分别表示为被预测变量的样本值和预测值，其计算方式为：

其中，为被预测变量的样本均值。

模型回测检验方法

依据上述对照模型预测结果，选定甲醇期货价格预测模型。根据甲醇期货价格回归模型预测结果构建回测交易规则，根据回测交易结果检验预测模型的有效性。本文选用的回归交易策略为经过修改的Aberration策略，并使用夏普比率来确定确定交易策略的相关参数。

Aberration交易策略分为两个部分，该策略的第一部分是相关参数的计算：

其中)为上轨线在t时刻的向上突破值，)为下轨线在t时刻的向下突破值，)为甲醇期货收盘价在t时刻的移动时长为N的移动平均值，)为甲醇期货收盘价在t时刻移动时长为N的移动标准差，和为阈值。

第二部分是交易规则，Aberration交易策略是建立在布林格通道Bollinger Bands理论的基础之上，Bollinger Bands理论是由John Bollinger在20世纪60年代提出，该理论指出金融市场产品价格将主要在由移动平均线以及在移动平均值基础上增加和减少一个单位的移动标准差形成的上轨和下轨线之间进行波动，当产品价格触及上轨或者下轨线时，价格将有很大可能性重新回到移动平均线附近。Keith Fitschen根据此理论进行了大量实验后发现布林格通道的上轨和下轨线其实是价格突破信号，即价格靠近上轨或下轨线附近的时候将有很大可能性超过上轨值或者跌破下轨值。本文采用的交易策略为对Aberration交易策略结合甲醇期货价格预测模型进行改进，具体方法为，使用甲醇期货价格预测模型预测后一交易日价格上涨超过上轨线时进行买多操作，并在价格回落到移动平均线下方时进行平仓操作；点那个第二交易日的预测价格跌破下轨线时进行买空操作，并在价格上升到移动平均线上方时进行平仓操作。

由于夏普比率是能够同时对收益和风险进行考虑，能够很好地评估回测交易效果，所以本文将根据夏普比率来进行交易策略的筛选，确定改进的Aberration交易策略参数、、，再根据确定后的Aberration交易策略来评估基于基本面特征的甲醇期货价格预测模型的有效性。具体步骤为首先根据两组对照模型的预测精度进行预测模型的选择，其次根据选定的预测模型进行甲醇期货价格的预测得出期货价格预测序列；再次，根据不同、、参数进行Aberration策略回测交易，本文中的取值为[20，60)，步长为1，的取值为[0.6, 2)，步长为0.1，的取值为[0.6, 2)，步长为0.1，并计算各交易策略的夏普比率；最后根据策略交易回测收益以及夏普比率选定交易策略作为甲醇期货价格预测模型的所能带来的收益，进而评估甲醇期货价格预测模型的有效性。为了提高交易策略的有效性，本文将使用选取的甲醇期货价格预测模型预测2016年5月1日至2020年5月1日时间区间内的甲醇期货价格，并使用此时间区间内的样本做为策略模型的筛选，然后使用2020年5月1日至2021年5月1日时间区间的样本进行选定的交易策略的测试，以检验交易策略的泛化能力。

甲醇期货价格预测与回测交易实证分析

1. 样本数据说明

本文选取2016年5月1日至2021年5月30日时间区间内甲醇期货主力合约收盘价作为被预测变量。备选预测变量为焦煤现货价格、产量、进口量、库存量；甲醇现货价格、产量、进口量、出口量、库存量、毛利；醋酸现货价格、进口量、出口量、库存量、消费量、毛利，共16个特征变量。其中现货价格数据来源于生意社网站，期货价格数据来源于大智慧行情系统，其他特征变量来源于钢联客户端。备选预测变量中包含日序列、周序列、月序列数据，本文采取的处理方式是交易日若无前一交易日对应的特征变量样本值则选用已有的离本交易日最近的样本值与之匹配，形成日序列预测向量。使用前移一个交易日的预测向量与交易日被预测变量进行匹配构成一个样本。

