

加入舆情信息是否可以有效提高汇率预测效果?

方思然¹, 郭明君¹, 魏云捷^{2,3}

(1. 中国科学院大学经济与管理学院, 北京 100190; 2. 中国科学院数学与系统科学研究院, 北京 100190;
3. 中国科学院预测科学研究中心, 北京 100190)

摘 要 随着信息技术的发展, 机构研报、财经新闻、搜索指数等多种舆情信息成为影响汇率变化的重要因素. 通过自然语言处理 (NLP) 技术处理非结构化舆情文本数据可以反映投资者行为、情绪和预期, 为汇率趋势预测提供数据基础. 本文基于 NLP 技术和深度学习的方法, 提出一种使用多种舆情信息预测人民币汇率的研究方法, 并首次使用研报对汇率进行预测. 本文收集了研报、新闻及搜索指数三类舆情信息, 首先, 使用 NLP 技术和基于词典的文本情感分析方法, 将研报和新闻舆情文本数据转化为可计算的专业舆情指数和社会舆情指数; 其次, 综合了人民币汇率的百度搜索指数和谷歌趋势, 构建了大众舆情指数; 最后, 设计了 8 组“汇率-舆情”预测组合, 使用长短期记忆网络 (LSTM) 模型对人民币汇率进行预测. 实证结果表明, 舆情信息可以预测汇率趋势; 三种舆情信息中, 新闻舆情对汇率的预测效果最好, 研报次之, 搜索指数最后; 加入多种舆情信息相比单一舆情信息, 预测效果显著提升; 此外, 相较于自回归综合移动平均 (ARIMA)、支持向量回归 (SVR) 和极限学习机 (ELM) 模型, LSTM 模型对于基于舆情的汇率预测效果最好.

关键词 汇率预测; 自然语言处理; 舆情信息; 情感分析

Is the Inclusion of Public Opinion Information Effective in Improving Exchange Rate Forecasting?

FANG Siran¹, GUO Mingjun¹, WEI Yunjie^{2,3}

(1. School of Economics and Management, University of Chinese Academy of Sciences, Beijing 100190, China;
2. Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China;
3. Center for Forecasting Science, Chinese Academy of Sciences, Beijing 100190, China)

收稿日期: 2023-01-17

基金项目: 国家自然科学基金 (72171223, 71801213, 71988101)

Supported by National Natural Science Foundation of China (72171223, 71801213, 71988101)

作者简介: 方思然, 博士研究生, 研究方向: 经济预测理论与方法, E-mail: fangsiran21@mailsucas.ac.cn;
郭明君, 硕士研究生, 研究方向: 艺术品金融, E-mail: guomingjun21@mailsucas.ac.cn; 通信作者: 魏云捷, 副研究员, 研究方向: 经济预测理论与方法, E-mail: weiyunjie@amss.ac.cn.

Abstract With the development of information technology, a variety of public opinion information such as institutional research reports, financial news, and search indexes have become important factors influencing exchange rate changes. By processing unstructured public opinion text data through natural language processing (NLP) technology, it can reflect investors' behavior, sentiment, and expectation, and provide a data base for exchange rate trend prediction. Based on NLP technology and deep learning, this paper proposes a research method to predict RMB exchange rate using various public opinion information, and for the first time uses research reports to forecast exchange rate. This paper collects three types of public opinion information: Research reports, news, and search indexes, and firstly, uses NLP technology and lexicon-based text sentiment analysis to transform research reports and news opinion text data into computable professional opinion indexes and social opinion indexes. Finally, eight sets of "exchange rate-public opinion" forecasting combinations were designed to forecast the RMB exchange rate using LSTM model. The empirical results show that public opinion information can predict exchange rate trends; among the three types of public opinion information, news opinion has the best prediction effect on exchange rate, followed by research reports and search indexes; adding multiple public opinion information significantly improves the prediction effect compared with single public opinion information; in addition, compared with ARIMA, support vector regression (SVR) and extreme learning machine (ELM) models, long and short-term memory network (LSTM) has the best prediction effect on exchange rate based on public opinion.

Keywords exchange rates forecasting; NLP; public opinion information; emotional analysis

1 引言

2022 年以来, 在全球通货膨胀飙升、美联储多次大幅度加息、俄乌冲突爆发、能源危机加剧等事件的冲击下, 全球金融市场动荡, 汇率也因此大幅波动。人民币汇率自八月以来迅速贬值, 在九月中旬离岸人民币汇率与在岸人民币汇率相继跌破“7”的大关。人民币汇率的巨幅波动对国民经济运行、国际贸易、对外投资等宏观经济调控都有着深刻的影响, 并可能造成国际收支不平衡, 引发通货膨胀压抑或激发全球经济的活力。鉴于目前国际金融市场波动频繁, 且经济全球化进展进一步加深, 国际贸易联系更加紧密, 在这样的背景下, 汇率在经济和金融中地位比以往更加重要, 对人民币汇率进行准确预测, 对于指导个人投资行为、帮助企业制定国际贸易决策、以及国家外汇储备和外汇风险防范等都有着重要的意义。

