

融合舆情的黄金期货价格预测：RFECV 特征选择与 DNN 模型

杨紫帆^{1*}, 赵艺晴¹, 徐浩然², 潘晨宇¹, 肖尧¹

1. 上海对外经贸大学 统计与数据科学学院, 上海 201620

2. 上海对外经贸大学 国际经贸学院, 上海 201620

DOI:10.61369/ASDS.2026010006

摘 要 : 本文提出了一个面向黄金期货价格预测的“情绪指标增强 + 特征筛选 + 深度学习”的统一建模框架, 重点探讨将多平台中文评论转化为可计算情绪特征, 并与传统市场变量融合以提升预测精度。首先, 基于情感词典与分词结果构建情绪强度, 并进一步给出情绪波动率与7日滚动相关的一致性系数, 使文本信息以时间序列形式进入模型输入。其次, 引入以 SVR 为基学习器的 RFECV 进行递归特征消除与交叉验证筛选, 保留关键价格变量、技术指标及情绪特征, 降低维度并抑制噪声。最后, 构建并对比 CNN- 线性回归混合模型与 DNN 模型进行预测评估, 结果表明加入情绪指标后两类模型的误差指标整体改善, 且对价格局部起伏与拐点刻画更贴近真实走势。该模型的作用与优点在于: 可将非结构化文本信号标准化为通用可迁移特征, 结合特征筛选与深度模型提高稳健性与泛化能力, 适用于多类金融时间序列的预测任务。

关 键 词 : 深度神经网络; 卷积神经网络; RFECV 递归特征消除; 支持向量回归; 黄金期货价格预测

Gold Futures Price Forecasting Incorporating Public Sentiment: RFECV Feature Selection and DNN Model

Yang Zifan^{1*}, Zhao Yiqing¹, Xu Haoran², Pan Chenyu¹, Xiao Yao¹

1.School of Statistics and Data Science, Shanghai University of International Business and Economic, Shanghai 201620

2.School of International Business, Shanghai University of International Business and Economic, Shanghai 201620

Abstract : This paper proposes a unified modeling framework for gold futures price forecasting, integrating “sentiment indicator enhancement + feature selection + deep learning.” It focuses on converting multi-platform Chinese comments into computable sentiment features and fusing them with traditional market variables to enhance prediction accuracy. First, sentiment intensity is constructed based on sentiment dictionaries and word segmentation results. Furthermore, a consistency coefficient is derived for the rolling 7-day correlation between sentiment volatility and price movements, enabling textual information to enter the model as time series data. Second, RFECV with SVR as the base learner is introduced for recursive feature elimination and cross-validation screening, retaining key price variables, technical indicators, and sentiment features to reduce dimensionality and suppress noise. Finally, a CNN-linear regression hybrid model and a DNN model are constructed and compared for prediction evaluation. Results demonstrate that incorporating sentiment indicators improves overall error metrics for both models and better captures local price fluctuations and inflection points aligned with actual market movements. The model's utility and advantages lie in: standardizing unstructured text signals into universal transferable features; enhancing robustness and generalization through feature selection and deep learning; and its applicability to diverse financial time series forecasting tasks.

Keywords : Deep Neural Network(DNN); Convolutional Neural Network (CNN); Recursive Feature Elimination with Cross-Validation (RFECV); Support Vector Regression (SVR); gold futures price forecasting

作者简介:

赵艺晴, 上海对外经贸大学统计与数据科学学院, 本科生, 研究方向为经济统计学;

徐浩然, 上海对外经贸大学国际经贸学院, 本科生, 研究方向为数字经济下的企业升级;

潘晨宇, 上海对外经贸大学统计与数据科学学院, 本科生, 研究方向为数字经济与大数据统计应用;