（二）预测变量的筛选结果

使用SPSS统计软件进行备选预测变量与被预测变量之间的相关性检验，检验结果如表1所示。在显著性水平0.01情况下，除了天然气进口量和冰醋酸消费量以外的备选预测变量均与甲醇期货价格显著相关，因而将天然气进口量和冰醋酸消费量以外的备选预测变量选为预测变量。

（三）两种预测模型的预测结果比较及选择预测模型

将经过预处理、筛选的样本数据输入到随机森林模型，并对样本数据进行标准化处理输入到随机森林模型构造对照模型。根据本文选定的模型评估指标对两个模型的参数进行调优，寻找相对最优模型。两个模型的实验结果为：

|  |  |  |
| --- | --- | --- |
| 误差 | R2 | MAPE |
| 模型1 | 0.990 | 0.012 |
| 模型2 | 0.991 | 0.103 |

比较上述模型的结果可以看出，在经过标准化处理后的样本数据在MAPE评估指标上要和模型1的结果相差7.5倍，而在R2这一评估指标上也并未能表现的更好，而仅仅只提高了0.1%左右，表明对样本数据进行标准化处理无法很好地提高模型的预测精度。因而本文选择未经标准化处理预测模型1作为甲醇期货价格预测模型。

（四）甲醇期货价格预测模型回测交易统计结果

使用模型1对所选样本时间区间的甲醇期货价格进行预测得出甲醇期货价格预测值，在2016年5月1日至2020年5月1日时间区间的甲醇期货价格模型预测值上使用前述改进的Aberration交易策略进行回测交易。由于不同的交易参数组合将产生不同的交易策略，进而可进行交易的次数也随之不同，本文除却那些在5年内交易次数小于15次以及交易次数大于25次并且买多或者买空次数低于5次的的交易策略。交易策略的统计结果为：符合要求的交易策略共274个，其中能够带来正收益的策略共139个；在能够带来正收益的策略中，仅有4个策略的K2值为0.7，其余均为0.6，并且这4个交易策略的收益分别为455、432、364、298，属于收益较小的策略，并且夏普比率均小于0.7，因而，将这4个策略排除作为最终策略模型的候选项，并将k2的值设定为0.6；将K2其他取值的策略舍弃后仅剩下135条策略，在所剩的135个策略中， K1取值为0.9时策略的收益和夏普比率均处于前列，并且两者的最大值均在其中。因而将K1的值设定为0.9；将K1其他取值的策略舍弃后仅剩30条策略；再去除夏普比率小于2的策略后仅剩下4条策略，分别为：

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| N | K1 | K2 | 收益 | 夏普比率 |
| 44 | 0.9 | 0.6 | 792 | 2.1884 |
| 43 | 0.9 | 0.6 | 766 | 2.0575 |
| 47 | 0.9 | 0.6 | 753 | 2.2465 |
| 48 | 0.9 | 0.6 | 731 | 2.1282 |

根据上表所示，当N的取值为44时，策略的夏普比率达到了2.1884，为所选策略中的第二大取值，但是此时的收益为所有策略中的最大值，因而N的取值可设定为44，即改进的Aberration交易策略中参数N、K1、K2取值分别为44、0.9、0.6时，模型的收益达到最优，策略的收益为792，将此策略的收益作为甲醇期货价格预测模型的收益。以在交易过程中需要的最大本金作为本次交易的本金，为2306，因而此次长期交易的收益率为34.35%，平均年收益率为6.87%，高于同期的10年期固定国债的利率3.1%一倍多。

为了检验上述回测交易策略的泛化能力，在2020年5月1日至2021年5月1日时间区间上应用该交易策略进行回测交易。回测交易的结果显示，依据此策略进行投资，投资收益可达414，年化收益率为21.25%。

可以看出，在结合本文提出的与之相应的Aberration交易策略后，基于基本面的期货价格预测模型具有很好的实用价值。

结论与启示

首先，本文借助经济学原理中的生产成本理论以及供需关系准则角度初步选取了模型的备选预测变量，从提高模型实用性角度出发将周序列、月序列数据细化为日序列数据，将被预测变量与向前偏移一个交易日的预测变量数据进行匹配形成样本向量，通过相关性检验后筛选出模型的预测变量。