由于汇率数据具有高波动、高噪音、非平稳等特点, 如何准确地对其进行建模和预测成为一种挑战。目前, 在汇率建模和预测方法中, 一些广泛使用的模型包括, 早期常用的 ARIMA (自回归综合移动平均)、ECM (误差修正模型)、VAR (向量自回归) 和 GARCH (广义自回归条件异质模型) 等计量经济学模型; 后期, 如模糊集理论、ANN (Artificial Neural Network)、SVM (Support Vector Machine) 等智能优化算法和深度学习技术被广泛推广。近年来, 随着信息技术的发展和互联网的普及, 新闻及平台经济等微观因素成为影响汇率变化

的重要因素之一。通过自然语言处理 (NLP)、图像识别等技术得到的财经新闻、社交网站、网络评论等非结构化舆情文本数据和视频图像数据可以用来代表投资者情绪和预期,反映投资者的行为趋势,为金融市场波动预测提供数据基础。机构研报、财经新闻、搜索指数等多种舆情信息成为影响汇率变化的重要因素。

基于财经新闻、社交网站、网络评论等非结构化舆情文本数据对金融时间序列进行预测的文章很多,对股价(黄丽明等(2019),尹海员和吴兴颖(2019),许启发等(2017),姚加权等(2021),Oliveira et al. (2017),Feuerriegel et al. (2018),Jabeen et al. (2022))进行预测的文章尤其多。基于舆情信息对汇率进行预测的文章相对较少(孙少龙等(2022),任仙玲和邓磊(2019),张杰等(2021),王吉祥等(2019),Huang et al. (2022))。通过对现有文献的梳理后发现,基于舆情对汇率趋势进行预测的文章大多数使用的舆情信息都为财经新闻。例如,孙少龙等(2022)对新浪财经网的外汇新闻进行情感挖掘,使用外汇新闻对汇率进行预测。张杰等(2021)基于汇通网的外汇新闻构建了时间卷积网络模型对汇率进行预测和交易策略制定。王吉祥等(2019),成舟等(2019)都使用来自华尔街见闻的汇率新闻作为互联网舆情数据,基于多种预测模型对汇率进行预测。Semirami et al. (2020),Zhang et al. (2005)等也都挖掘新闻中的情感以对汇率进行预测。也有少量文献使用除新闻外的舆情数据对汇率进行预测。例如,应用百度指数作为舆情数据对汇率进行预测的文章有两篇(杨超等(2019),李佳等(2019))。任仙玲和邓磊(2019)使用微博平台的评论作为舆情数据研究中美贸易摩擦事件对人民币汇率的冲击效应。国外学者在使用新闻外的舆情数据时,大多采用社交媒体(如推特)的评论数据(Colonescu(2018),Yasar(2020))和谷歌趋势(Bulut(2018),Wilcoxson(2020))。也有少量文献将两种舆情数据结合,例如Crone and Koeppel(2014)使用新闻及社交媒体评论两种舆情数据构建情绪指标来解释汇率回报的市场波动;Rao and Srivastava(2013)采用推特评论情绪与谷歌搜索量指数两种舆情数据来建立外汇的预测模型。

从上述文献可以看出,绝大部分文献都采用财经新闻作为舆情数据对汇率进行预测,少部分文献采用搜索引擎的搜索指数和网络平台的评论作为舆情数据。现有文献在对汇率基于舆情进行预测时,存在以下几个问题:

1) 舆情数据种类和来源单一。舆情信息可以分为多个类型,如有来自社交媒体、财经新闻类的社会舆情信息,来自各大证券公司的机构研报的专业舆情信息,能够反映投资者关注度的大众舆情信息(如百度搜索指数、微博评论)等。现有文献目前大多使用单一类型舆情作为舆情信息,有少数文献会采用两种舆情数据,但每种舆情数据也都为同信息源的舆情数据。虽然成舟等(2019)的文章里将华尔街见闻的汇率新闻分类为了金融市场舆情数据和宏观经济舆情数据等分级的舆情数据,但舆情信息的来源仍都为同一财经网站的新闻类舆情。

2) 没有使用研报作为舆情信息对汇率进行预测的文章。研究报告属于反映机构、专家意见的一种专业类舆情信息。目前,各大证券公司、基金公司、银行投资部门等都会出具汇率相关的研报,为企业和投资者提供相关的专业资讯。与财经新闻、平台评论、搜索指数等舆情信息相比,研报所提供的舆情信息更加专业、客观,并具有很强的权威性。因此,研报对企业和投资者的行为也会产生潜移默化的影响,是一种很重要的舆情信息。然而现有文献里,仅有少量文章使用研报对股价进行预测(吴武清等(2020),饶东宁等(2017)),没有使用研报作为舆情信息对汇率进行预测的文章。

3) 现有文献中没有文章对比不同类型的舆情信息对汇率的预测效果的差别, 如新闻类舆情数据、微博评论类舆情数据、搜索指数类舆情数据等, 不同种类的舆情数据对于市场的影响程度、时间、对象都有所不同. 因此, 不同种类的舆情信息对汇率预测精度影响程度也许不同, 哪种类型的舆情数据对汇率的预测效果更好值得探究.