肖尧, 上海对外经贸大学统计与数据科学学院, 本科生, 研究方向为数据科学与大数据技术。

通讯作者: 杨紫帆, 上海对外经贸大学统计与数据科学学院, 本科生, 研究方向为统计学与机器学习。

引言

本文聚焦于金融时间序列在噪声强、驱动因素复杂条件下的短期预测，旨在构建一个可同时利用市场数值信息与网络文本信号、并具备可迁移性的算法模型框架^[1]。首先，从通用预测任务出发指出：仅依赖价格等结构化变量的模型往往难以反映突发信息带来的预期变化，而仅用文本信号又容易受平台差异与时序对齐影响，特征稳定性不足^[2]。其次，本文提出“情绪特征构建—特征筛选—深度预测”的流程：将多源中文评论转化为可滚动更新的情绪强度与一致性指标，并与技术指标共同作为输入；再通过递归特征消除与交叉验证筛去冗余变量，降低噪声干扰^[3]。最后，在统一输入与筛选机制下构建并对比不同深度模型，以黄金期货价格为例展示框架的通用性与可扩展性^[4]。实验结果表明，该框架在波动刻画与整体误差上均优于仅使用传统变量的基线方法。

一、理论方法介绍

（一）情绪强度指标构建

情绪强度用于衡量文本情感倾向的方向与强弱程度：正值表示积极，负值表示消极，绝对值越大说明情绪表达越强。本文对单条评论的情绪得分采用平均情感权重法（Average Sentiment Weighting, ASW），步骤如下^[5]：

文本分词：使用 jieba 分词对中文评论进行分词，得到词语序列；

情感词匹配：遍历词语，若该词出现在情感词典中，则提取其情感得分；

情感得分计算：将匹配到的情感词得分取算术平均，作为该条评论的最终情绪得分。其表达式为：

$$S = \frac{1}{N} \sum_{i=1}^N \text{sentiment}(W_i) \quad (1)$$

其中： S 为评论的情绪得分， N 为评论中匹配到的情感词数量， $\text{sentiment}(W_i)$ 为词语 W_i 在情感词典中的得分。若评论中未出现任何情感词，则令 $S = 0$ 。

采用平均情感权重法的主要原因是它更关注“情绪表达”的强度水平，而不是情感词出现的数量，从而减少评论长短差异带来的偏差。将评论情绪得分按时间汇总后，可以把市场心理转化为可建模的量化序列。在实际市场中，情绪强度往往会在重大事件前后出现更明显的变化，并可能与波动放大和阶段性趋势有关。

（二）情绪波动率的构建

情绪波动率用于衡量市场情绪在短期内变化的幅度，反映投资者情绪的不稳定性与突变特征^[6]。设情绪强度时间序列为 $s_t, t=1^T$ ，其中 s_t 表示第 t 日的情绪强度值。本文用相邻两期情绪强度的相对变化幅度定义情绪波动率，公式为：

$$v_t = \left| \frac{s_t - s_{t-1}}{s_{t-1}} \right| \times 100\% \quad (2)$$

当 $s_{t-1} = 0$ 时，可采用数值平滑技术，避免分母为 0 造成计算不稳定。

情绪波动率序列 $v_{t=2}^T$ 作为相对指标，可以削弱情绪强度量纲的影响，便于不同阶段进行比较；同时该指标恒为非负，通常更容易呈现右偏分布，这也更贴近金融时间序列中“波动”的常见统

计形态。

（三）情绪一致性的构建

情绪一致性是衡量情绪强度序列在短期内是否“延续”或“反转”的关键指标。它关注前一日的情绪对后一日情绪的影响方向与程度，可看作在滚动窗口内对一阶相关结构的刻画，从而识别情绪变化的持续性特征或反转性特征^[7]。

1. 数学模型与计算框架

设情绪强度时间序列数据集为有序集合： $D = (d_i, s_i)_{i=1}^T$ 。其中 d_i 为时间索引（通常为日期）， s_i 为对应日期的情绪强度值，且 $d_1 < d_2 < \dots < d_T$ 。

