文献综述

对期货的价格进行以下描述 （类似虚拟货币对于规避风险，投资组合的多样性有很大的帮助） 类似于背景的介绍

期货产品是怎样的

一、引言于文献综述

期货作为全球金融体系的重要的组成部分，是对股票、债券市场的一个重要补充，在调节经济方面具有重要的作用，对期货市场进行研究非常有意义。目前对期货的研究主要集中在能化产品及黑色系产品，能化产品主要集中在原油、PTA等，黑色系产品主要为螺纹钢、铜等，对能化产品职工的甲醇研究较少。

梳理国内外研究文献发现，对金融市场的的研究方法主要分为传统数理统计与计量经济学模型和机器学习算法模型。其中数理统计与计量经济学模型包括自回归模型、差分自回归移动平均模型、广义自回归条件异方差模型等。机器学习算法模型包括决策树、支持向量机回归模型、随机森林、人工神经网络等。随着计算力的提升，机器学习被逐渐应用到了金融领域的研究当中，并表现出良好的研究性能。

开始写文献研究的内容

王锐和陈倬（2011）运用VAR模型研究了各种因素对中国农产品价格波动的影响。

吴玉霞和温欣（2016）于志军和杨善林（2013）分别运用ARIMA模型和误差修正GARCH模型对股价进行了预测研究。

期货市场 价格受标的物价格、供求状况、信息不对称、周期变 化、社会政策、不可抗力等因素影响，波动极其不规律，导致试验数据在测试过程中存在显著噪声并表 现出不平稳特征。这些问题使得线性研究方法在对 该类型数据进行预测时显得有些力不从心，尤其在 面对更高精度预测需求时，线性模型更是会面临瓶颈，较大程度限制时间序列模型的运用与拓展。

1.基于动态模型平均的大豆期货价格预测研究

2. 中国甲醇期货市场的价格发现功能研究

与传统数理统计及计量经济学方法相比，机器学习算法对数据的前提假设较少，具有较好的实验效果。陈标金和王锋（2019）采用随机森林算法构建出有效的国债期货量化投资模型。周亚（2021）提出基于XGBoost和弹性网络回归的集成模型来有效识别影响股票指数的日极差预测的重要因素。范俊明、刘洪久和胡彦蓉（2021）采用LSTM网络模型对大豆期货价格进行了良好的预测。王纯杰、温丽男和马元嘉（2020）分别利用岭回归和Lasso回归两种方法消除解释变量之间的共线性，构建出修正的多元回归模型对螺纹钢价格的走势和预测具有重要的参考价值。R.KJana，Indranil Ghosh和Debojyoti Das（2021）提出基于差分进化算法（DE）的回归模型框架来对一个交易日后的比特币价格进行精准的预测。

Winky.K.o.Ho等（2020）使用随机森林、随机梯度上升和支持向量机对放假价格进行预测，得出随机森林和梯度上升算法能够较为准确的预测中国商品房的价格。

Feggella (2019), machine learning is about ‘the science of getting computers to learn and act like humans do, and improve their learning over time in autonomous fashion, by feeding them data and information in the form of observations and real-world interactions’.

Feggella（2019）认为机器学习是一门让计算机能够像人类一样行为，通过输入数据，不断地进行观察和学习来提升自己与真实世界的交互能够。

In econometrics research, economic models are built up to focus on a number of parameters that represent the effects of change in demand and/or supply factors. Researchers have to correct for heterogeneity and attempt to obtain heteroskedasticity-consistent standard errors and covariances for the estimated parameters (Einav & Levin, 2014). They also have to evaluate the robustness of the results by estimating several alternative model specifications. In contrast, machine learning approaches mostly put emphasis on the out-of-sample predic- tions, and rely on the data-driven model selection to unveil the features with the most explanatory power (Vespignani, 2009).

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是能够代表供需变化的因素上面。研究人员需要对异质性进行校正，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结果的鲁棒性。而机器学习方法以数据为驱动，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（Vespignani，2009）。

due to many decision trees participating in the process, random forest (RF) is used because it can reduce over-fitting of data. Previous property research has also demonstrated that random forest is a robust algo- rithm which provides accurate predictions (Mohd et al., 2019; Mullainathan & Spiess, 2017; Pérez-Rave et al., 2019).

Misra et al. [21] discussed linear classification with support vector machine (SVM) and artificial neural networks to accurately forecast the market’s daily movements.

随机森林算法是集成多棵决策树的集成模型，可以有效地避免模型的过拟合，Mohd et al.2019，Mullainathan、Spiess，2017和Pérez-Rave et al., 2019均指出随机森林是一个能够做出精确预测的强鲁棒性算法。

**Alexandre Lucas等（2020）使用梯度提升、随机森林和XGBoost对居民电价进行预测得出XGBoost对能源的价格能够较好的预测**

The model returns a Mean Absolute Error (MAE) of 7.89 £/MWh, a coefficient of determination (R2 score) of 76.8% and a Mean Squared Error (MSE) of 124.74. The variables that contribute the most to the model are the Net Imbalance Volume, the LOLP (aggregated), the month and the De-rated margins (aggregated) with 28.6%, 27.5%, 14.0%, and 8.9% of weight on feature importance respectively.

Improvement in computing technology has made it possible to examine social informa- tion that cannot previously be captured, processed and analysed. New analytical techni- ques of machine learning can be used in property research. This study is an exploratory attempt to use three machine learning algorithms in estimating housing prices, and then compare their results.

说一下目前的问题（比如说目前比特币的return behavior is chaotic and multifractal）

（如比特币目前存在的问题，其实也就是本文研究的目的或者主题）

一些研究研究处理期货价格和一些因素有关（如现货价格、库存、毛利等）

由于高度的非静态性、非线性、多种多样的行为和一些无法控制的因素影响着期货价格的评估或者预测（看有没有相关方面的论文可以引用）

传统的预测模型例如MA，Autoregressive（AR），autoregressive integrated moving average（ARIMA）无法正确的进行预测比特币的价格

并指出这些传统模型的缺点是要求数据是静态的，homoscedasticity of errors

研究人员提出机器学习，深度学习以及动态模型，计算机智能可以减轻这方面的限制

本文呈现出一个有效的基于回归的预测模型框架来预测后一个交易日的比特币的价格

一个精确的后一个交易日的预测模型可能为日常交易者做出更好的投资决策并在投资过程中获利

介绍一下本文的研究工作

含交叉项的多项式回归，随机森林，XGBoost作了哪些工作

对交叉项多项式回归，随机森林，XGBoost的超参数进行调优是必要的，in order to 确保较高准确度的预测模型

当然我们也要评估多项式回归的系数

使用MLP，RF，ANFIS，SVR，MARS

描述一下本研究的重要贡献是什么

可以说是找出了影响甲醇期货价格重要特征

同时有了这点之后还可以说一下未来的论文的研究方向是能否找出更为重要特征来预测模型

本文的下述内容分为了这样的几个部分：section2 展示了之前的研究

Section3 描述了数据和关键的统计特征of bitcoin prices

Section*4* 展示一下研究的方法 methodology

Section5 报告结果（obtained from the proposed framework）

Section6 总结paper