4) 现有文献中没有文章考虑是否采用多种类型舆情信息, 综合多方位舆情信息, 相较于单一类型舆情信息, 能够更全面且系统地刻画汇率的波动, 对汇率的预测效果更好.

因此, 本文创新性地从纳入多方位舆情信息的角度出发, 构建了一个基于多种舆情信息的汇率预测方法, 并首次使用研报对汇率进行预测. 我们首先将舆情信息分为三种: 来自社交媒体、财经新闻类的社会舆情信息, 来自各大证券公司机构研报的专业舆情信息, 能够反映投资者关注度的大众舆情信息 (如百度搜索指数、谷歌趋势). 其次, 使用 NLP 技术和基于词典的文本情感分析方法, 将舆情数据中的文本数据转化为可计算的舆情指数; 最后, 设计了 8 组“汇率-舆情”预测组合, 使用 LSTM 模型对人民币汇率进行预测. 本文通过首次纳入多方位的舆情信息, 对比不同类型、单种及多种舆情信息的预测能力, 为提高汇率预测精度提供了一种新的视角和思路, 弥补了目前基于多种舆情对汇率进行预测的文献的空白, 对帮助市场从多种舆情信息角度更好地理解外汇市场价格的形成机制, 对相关政策制定者和市场参与者做决策有着重要的意义. 本文主要探究了以下问题: 1) 舆情信息能否对汇率趋势进行预测; 2) 三类舆情信息中 (专业舆情、社会舆情、大众舆情), 哪一类舆情信息对汇率的影响更大, 对汇率的预测效果更好; 3) 相对单一舆情信息, 是否采用多种舆情信息对汇率的预测效果更好.

本文剩余部分结构安排如下: 第二章设计了一种基于多种舆情数据的汇率预测研究方法, 其中详细介绍了舆情信息的选取、收集、预处理、如何通过文本数据情感分析将文本数据转化为可计算的情感指数, 以及对本文使用的汇率预测模型 LSTM 进行了介绍; 第三章构建了“汇率-舆情”的预测组合, 对基于多种舆情数据的汇率预测结果进行了分析; 第四章对文章进行了总结.

2 研究设计

本文提出一种基于多种舆情信息对人民币汇率趋势进行预测的研究方法, 主要分为四部分, 研究方法流程图如图 1 所示.

1) 数据提取. 数据的提取包括历史人民币汇率数据及三种类型舆情数据. 从数据的全面性、专业性以及时效性角度考虑, 本文的三类舆情信息, 专业舆情信息来自东方财富网的宏观研报板块、社会舆情数据来自于 Wind 数据库, 宏观舆情主表中汇率相关的新闻、大众舆情数据选取了关键词“人民币”“人民币汇率”的百度搜索指数, 及“RMB”“Chinese Yuan”谷歌趋势的搜索量. 汇率的数据选取美元兑人民币的汇率日度数据, 数据来源于 Wind 数据库. 四种数据中, 汇率历史数据与指数类舆情数据为结构化数据, 研报及新闻类舆情数据为非结构化数据.

2) 文本数据预处理. 舆情数据中, 研报及新闻类舆情数据为非结构化文本数据, 在进行数据计算前, 须先采用 NLP 技术对其进行文本数据预处理, 包括过滤、合并、分词、去停用词等步骤, 使其转化为由词汇向量构成的矩阵, 以便后续进行文本数据情感分析.

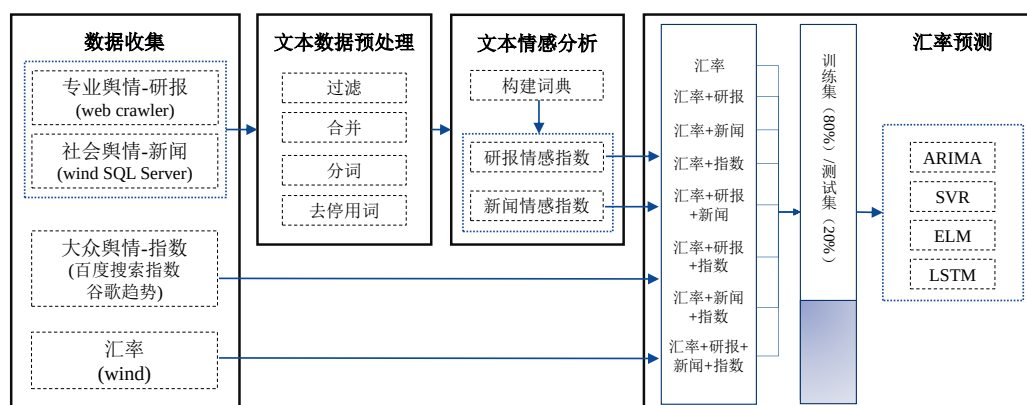


图 1 研究方法流程图

3) 文本情感分析. 本文采用基于语义, 与词典匹配的情感分析方法对研报及新闻文本数据进行分析. 情感分析主要分为两步, 首先构建一个情感词典, 其次将预处理后的文本数据与词典进行匹配, 提取出文本数据中的情感词汇及修饰词, 计算各类舆情的感情指数, 将文本数据转化为计算机可以处理的二维数据.