对于每个足够靠后的时间点 $t \geq 7$ ，定义一个长度为 7 的滚动窗口 W_t ，包含当前及前 6 天的情绪强度： $W_t = \{s_{t-6}, s_{t-5}, s_{t-4}, s_{t-3}, s_{t-2}, s_{t-1}, s_t\}$ 。从向量表示角度，窗口也可写为 $W_t = [s_{t-6}, s_{t-5}, s_{t-4}, s_{t-3}, s_{t-2}, s_{t-1}, s_t]^T \in R^7$ 。窗口长度取 7 的主要考虑两点因素：一是 7 天的窗口能覆盖一周的情绪变化节奏，包含了周末信息与情绪的累积效应；二是样本量足够用于相关计算，又能保持对近期变化的敏感度。

2. 一致性系数的计算过程

对每个窗口 W_t ，将其中 7 个观测值依次记为 $x_1, x_2, x_3, x_4, x_5, x_6, x_7$ ，其中 $x_i = s_{t-7+i}$ （ $i=1, 2, \dots, 7$ ）。在此基础上构造两个长度为 6 的序列：

原始子序列（ X ）：

$$X = [X_1, X_2, X_3, X_4, X_5, X_6] = [x_1, x_2, x_3, x_4, x_5, x_6]；$$

滞后一期子序列（ Y ）：

$$Y = [Y_1, Y_2, Y_3, Y_4, Y_5, Y_6] = [x_2, x_3, x_4, x_5, x_6, x_7]。$$

两者对应 6 组相邻日配对观测： $(X_k, Y_k) = (x_k, x_{k+1})$ ， $k=1, 2, \dots, 6$ ，每一组配对都表示相邻两日的情绪强度关系。

窗口 W_t 的一致性系数 c_t 定义为序列 X 与 Y 的样本皮尔逊相关系数：

$$c_t = r_{XY} = \frac{\sum_{k=1}^6 (X_k - \bar{X})(Y_k - \bar{Y})}{\sqrt{\sum_{k=1}^6 (X_k - \bar{X})^2} \cdot \sqrt{\sum_{k=1}^6 (Y_k - \bar{Y})^2}} \quad (3)$$

其中 $\bar{X} = \frac{1}{6} \sum_{k=1}^6 X_k$ ， $\bar{Y} = \frac{1}{6} \sum_{k=1}^6 Y_k$ 分别为两个序列的样本均值。

这一过程本质上是在每个滚动窗口内刻画情绪强度的“一阶相关”程度，用来反映相邻日期情绪变化是否呈现稳定的线性

依赖。

3. 系数的统计解释

一致性系数 C_t 的取值范围为 $[-1, 1]$ ：

当 C_t 接近 +1 时，说明窗口 W_t 内情绪具有较强延续性：前一日偏高（偏低）时，后一日往往也偏高（偏低）。

当 C_t 接近 0 时，说明情绪变化缺乏稳定模式，更接近随机波动。

当 C_t 接近 -1 时，说明情绪更偏向反转：前一日偏高（偏低）往往对应后一日回落（回升）。

最终可得到新的时间序列 $C_{t=7}^T$ ，其中每个 C_t 对应日期 d_t 的情绪一致性度量，为分析情绪的短期传导、惯性与修正特征提供时间维度上的依据。

综上，情绪波动率和情绪一致性从“变化幅度”和“变化规律”两方面刻画投资者情绪。情绪波动率关注相邻时点情绪强度的相对变动大小，用来识别情绪剧烈波动与相对平稳的阶段，反映市场对信息冲击的即时反应程度。情绪一致性则基于滚动窗口内相邻日情绪的相关性，判断情绪是延续还是反转，从而体现情绪的惯性与自我修正特征

（四）RFECV 方法介绍

RFECV（Recursive Feature Elimination with Cross-Validation）是一种嵌入式特征选择方法。其核心思路是先用基学习器训练模型，根据特征重要性逐步剔除贡献较小的特征；每删一轮，就用交叉验证评估当前特征子集的效果，最后选择验证表现最好的那一组特征及其数量^[8]。

