摘要：

本论文将采用机器学习方法从长期和短期两个维度分析股票周期性。长期上以发现股票长线周期性与宏观经济指标之间的联系为目标，分析与股票市场联系紧密的指标，包括:通货膨胀率、利率、货币同比增长率等，寻找指标和长期指数波动之间的相关性。同时加入股民心态分析，寻找周期性的转折点和投资者从众心理之间的联系。短期上以日为单位，使用机器学习的方法对股价进行预测，并且加入，成交曲线周期分量的提取，对模型进行优化和比较。模型将学习几种重要的经济金融指标影响股市的具体趋势,使经济金融数据可以直接用来指导股市的投资。

This paper will use the machine learning method to analyze the periodic stock of the stock from two dimensions: the long-term and short-term. In the long term, the link between the long-term cyclical and macroeconomic indicators of the stock is found to be the target of the analysis of the stock market, including the inflation rate, the rate of interest rates, the growth rate of the currency, and the correlation between the indicators and the long-term index fluctuations. At the same time, I will also participate in the analysis of the attitude of the shareholders, and find the link between the cyclical turning points and the investors' psychology. In the short term, the stock price is predicted by using the method of machine learning, and the extraction of the curve period component of the transaction curve is optimized and compared. The model will learn several important economic and financial indicators that affect the specific trends of the stock market, enabling economic and financial data to be used directly to guide the investment of the stock market.

总结

长期周期性上，通过分析我国股票市场的周期性，发现在投资者最疯狂的时刻恰恰是周期性的转折点，投资者从众行为非常严重。货币M1同比增长率，外汇占款与上证指数之间具有强周期性，该指标可以很好地预测周期性的行情。股市周期与经济周期具有很强的相关性，并且大部分时间先于经济周期性。

通过对成交数据的识别、修复、数据整合等操作，为股市短时预测奠定了数据基础。根据对交易日内两两Ｐｅａｒｓｏｎ相关系数分析，为后期提取周期性分量提供了理论基础。在短时预测中，引入傅里叶级数理论表示周期性分量。随着傅里叶级数展开项的增加，逐步分离剩余残余项，并运用长短时记忆模型分别输出残差项的预测结果，并和周期性分量部分合并为流量的预测值。运用本文提出的基于周期性分量提取的改进模型进行短时预测，根据傅里叶级数展开项数的不同，分析单步预测和多步预测的结论：

1. 在选取的短时走势预测中，长短时记忆神经网络预测模型表现出与支持向量机预测模型相近的性质，具备较高的鲁棒性和准确性。在选取的多个实证断面中，随着傅里叶级数展开项数的增多，长短时记忆神经网络预测的精度略优于支持向量机模型。作为尚在发展阶段的机器学习方法，长短时记忆神经网络在股市预测存在适用性和发展前景。
2. 在周期性分量提取的过程中，当展开项数较少时，傅里叶级数难以准确描述周期性部分，因此基于周期性分量提取的ＳＶＭ和ＬＳＴＭ预测指标具有一定的波动性。随着傅里叶级数展开项数的增多，基于周期性分量提取的改进模型预测指标最终均趋于平稳。
3. 对于三步以内的短时预测，ＳＶＭ和ＬＳＴＭ模型预测效果良好，考虑数据周期性分量提取的必要性不大。随着预测步数的增加，基于周期性分量提取的改进模型具有更高的准确性。
4. 对于六步及其以上的多步预测，随着预测步数的增加，基于周期性分量提取的ＳＶＭ和ＬＳＴＭ模型预测效果越来越好。且随着周期性分量提取，基于周期性分量提取的机器学习模型的预测优势比起原始模型更为显著。

展望

无

在股票走势数据分析中，发现数据具有一定的Pearson相关系数，即股票走势具有一定的周期性。借鉴对具有周期性速度序列的处理方法，考虑将股票走势分为周期性部分和剩余残差项。面临的问题是选用合适的方法分别预测周期性部分和剩余残差项。本章以周期性分量提取作为基础对走势数据周期性部分和剩余残差项进行分离，结合前沿研究深度学习，建立基于周期性分量提取的自回归积分滑动平均模型、支持向量机模型和长短时记忆神经网络模型。