4) 汇率预测. 本文基于多种舆情数据对汇率进行预测, 因此, 美元对人民币的汇率为因变量, 研报、新闻及指数类舆情数据的感情指数为自变量, 构成 8 组“汇率-舆情”预测组合. 我们选取了以深度学习方法为基础的 LSTM 为主要预测模型, 预测人民币汇率的趋势. 最后, 使用 MAE、RMSE、MAPE 和 DS 作为预测误差的评价指标对模型的预测效果进行对比和综合评估.

2.1 舆情数据收集

本文收集的数据分为两种: 舆情信息数据和历史人民币汇率数据, 其中舆情信息数据又能分为三类, 代表专业意见的研报类舆情数据, 代表社会焦点的新闻类舆情数据, 以及能够反映投资者关注度的指数类舆情数据 (如百度搜索指数). 对于不同类型数据, 具体的数据收集方法如下.

2.1.1 专业舆情数据-研报

研究报告属于专业舆情信息的一种. 目前, 各大证券公司、基金公司、银行投资部门等都会出具汇率相关的研报, 为企业和投资者提供相关的专业资讯. 与新闻和指数类舆情信息相比, 研报所提供的舆情信息更加专业、客观, 并具有很强的权威性, 因此, 对企业和投资者的行为也会产生潜移默化的影响.

本文从研报的全面性、专业性以及时效性的标准考虑, 选择了东方财富网的宏观研报作为专业舆情数据. 东方财富网是国内用户活跃度高、财经资讯专业度高、覆盖范围广的权威性财经类网站. 东方财富网的宏观研报板块提供包括中银证券、平安证券、国泰期货等各大券商的研究报告, 覆盖汇率、货币、财政、国际金融等多方面的宏观基本面信息, 保证了专业舆情数据来源的全面性和专业性. 且东方财富网在每个工作日都会更新大量的宏观研报, 保证了专业舆情数据的时效性.

本文通过 python 语言、自动化测试库 Selenium 及爬虫框架 Scrapy 实现了对东方财富网的宏观研报的爬取. 具体爬取过程为, 首先通过自动化测试库 Selenium 爬取宏观研报板块每一篇研报的报告名称, 发布时间及 URL, 再根据每一篇研报的 URL 通过爬虫框架 Scrapy 对研报正文内容的摘要进行爬取. 东方财富网对于宏观研报提供的最长期限为两年内的研报, 再过滤掉一些文本内容较少的研报, 最终本文一共爬取了 2020 年 11 月 2 日至 2022 年 9 月 20 日的 15042 篇宏观研报, 爬取到的研报数据示例如表 1 所示. 本文仅使用宏观研报的摘要部分作为专业舆情数据, 因为摘要部分包含了研报足够的键信息, 相比研报全文包含更少的重复和不相关的词, 且更容易检索和爬取.

表 1 研报数据示例			
时间	标题	正文摘要	来源
2022/9/20	【资产配置系列】央行下调外汇存款准备金率	8 月中旬至今, 人民币汇率出现了大幅走贬, 从 6.74 迅速贬值至 6.89……	中泰证券
2022/9/20	英大宏观评论 (2022 年第 32 期, 总第 116 期): 人民币汇率破 “7” 点评: 央行大概率重启逆周期因子, 人民币汇率反转时点有可能在 2023 年春季	9 月 15 日离岸人民币汇率破 “7”; 9 月 16 日在岸人民币汇率盘中破 “7” ……	英大证券
2022/9/7	降低外汇存款准备金率象征意义较强, 后续人民币走势将取决于经济数据表现	央行年内再次下调外汇存款准备金, 应对人民币贬值压力……	中航证券
2022/9/6	流动性洞见系列十: 人民币汇率急贬, 未来怎么走?	核心观点: 今年以来, 中美经济和通胀周期错位导致政策周期错位……	光大证券
2022/8/12	宏观大类日报: 人民币汇率小幅走强关注对 A 股的支撑	近期人民币汇率小幅走强, 从 2014 年至今的历史表现来看, 汇率走强对 A 股将提供一定支撑……	华泰期货
...	...	...	...

2.1.2 社会舆情数据-新闻

来自财经新闻、社交媒体的新闻类舆情数据相比研报类专业舆情数据和大众类舆情数据, 更能反映社会关注焦点. 财经新闻网站每天都会更新最新的金融财经新闻, 提供更具时效性的舆情数据. 目前, 互联网中有许多提供汇率新闻专栏的财经网站, 但这些网站大都只能提供最长期限为几个月的汇率新闻 (如新浪财经、网易财经、东方财富网). 因此, 难以在财经网站获取一年以上的汇率新闻.

最终我们选取 Wind 新闻数据库的宏观舆情主表中汇率相关的新闻作为社会舆情数据. 选取该数据库的原因是, Wind 新闻数据库的宏观舆情主表提供来自 Wind、网易、财联社、第一财经等多个专业财经网站的两年以上的每日汇率新闻, 保证了社会类舆情数据的专业性高, 覆盖度广以及时效性强.