设训练数据集为 $D = (x_i, y_i)_{i=1}^N$ ，其中 $x_i \in R^p$ 表示包含 p 个特征的输入变量， y_i 为目标变量。RFECV 在每次迭代中以预设的基学习器 $f(\cdot)$ 训练模型，并利用交叉验证得分函数

$$S_k = \frac{1}{K} \sum_{j=1}^K s(f_{-j}, D_j), \quad (4)$$

表 1：情绪指标描述性分析

指标	样本数	均值	标准差	最小值	中位数	最大值	偏度	峰度
情绪强度	2023	0.400	0.061	0.198	0.401	0.633	0.086	0.349
情绪一致系数	2021	-0.156	0.362	-1.000	-0.162	0.909	0.195	-0.560
情绪波动率	2022	2.107	21.129	-55.582	-0.538	105.432	0.680	0.956

从表 1 可以看出：

情绪强度整体分布较为平稳。均值约为 0.400，说明样本期内情绪整体偏正向；标准差为 0.061，波动不大。偏度 0.086 接近 0，整体较为对称；峰度 0.349 略高于正态，意味着在少数时点可能出现相对“更极端”的情绪值，但不算突出。

情绪一致系数的均值为 -0.156，整体呈现轻微的“反转倾向”，即情绪在相邻时段并不总是延续。标准差 0.362，说明不同时期差异明显。偏度 0.195 为轻微右偏；峰度 -0.560 较为平坦，极端值的集中程度不高。取值范围从 -1.000 到 0.909，说明样本中既有延续阶段，也有反转阶段，情绪演化并不单一。

情绪波动率的波动最强。均值约 2.107，但标准差高达 21.129，说明该指标本身非常不稳定；偏度 0.680 右偏、峰度 0.956 呈尖峰厚尾，更符合金融数据中“平时平稳、关键时点跳

其中 K 为交叉验证折数， $s(\cdot)$ 表示模型评价指标（如 R^2 、MSE 等）。通过比较不同特征子集下的交叉验证得分，可以得到更合适的特征组合，从而在模型效果与特征维度之间取得平衡。

（五）卷积神经网络介绍

卷积神经网络（CNN）是一类常用于处理网格状或具有局部结构数据的深度学习模型。其核心设计思想源于对生物视觉机制的模拟，通过局部连接、权重共享和空间下采样等机制，能够高效地从输入数据中逐层提取从低级到高级的抽象特征。

（六）深度神经网络介绍

深度神经网络（DNN）是一种由多层结构组成的人工神经网络模型，其核心思想在于通过多层非线性映射与特征变换，从复杂的数据中自动提取出具有代表性的高层特征。

二、实证分析

（一）情绪指标的构建

本研究基于 BosonNLP 情感词典，对多平台中文评论文本进行情绪量化，并构建日度情绪指标^[9]。流程如下：

对收集到的非结构化评论文本进行清洗与预处理，使用 jieba 分词得到规范化词序列；

将分词结果与 BosonNLP 情感词典匹配，识别情感词并提取对应得分；对单条评论中命中的情感词得分取算术平均，作为该评论的情绪强度。若未识别到情感词，则情绪得分记为 0；

以日期为索引，对各平台、各关键词下的评论情绪得分进行汇总，通过计算日度平均值，得到反映当日整体情绪强度的综合指标。

在此基础上，进一步构建情绪一致系数与情绪波动率。然后对三类情绪指标做描述性统计与分析。

1. 情绪指标描述性分析

动很大”的特征。最小值为 -55.582、最大值为 105.432，进一步体现极端波动时点的存在，往往对应重大事件冲击或情绪集中转向。

2. 情绪指标时间序列动态特征

情绪强度序列存在明显的波动聚集现象，在部分阶段会持续走高或走低，说明市场情绪并非随机波动，而是带有一定的阶段性特征。尤其在重大市场事件前后，情绪强度往往出现更突出的峰值，反映投资者对关键信息的集中反应。

情绪一致性（7 天一致性系数）在样本期内起伏较大，且正负值频繁切换，说明情绪在“延续”与“反转”之间不断转换。滚动均值结果也支持这一点：有些阶段一致性保持为正，意味着情绪更容易延续，前一日情绪对后一日具有较明显的正向带动；而在另一些阶段，一致性下降甚至为负，更像是情绪出现了回调与

