Kaggle Optiver 实际波动率预测竞赛银牌算法概览

所用算法：

1. 基于股票的实际报价和交易数据，提取交易价差、报价价差、交易量等一系列特征组合F1；利用聚类算法，将113类股票聚成7类，并计算这7类股票的上述特征组合F2；利用KNN算法计算找到时序上最相近的数据点，并计算最相近时间上的特征组合F3；将F1、F2、F3这三类特征组合在一起得到最后的特征向量F。
2. 在上述特征基础上，建立机器学习LightGBM和神经网络模型，使用RMSPE作为损失函数，使用sklearn、lightgbm、tensorflow等模块搭建模型进行训练，达到当给定交易数据即可预测实际波动率的效果。
3. 将LightGBM和神经网络模型结果等权平均，即可达到top4%的排名。