Wind 新闻库数据存放于 SQL Server 数据库中, 本文下载了宏观舆情主表中从 2020 年 11 月 2 日至 2022 年 9 月 20 日汇率相关的新闻共 1945 篇, 爬取到的新闻数据示例如表 2 所示. 本文仅使用新闻的标题作为社会类舆情数据, 因为新闻标题包含了简洁、充足的关键信息, 相比新闻全文后续更容易进行准确的文本情感分析.

表 2 新闻数据示例

时间	标题	来源
2022/7/14	中国 6 月末央行外汇占款 21.3 万亿元人民币, 环比减少 41.1……	Wind
2022/7/13	我国外汇市场韧性显著增强更有基础和条件抵御外部冲击风险	中国金融新闻网
2022/7/12	12 日人民币对美元汇率中间价 6.7287 元下调 327 个基点	中国新闻网
2022/7/10	人民币周报: 美元指数一度创新高, 人民币汇率总体保持稳定	Wind
2022/7/7	外汇局王春英: 6 月外汇市场运行平稳境内外汇供求延续基本平衡	顶尖财经网
2022/7/5	人行中卫市中支: 推进跨境人民币结算助力涉外企业发展	中国金融新闻网
2022/7/4	债券通五周年, 境外持债规模达 3.66 万亿, 人民币债券吸引力强劲	Wind
...	...	...

2.1.3 大众舆情数据-指数

通过搜索引擎搜索关键词获得的搜索指数是反映大众对事物关注度的重要指标, 搜索量的变化体现了大众的关注热点, 对于汇率等经济数据的预测具有很好的前瞻性. 因此, 本文选取国内及国外两个用户数最多, 搜索量最大的搜索引擎的搜索指数, 百度搜索指数及谷歌趋势作为大众舆情数据, 以反映大众投资者对汇率的关注度.

本文选取了“人民币”及“人民币汇率”作为百度搜索指数的搜索关键词, “RMB”和“Chinese Yuan”作为谷歌趋势的搜索关键词, 获取了 2020 年 11 月 2 日至 2022 年 9 月 20 日关于人民币汇率的百度搜索指数和谷歌趋势.

2.2 文本数据处理

在舆情信息数据中, 研报和新闻舆情数据属于文本数据, 为后续对文本数据进行情感分析, 需要先对文本数据进行预处理. 本文采用自然语言处理 (NLP) 技术对文本数据进行过滤、合并、分词、去停用词、构建词典及情感分析等处理工作, 具体流程如图 2 所示.

2.2.1 文本数据过滤、合并及分词

1) 过滤

文本数据进行预处理的第一步为对收集到的文本数据进行过滤. 收集到的文本数据中可能存在一些网页链接、图片等非有效文本数据, 需要将这些文本数据过滤掉. 对于研报文本数据, 本文收集到的研报为宏观研报, 其中包括汇率、货币、财政、国际金融等多方面

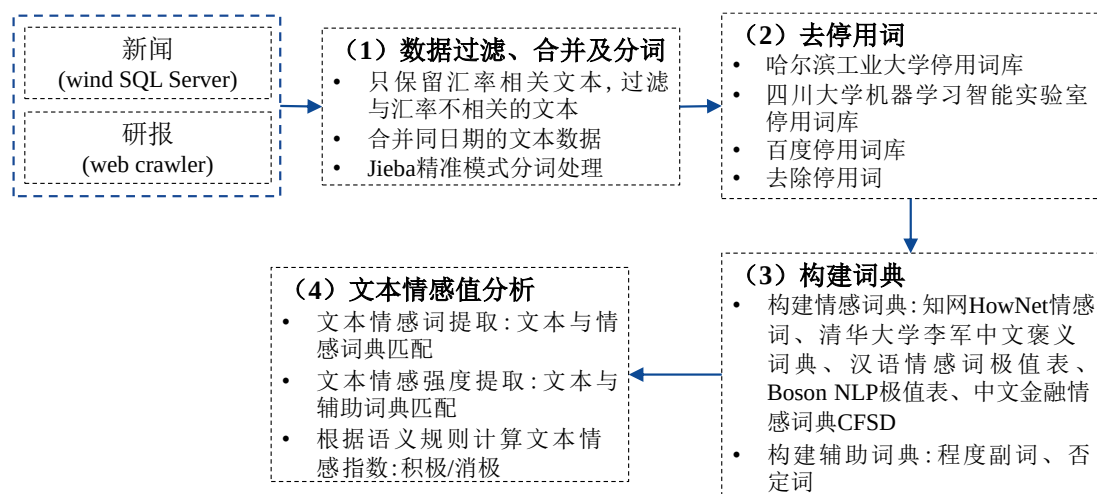


图2 文本处理流程图

的宏观基本面信息。而本文对汇率进行预测, 因此对研报文本数据通过匹配“人民币汇率”“USD/CNY”等相关的汇率关键词, 只保留与汇率相关的文本, 过滤掉与汇率不相关的文本。

2) 合并

文本数据预处理的第二步为合并文本数据。本文收集到2020年11月2日至2022年9月20日, 除休息日外, 共461天的研报及新闻文本数据, 其中研报共15042篇, 每天平均收集到研报约33篇, 新闻共1945篇, 平均每天约4篇。本文欲对人民币汇率进行日度预测, 需要获取日度的舆情情感指数。因此, 须将每日收集的研报及新闻文本数据合并, 得到461篇研报文本数据及461篇新闻文本数据。

