

C 题 商品期货涨跌幅预测问题

商品期货（如螺纹钢、铁矿石、焦炭、焦煤等）是金融市场中的重要交易品种，其价格波动受到多种因素的影响，包括供需关系、宏观经济政策、国际市场变化等。若能利用历史数据预测商品期货未来的涨跌幅，则可帮助投资者更好地进行交易决策。

现有数据集为 1 分钟级数据，包括时间戳、开盘价、最高价、最低价、收盘价、成交量、持仓量等。请基于该数据集建立数学模型，预测商品期货未来 30 分钟的涨跌幅。涨跌幅定义为

$$\text{涨跌幅} = \frac{P_{t+30} - P_t}{P_t} \times 100\%,$$

其中 P_t 是当前时刻的价格， P_{t+30} 是 30 分钟后的价格。要求从 1 分钟级数据中提取出可能影响 30 分钟涨跌幅的特征，选择合适的机器学习模型对未来 30 分钟的涨跌幅进行预测。解释模型的选择理由，并使用适当的评价指标评估模型的性能，讨论模型的局限性及可能的改进方向。

提交内容：

1. 数据预处理和特征提取的详细步骤。
2. 模型选择的理由及模型的具体实现。
3. 模型训练和验证的过程及结果。
4. 模型的预测效果分析及改进建议。
5. 程序源码及文档。

附录 数据集

下载地址：<https://pan.baidu.com/s/14kb1ZSvAIDMUxOrQw6GJoQ>（提取码：esb1）

数据说明：data 文件夹下包含了从 2017 年 1 月 3 日至 2025 年 4 月 18 日的商品期货 1 分钟级主力合约数据，包括不锈钢（SS）、硅铁（SF）、锰硅（SM）、螺纹钢（RB）、热轧卷板（HC）、铁矿石（I）、焦煤（JM）。此处的主力合约都是按持仓排序筛选出来的主力。如果某一天主力合约发生变化，那么前一天的主力合约和当天的主力合约数据都会出现在前一天数据里面。