4.1周期性分量预测

结合股票走势数据的周期性特点，根据己有研究具有周期性时间序列的处理方法，可以将交易日期间的走势分为两个部分，一部分是周期性部分即工作日期间趋势的确定值，另一部分是走势真实值去除周期性部分的剩余残差项，表达式如式(4-1)所示。

Qt=Mt+Qtr （4-1）

式中，Qt为t时刻的观测走势;Mt为周期性部分;Qtr为剩余残差项。

在周期性分量Mti的求解中，根据离散傅里叶变换，使用正弦曲线组合的三角函数回归模型进行表示具有固定时间周期性变化和周期性模式的时间序列。在本章的改进模型中，Mti作为傅里叶级数展开项数为i的周期性分量，Qtr为周期性分量提取后的残余项，应用短时预测的线性模型和机器学习模型分别进行预测。

本文对周期性部分Mt，和周期性分量Mti两个概念描述如下。周期性部分Mt是一个根据某个断面多日真实走势求得的确定序列，用于确定傅里叶级数的各项系数。根据傅里叶级数展开项数i的不同表示出该傅里叶级数展开项数下的周期性分量Mti。其中，周期性部分Mt，仅用于傅里叶级数系数的求解，在后期预测过程中用到的均是周期性分量Mti。

4.1.2周期性分量求解

选取同一断面p天具有周期性的走势序列，以同一时刻多日的平均值作为该断面在这一时刻走势的周期性部分Mt，即

(4-13)

式中，mtd为该断面d日t时刻走势的平均值，d=1,2,...p。

对确定的周期性部分M，序进行傅里叶级数展开得到周期性分量，如式(4-14)所示。

其中，根据既有的周期性部分从，结合式(4-8)、式(4-11)和式(4-12)计算傅里叶级数各项系数。选择傅里叶级数展开项数i，进而确定周期性分量的表达式，实现未来时刻周期性分量部分的预测。

本章根据工作日中同一走势的周期性，提出基于周期性分量提取的短时预测方法。应用傅里叶级数理论进行走势周期性分量的预测。针对逐步分离周期性分量的残差项，引入目前处于快速发展阶段的长短时记忆神经网络模型进行预测，并选择经典的自回归积分滑动平均模型和支持向量机模型进行对比，最后将周期性分量与残差项进行合并，得到未来时刻走势的预测值。

实证：

根据基于周期性分量提取的短时预测模型，结合成交价数据，针对交易日走势进行实证性研究。利用时间序列{q1,q2,...qr}对未来ts个时刻后的走势值qT+ts预测，其中qr为第T个时刻的走势值，is为预测步长。当ts=1时，该预测为单步预测，当ts > 1时，该预测为多步预测。本文原始走势数据时间间隔为2 min，步长分别为1, 3.6, 9, 12，分析比较不同预测步长下基于周期性分量提取的各种预测模型。

针对交易日走势，进行基于周期性分量提取的实证性分析。模型经过添加周期分量进行预测，根据单步预测和多步预测结果，总结出基于周期性分量提取的改进模型适用条件和预测结论。

根据选取的断面预测结果显示，随着傅里叶级数展开项数的增大，基于周期J性分量提取的自回归积分滑动平均模型、支持向量机模型和长短时记忆神经网络模型均呈现出预测指标稳定的趋势。傅里叶级数展开项数和预测步长越多，周期性分量的提取对SVM模型和LSTM模型的改善效果越显著，同时，基于周期性分量提取的LSTM模型预测精度略优于同等傅里叶级数展开项数下的SVM模型。

本章：

由于机器学习的模型结构可更有效地挖掘时间序列内在的非线性规律，支持向量机模型和长短时记忆神经网络显示出比自回归积分滑动平均模型更好的预测精度和变化趋势。单步预测中，改进前后的模型预测指标差别不明显，随着预测步数的增加，基于周期性分量提取的SVM模型和LSTM模型具有更好的预测效果。同时，基于周期性分量提取的改进模型对出走势的预测优化效果同样适用，一定程度上反映出基于周期性分量提取的改进预测模型具有较强的鲁棒性和推广性。

本文：