3) 分词

在获取日度文本数据后, 须对文本数据进行分词处理。基于词典的分词方法, 本文采用jieba分词器, 使用精准模式对文本数据精确分词, 将文本数据转化为由词汇向量构成的句子矩阵, 使研报及新闻文本数据表示为计算机可处理的形式。

2.2.2 去停用词

停用词指文本中一些出现频率高, 但对文本情感分析没有帮助的词语, 例如“了”“的”, 人名, 时间等词汇。因此在对文本数据进行情感分析前, 先将文本数据中的停用词去掉, 可以大幅提高文本情感分析的效率。本文使用哈尔滨工业大学停用词库、四川大学机器学习智能实验室词库, 和百度停用词库, 构建一个全面的停用词库。将分词后的文本数据与停用词词典相匹配, 去除文本数据中的停用词(本文构建的停用词库示例如表3所示)。

2.3 文本数据情感分析

本文采用基于语义, 与词典进行匹配的情感分析技术对文本数据进行分析。因此, 对研报及新闻的情感分析主要分为两步, 1) 构建一个词典; 2) 将预处理后的文本数据与词典进行匹配, 提取出文本数据中的情感词汇及修饰词, 计算情感指数。

2.3.1 构建词典

为进行情感分析,首先需要构建一个词典,词典中文本的词性主要分为情感词及辅助词,情感词分为:积极情感词和消极情感词;辅助词分为:程度级别词汇(程度副词)和否定词.本文所建立的情感词典示例如表3所示,对情感词及辅助词的具体解释如下:

1) 情感词

情感词分为积极情感词及消极情感词,积极情感词如“喜欢”“痴迷”“爱护”等,消极情感词如“忧愁”“败坏”“下流”等.本文采用的情感词词典为目前被广泛应用的知网 HowNet 情感词典、清华大学李军中文褒义词典和汉语情感词极值表,这三个词典较为齐全的涵括了中文的情感词汇.除基础的中文词典外,因为本文研究的是外汇相关的文本,因此在上述词典的基础上添加了中文金融情感词典 CFSD.该词典包含积极的外汇情感词如“坚挺”“缓和”“看涨”等;消极情感词如“下跌”“加息”“看跌”等.最终,本文结合中文金融情感词典和基础中文情感词典构建了较为全面的外汇预测情感词典,其中积极情感词共 7186 个,消极情感词共 12085 个.

2) 辅助词

辅助词主要包括两种:程度副词和否定词.两种词汇的具体解释如下:

程度副词是指影响文本情感倾向程度的词语,例如“非常”“加倍”“明显”等.本文采用知网 HowNet 情感词典,共包括 221 个程度级别词汇,并且对情感词汇划分了 4 个级别,即:最强烈的 (mostdict)、次强烈的 (verydict)、一般强烈的 (moredict)、弱强烈的 (ishdict),程度副词的级别依次下降.

否定词指改变文本情感倾向的词汇,例如在“喜欢”前加上否定词“不”,语义变为“不喜欢”,将喜欢的含义完全调转,因此识别文本中的否定词汇非常重要,忽略如“不”“非”“没”这样的否定词汇会完全扭转文本的情感倾向.本文所建立的否定词词典共 15 个词语.

表 3 外汇预测词典

情感词		辅助词		停用词
积极情感词	消极情感词	程度级别词汇	否定词	
喜爱	极度痛苦	非常	不是	的
敬礼	猛撞	极	不能	了
荣耀	苦味	极度	不可	得
纯洁	不稳定	极端	没有	!
文雅	不甘心	极其	不要	(
奇妙	诟病	万	勿	)
坚挺	严重	万般	休	#
缓和	新低	万分	甬	/
缓解	下跌	万万	非	,
好转	加息	无比	没	*
...	...	...	...	...

2.3.2 文本情感指数计算

在词典建立好后, 本文采用基于语义, 与词典进行匹配的情感分析技术对文本数据进行分析. 将预处理后的文本数据与词典进行匹配, 提取出文本数据中的情感词汇及修饰词, 计算情感指数. 具体的文本情感指数计算步骤如下所示 (文本情感值计算步骤流程图如图 3 所示):

1) 将预处理过的研报加载到构造好的词典中, 对一篇研报和新闻中的每个句子在词典中进行匹配, 并设定一个积极感情初始值 poscount , 令 $\text{poscount}=0$ (每一个积极情感词都会有一个 poscount , 一句话中若有多个积极情感词, 则有多个 poscount); 设定一个消极情感初始值 negcount , 令 $\text{negcount}=0$ (每一个消极情感词都会有一个 negcount , 一句话中若有多个消极情感词, 则有多个 negcount). 如果在该句中检测到情感词或感叹号 “!”, 则进行到下一步骤, 若未检测到情感词, 则跳过此句, 对下一句进行匹配.