其次，考虑到变量的不同量纲可能会对模型造成负面影响，本文采用了标准化的方式对边量进行处理，构造对照模型。根据两组对照模型的R2和MAPE模型评估指标结果选择甲醇期货价格预测模型。结果发现，经过使用经过标准化处理后的变量作为预测变量的预测模型并没有提高模型的精确度，反而大大增大了模型的误差。

最后，基于Aberration交易策略思想，使用2016年5月1日至2020年5月1日时间区间内的甲醇期货价格模型预测的结果进行回测交易，并使用夏普比率对交易策略进行筛选，并根据最终交易策略的表现结果来评估预测模型的实际价值。研究发现，改进的Aberration交易策略中周期取值为44，上轨线的阈值取为0.9，下轨线的阈值取为0.6时，交易策略的保守收益可以达到较大值，平均年化收益率可以达到6.87%，超过同期10年起国债利率1倍多。将此策略应用于2020年5月1日至2021年5月1日时间区间上的甲醇期货价格预测模型预测值所获得的收益达到414，年化收益率达到21.25%。策略的泛化结果表明，上文提出的甲醇期货价格预测模型在结合使用改进后的Aberration交易策略能够实现很好的投资收益。

研究表明，以甲醇产业链上下游相关产品的基本面特征作为模型的预测变量来预测甲醇的期货价格具有很高的使用价值，能够很精准对甲醇期货价格进行预测，结合改进的Aberration交易策略能够达到稳定且较高的交易收益。

参考文献

[1]周亚.基于XGBoost与弹性网络回归的集成模型对上证指数日极差的预测分析[J].中国物价,2021(05):68-71.

[2]李燕,何如海,廖宜静,聂雷.聚氯乙烯期货价格与现货价格传导关系实证研究[J].宿州学院学报,2013,28(01):14-18.

[3]范俊明,刘洪久,胡彦蓉.基于LSTM深度学习的大豆期货价格预测[J].价格月刊,2021(02):7-15.

[4]陈标金,王锋.宏观经济指标、技术指标与国债期货价格预测——基于随机森林机器学习的实证检验[J].统计与信息论坛,2019,34(06):29-35.

[5]范俊明,刘洪久,胡彦蓉.基于LSTM深度学习的大豆期货价格预测[J].价格月刊,2021(02):7-15.

[6]Liran Einav,Jonathan Levin. Economics in the age of big data[J]. Science,2014,346(6210).

[7]Alessandro Vespignani. Predicting the Behavior of Techno-Social Systems[J]. Science,2009,325(5939).

[8]李航. 统计学习方法[M].清华大学出版社:, 201905.403.

[9]Breiman, L. Random Forests. Machine Learning 45, 5–32 (2001).

[10]Sendhil Mullainathan,Jann Spiess. Machine Learning: An Applied Econometric Approach[J]. Journal of Economic Perspectives,2017,31(2).

[11]Pérez-Rave,Correa-Morales,González-Echavarría. A machine learning approach to big data regression analysis of real estate prices for inferential and predictive purposes[J]. Journal of Property Research,2019,36(1).

[12]Lucas Alexandre,Pegios Konstantinos,Kotsakis Evangelos,Clarke Dan. Price Forecasting for the Balancing Energy Market Using Machine-Learning Regression[J]. Energies,2020,13(20).

[13]马郑玮,琚心然,姚瑜,冯异立.基于产业链视角对石油期货与PTA期货价格关联性的分析[J/OL].现代化工:1-16[2021-06-29].<http://kns.cnki.net/kcms/detail/11.2172.tq>.20200910.084

5.005.html.

[14]方匡南,吴见彬,朱建平,谢邦昌.随机森林方法研究综述[J].统计与信息论坛,2011,26(03):32-38.

[15]Q Chen. XGBoost: A scalable tree boosting system. The 22nd ACM SIGKDD International Conference ACM, 2016.

[16]周亚.基于XGBoost与弹性网络回归的集成模型对上证指数日极差的预测分析[J].中国物价,2021(05):68-71.