修正, 市场情绪的自我调整更明显。

情绪波动率(日变化率)的序列同样呈现典型的波动率聚集特征, 高波动期与低波动期交替出现。从30天滚动均值与滚动标准差来看, 波动率在部分阶段会快速抬升, 通常对应市场不确定性上升或信息冲击更频繁的时期; 滚动标准差区间较宽也说明该指标本身不稳定, 容易在特定时点被外部冲击拉大。

3. 情绪指标间相关性结构

为考察三类情绪指标之间的关系, 本文进一步进行了相关性分析。结果显示: 情绪强度与情绪波动率呈微弱负相关(约-0.15), 说明情绪达到极端时, 波动率未必同步处于最高水平, 可能存在“情绪冲击后短暂平静”的阶段。情绪强度与情绪一致性相关性较弱(约0.09), 表明情绪水平的高低与情绪是否延续并不存在明显对应关系。情绪波动率与情绪一致性呈负相关(约-0.28), 说明情绪变化越剧烈, 越难形成稳定的延续结构, 一致性更容易被打破。

(二) RFECV 特征变量筛选

在黄金期货价格预测建模中, 为提升训练效率并减少无效特征带来的干扰, 本文在深度学习建模前引入基于支持向量回归(SVR)的递归特征消除交叉验证(RFECV)进行特征筛选^[10]。

实验中选用线性核 SVR 作为基础学习器。该模型结构简单、结果更便于解释, 可为特征重要性评估提供相对稳定的依据。考虑到黄金期货数据具有明显的时间序列属性, 筛选过程中采用时间序列交叉验证而非随机划分, 以保证训练集与验证集在时间上保持前后顺序, 降低信息泄露风险。

目标变量设为“次日收盘价”, 其余变量作为候选输入特征参与筛选。通过 RFECV 的逐步迭代, 保留了一组核心特征用于后续建模。

根据筛选结果, 本文将“期货收盘价(连续): 黄金”“期货开盘价(连续): 黄金”“期货最高价(连续): 黄金”“期货最低价(连续): 黄金”“黄金储备: 中国”“现货价: 黄金”“RSI12”“SMA30”等特征变量, 与构造的三个情绪指标一并作为模型输入, 用于后续预测实验。

总体来看, RFECV 在本文中主要起到两方面作用: 一是减少输入维度, 提高 DNN 与 CNN- 线性回归混合模型的训练效率与收敛速度; 二是降低冗余与噪声特征的影响, 缓解过拟合, 使模型在金融时序数据下的表现更稳定。

(三) 模型实验阶段

在黄金期货价格预测的实证研究中, 本文构建并比较两类模型: CNN- 线性回归混合模型与深度神经网络(DNN), 用于评估不同结构在预测任务中的表现差异。

1. CNN+ 线性回归混合模型

该模型将局部特征提取与回归预测结合起来: 首先通过一维卷积层(Conv1d)提取局部模式, 设置64个滤波器、卷积核大小为3, 并采用 padding=1 保持维度; 卷积层后接 ReLU 激活, 再通过最大池化层(MaxPool1d)降维(池化核大小为2)。为抑制过拟合, 在卷积模块后加入 Dropout(丢弃率0.3)。卷积

输出的特征图展平后输入全连接网络, 全连接部分包含三层线性层, 隐藏层维度为100与50, 层间使用 ReLU 与 Dropout(丢弃率0.2), 最终输出单值预测。训练时采用 MSE 损失与 Adam 优化器, 学习率0.001, 并加入 L2正则化(权重衰减 $1e-4$)与梯度裁剪(最大范数1.0)提高训练稳定性; 学习率调度使用 ReduceLROnPlateau, 当验证损失进入平台期超过20轮时学习率减半。训练300轮, batch size 为32。

2. 深度神经网络(DNN)模型

DNN 采用四个隐藏层, 维度依次为256、128、64、32, 形成逐层递减的结构。每个隐藏层后加入批归一化以加速训练并提升泛化能力, 激活函数使用 LeakyReLU(负斜率0.1)以降低“神经元失活”风险; 各层设置 Dropout(丢弃率0.2)。权重使用 Kaiming 正态初始化以匹配 LeakyReLU。损失函数采用 SmoothL1Loss, 提高对异常值的鲁棒性; 优化器为 AdamW, 学习率0.001, 权重衰减 $1e-3$, 并结合余弦退火热身重启调整学习率($T_0=50$, $T_{mult}=2$)。训练时设置早停: 连续30轮无改善则提前结束, 最大训练轮数为500。