2) 若在 1) 中匹配到情感词, 当该情感词为积极情感词汇时, 令 $\text{poscount}+1$; 检查该情感词的前后词汇:

- 若前面词汇为程度级别词汇, 根据程度副词的级别, 对 poscount 放大不同的系数. 程度副词级别越高, 放大的系数越大.
- 若前面词汇为否定词, 当否定词个数为奇数时, 对 $\text{poscount}*(-1)$, 否则不对 poscount 进行修改.

3) 若在 1) 中匹配到情感词, 当该情感词为消极情感词汇时, 令 $\text{negcount}+1$; 检查该情

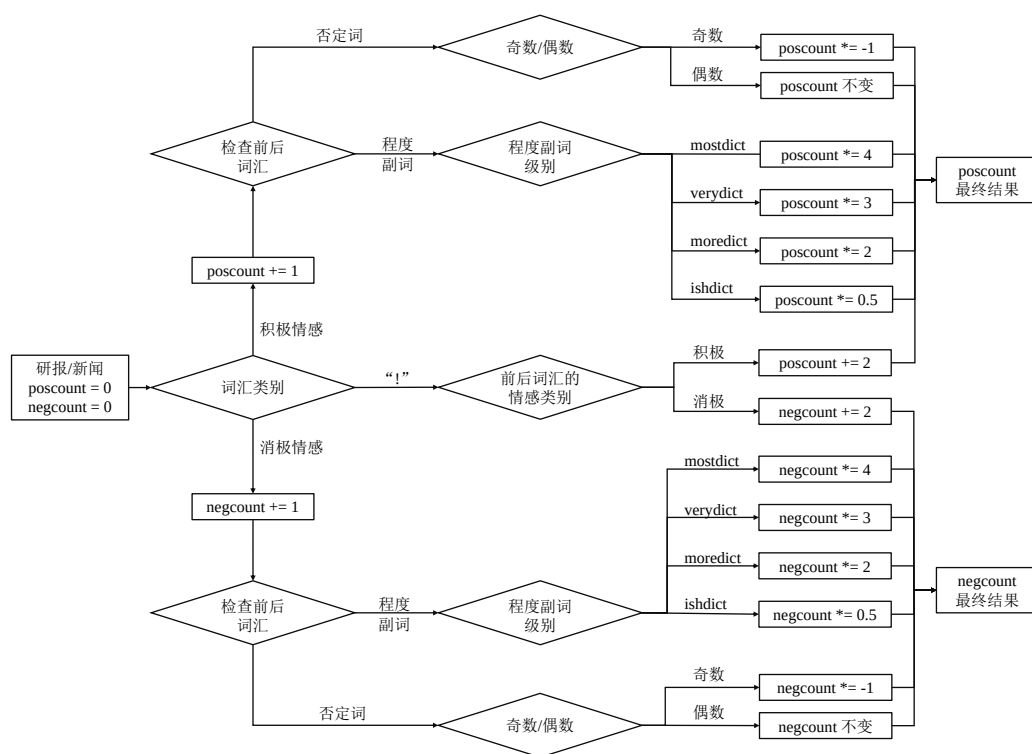


图 3 文本情感值计算步骤流程图

感词的前后词汇:

- 若前面词汇为程度级别词汇, 根据程度副词的级别, 对 negcount 放大不同的系数. 程度副词级别越高, 放大的系数越大.
- 若前面词汇为否定词, 当否定词个数为奇数时, 对 $\text{negcount} * (-1)$, 否则不对 negcount 进行修改.

4) 若在 1) 中匹配到感叹号 “!”, 当 “!” 出现在积极情感词存在的句子中时, 该句的 $\text{poscount}+2$, 当 “!” 出现在消极情感词存在的句子中时, 则对该句的 $\text{negcount}+2$.

- 将每日的文本数据积极情感值 poscount 和消极情感值 negcount 加总, 得到每日总积极情感值 Pos 和总消极情感值 Neg ;
- 将 Pos 和 Neg 根据每日的文本句数取平均得到平均积极情感值 AvgPos 和平均消极情感值 AvgNeg ;
- 将 Pos 和 Neg 根据每日的文本句数取标准差得到积极情感标准差 StdPos 和消极情感标准差 StdNeg .

通过上述四步计算后, 可以得到每日文本数据的积极情感值和消极情感值, 文本情感分析效果示例见表 4. 考虑每日研报及新闻的数量也可以反映舆情信息, 研报及新闻数量越多时, 代表这段时间舆情对汇率的关注度更高. 因此本文主要采用每日总积极情感值 Pos 和总消极情感值 Neg 计算每日文本数据情感指数, 以更好地传达文本数据的情感. 参考孙少龙等 (2022) 年的论文, 如下计算第 t 日的文本数据 (研报或新闻) 的情感指数 Sentiment_t :

$$\text{Sentiment}_t = \ln \left(\frac{1 + \text{Pos}_t}{1 + \text{Neg}_t} \right). \quad (1)$$

其中, Pos_t 为第 t 日的文本数据 (研报或新闻) 总积极情感值, Neg_t 为第 t 日总消极情感值.