3. 测试集表现与区间差异

从测试集拟合效果看, 两种模型在不同价格区间的表现存在差异。DNN 在480-550区间拟合较好, 预测曲线与真实价格更接近; CNN+ 线性回归模型在520-580的中位价格区间表现相对稳定, 卷积层提取的局部特征有助于识别短期价格形态。

但在价格极端区域, 两者都出现明显偏差: 当价格突破650后, 预测值普遍低于实际值约5%-8%, 形成持续的负向偏差, 不过仍能较好跟随价格波动形态。一个可能的原因是训练数据在高价区样本较少, 模型在学习绝对价格水平时更偏保守, 导致高价阶段系统性低估。

(四) 模型对比

为比较 CNN+ 线性回归混合模型与深度神经网络(DNN)在黄金期货预测任务中的效果, 本文采用 MAE、RMSE、 R^2 与 MAPE 四项指标进行评估, 结果见表2。

表2: 两模型指标对比

评估指标	CNN+ 线性回归	深度神经网络
MAE	20.913	17.425
RMSE	24.075	21.07
R^2	0.941	0.954
MAPE (%)	0.954	2.688

从结果看, 两种模型侧重点不同。DNN 在 MAE、RMSE 和 R^2 上更优, 说明其对绝对误差的控制更好, 整体拟合能力也更强; CNN+ 线性回归模型在 MAPE 上更低, 说明其在相对误差控制方面更占优势。也就是说, 如果更关注“误差的绝对大小”, DNN 更合适; 如果更关注“相对误差比例”, CNN+ 线性回归模型表现更稳。

(五) 情绪指标贡献

为检验情绪信息的增量价值, 本文对比了“加入情绪指标”和“未加入情绪指标”两种情况下的模型表现, 结果如表3所示。

表 3：两模型有无情绪指标模型指标对比			
评估指标	模型类型	有情绪指标	无情绪指标
MAE	CNN+ 线性回归	15.011	20.913
	深度神经网络	17.425	29.536
RMSE	CNN+ 线性回归	24.075	26.344
	深度神经网络	21.07	37.616
R^2	CNN+ 线性回归	0.941	0.93
	深度神经网络	0.954	0.857
MAPE	CNN+ 线性回归	0.954	2.46
	深度神经网络	2.688	4.729

先看总体趋势：对 DNN 来说，引入情绪指标后 MAE 与 RMSE 明显降低， R^2 提升，MAPE 也下降，说明情绪信息确实增强了模型对价格波动的刻画能力，预测更贴近真实走势。对 CNN+ 线性回归模型而言，RMSE、 R^2 与 MAPE 的变化同样显示出情绪指标带来的改善。

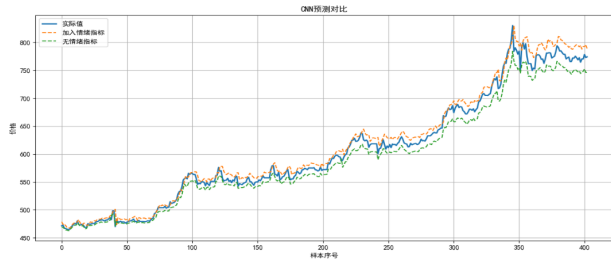


图 1：CNN+ 线性模型有无情绪指标预测结果对比

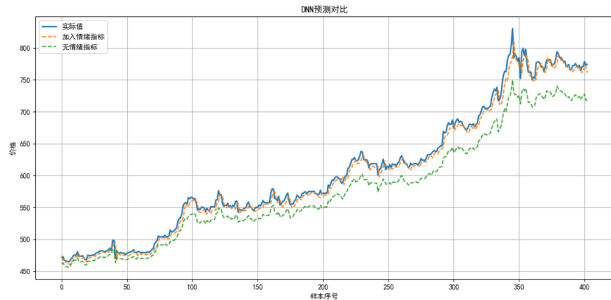


图 2：DNN 有无情绪指标预测结果对比

由图 1 和图 2 可知，情绪指标对模型预测的改进不仅体现在数值精度上，还表现在对价格波动形态的刻画上，在加入情绪特征后，模型对局部起伏、拐点变化的刻画更贴近真实走势，证明了情绪因素作为价格先行指标的有效性。

综上，情绪指标的引入能够降低预测误差、提高了模型拟合优度，还增强了模型对不同市场环境的适应能力，显著提升了 DNN 和 CNN 两类模型在黄金期货价格预测中的综合性。这也意味着，将情绪信息作为补充特征，有助于提高模型输出在实际量化分析与策略辅助中的可用性与稳定性。

三、结束语

本文提出了一个融合多源舆情情绪特征与市场结构化变量的黄金期货价格预测框架，能够将行为金融中的“情绪冲击”转化为可计算信号，并提升预测的稳定性与泛化性。首先，通过情绪强度、情绪一致系数与情绪波动率刻画市场预期变化，其中情绪波动率标准差达 21.129、区间为 [-55.582, 105.432]，对突发波动更敏感。其次，采用以 SVR 为基学习器的 RFECV 进行特征筛选，减少冗余与噪声，使输入更精炼、迁移更方便。最后，在统一管线下对比深度模型，DNN 取得 MAE=17.425、RMSE=21.07、($R^2=0.954$)，显示出更好的整体拟合与波动刻画能力。未来研究可进一步融合事件级文本与跨市场信息以增强鲁棒性。

参考文献

[1] 范彩云, 童君逸, 程俊彦, 等. 基于 ML-DMA 的黄金期货价格预测研究 [J]. 数理统计与管理, 2024, 43(03): 541-558. DOI: 10.13860/j.cnki.sltj.20240509-009.

[2] 杨文清. 基于文本挖掘的投资者情绪对股票收益率的影响 [D]. 哈尔滨工业大学, 2021. DOI: 10.27061/d.cnki.ghgdu.2021.001613.

[3] 孙晶. 基于 RFECV 特征选择和随机森林预测模型的应用与优化 [J]. 数字通信世界, 2024, (09): 114-116.

[4] 陈湘源, 秦伟, 刘晏驰, 等. 融合卷积神经网络与线性回归的带式输送机托辊故障音频识别方法 [J]. 煤炭科学技术, 2025, 53(S1): 389-398.

[5] 黄奇. 基于多模态信息融合的情感分析机制研究与实现 [D]. 南京邮电大学, 2022. DOI: 10.27251/d.cnki.gnjdc.2022.001457.

[6] 吕成双, 王彤, 孙浩然. 融合情绪指标的股价波动率预测研究 - 基于微调大语言模型与 GAT-TCN 网络 [J/OL]. 运筹与管理, 1-8[2025-12-09]. <https://link.cnki.net/urlid/34.1133.G3.20250106.1142.002>.

[7] 朱晓彤. 基于 Transformer 模型的金融时间序列特征分析与预测方法研究 [D]. 山东财经大学, 2025. DOI: 10.27274/d.cnki.gsdjc.2025.000832.

[8] 陈亦真, 宋旭峰, 李超顺. 基于 RFECV-RF 和 PCA-LightGBM 的水电机组劣化趋势评估 [J]. 水电能源科学, 2025, 43(03): 196-199+172. DOI: 10.20040/j.cnki.1000-7709.2025.20240361.

[9] 肖璐, 陈果, 刘继云. 基于情感分析的企业产品级竞争对手识别研究 - 以用户评论为数据源 [J]. 图书情报工作, 2016, 60(01): 83-90+97. DOI: 10.13266/j.issn.0252-3116.2016.01.012.

[10] 郭佳伦, 钟浩琨, 赵俊博, 等. 基于支持向量回归 (SVR) 的马尾松木材脱脂率预测 [J]. 北京林业大学学报, 2025, 47(03): 151-161.