从表 4 可以看出, 本文构建的文本数据情感挖掘基本准确, 例如, 表 4 展示的第 1 条、第 5 条语句, 文本内容 “加强” “有效防范” “韧性” “有效增强” “保持” 等词汇都呈现出较为积极的情感, 且这两条语句中没有消极的情感词汇, 因此相应地计算出的情感分数, 总积极情感值 Pos , 平均积极情感值 AvgPos , 积极情感值标准差 StdPos 都为正值, 总消极情感值 Neg , 平均消极情感值 AvgNeg , 消极情感值标准差 StdNeg 都为 0.

与其相对地, 第 2 条语句, 文本内容 “通胀” “紧缩” “贬值” 等词汇都呈现出较为消极的情感, 且这两条语句中没有积极的情感词汇, 因此相应地计算出的情感分数, 总消极情感值 Neg , 平均消极情感值 AvgNeg , 消极情感值标准差 StdNeg 都为正值, 总积极情感值 Pos , 平均积极情感值 AvgPos , 积极情感值标准差 StdPos 都为 0. 从表 4 所展示的示例可以看出, 本文所构建的情感分析算法基本合理, 可以较为准确地反映外汇研报及新闻文本情感倾向.

2.4 长短期记忆神经网络 (LSTM)

本文选取了以机器学习方法为基础的长短期记忆网络模型 (LSTM) 为主要预测模型. LSTM 是由时间循环神经网络 (RNN) 改进而来的, 最早由 Hochreiter and Schmidhuber (1997) 提出. LSTM 模型通过引入记忆单元发掘长期依赖关系, 可以克服循环神经网络中梯度爆炸和梯度消失的问题 (倪铮和梁萍 (2018)), 因此适用于时间间隔较长或者有延迟特性的时间序列预测.

表 4 文本情感分析效果示例

第 n 句	句子内容	Pos	Neg	AvgPos	AvgNeg	StdPos	StdNeg
1	外汇局将加强外债形势监测和分析, 有效防范可能出现的风险.	6	0	0.4	0	0.4899	0
2	面对多年来少见的通胀, 全球大部分经济体加快货币紧缩步伐, 新兴市场面临资本外流、汇率贬值的压力.	0	4	0	0.16	0	0.463
3	人民币对美元的汇率出现了一定程度贬值, 但是与国际主要货币相比, 人民币币值的稳定性仍然较强.	3	15	0.125	0.625	0.3307	0.4841
4	从汇率预期来看, 外汇远期和期权相关指标显示, 人民币汇率没有明显的升值和贬值预期, 市场主体总体保持理性有序的交易模式.	20	27	0.6452	0.871	0.7426	0.9067
5	在我国外贸保持韧性、人民币资产吸引力强、跨境资本双向有序流动、人民币汇率弹性显著增强等因素支撑下, 人民币汇率仍有望保持在均衡水平附近, 双向波动常态化.	57	0	1.3902	0	1.3412	0

RNN 的核心思想是将同一神经网络进行多次复制, 通过不同神经网络模块的信息传递刻画非线性时间序列的特征. 但目前 RNN 有如下缺陷: 一方面, RNN 存在梯度消失与梯度爆炸的问题, 从而不能很好地刻画有延迟特征的时间序列; 另一方面, 延迟窗口长度需要在训练模型之前确定下来, 但是在实践中难以自动找到最优的窗口值. 这些缺陷使得 RNN 的实用性减弱.

LSTM 则克服了 RNN 的主要缺陷. LSTM 基于 RNN 发展出了含一个记忆单元和三个门的神经元的模型细胞结构, 具体包括记忆单元 c 、输入门 i 、输出门 o 、遗忘门 f .

输入门 i 包括两个部分, 首先是由 sigmoid 函数决定输入的值, 接着通过 tanh 函数将新创建的候选值向量加入到状态 c_t , 公式如下:

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i).$$

(2)

遗忘门 f 的功能是读取上一层输出 h_{t-1} 和当前的输入 x_t , 并返回一个在 0、1 之间的数值 f_t , 并将值储存在当前细胞状态中, 公式如下:

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f).$$

(3)

更新门功能是更新旧细胞的状态, 由 c_{t-1} 更新至 c_t , 公式如下:

$$c_t = f_t \cdot c_{t-1} + i_t \tanh(W_c \cdot [h_{t-1}, x_t] + b_c).$$

(4)

输出门 o 的输出内容基于过滤后的细胞状态进行输出. 首先, 使用 sigmoid 函数确定将要输出细胞中哪个部分的状态, 其次, 将细胞状态转换成 -1、1 之间的值, 并与输出部分相

乘, 最终得到输出数值, 公式如下:

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_{t-1}] + b_o), \quad (5)$$

$$h_t = o_t \cdot \tanh(c_t). \quad (6)$$

x_t 为 t 时刻的输入向量, W_i 、 W_f 、 W_c 、 W_o 为循环层的权重矩阵, b_i 、 b_f 、 b_c 、 b_o 为偏执向量, h_t 为 LSTM 层输出向量.

LSTM 神经元结构如图 4 所示 (Graves (2012)). 记忆单元用于储存神经元状态, 输入门以及输出门用于接受、输出并修正参数, 遗忘门用于移除上一单元状态中的无用历史信息.

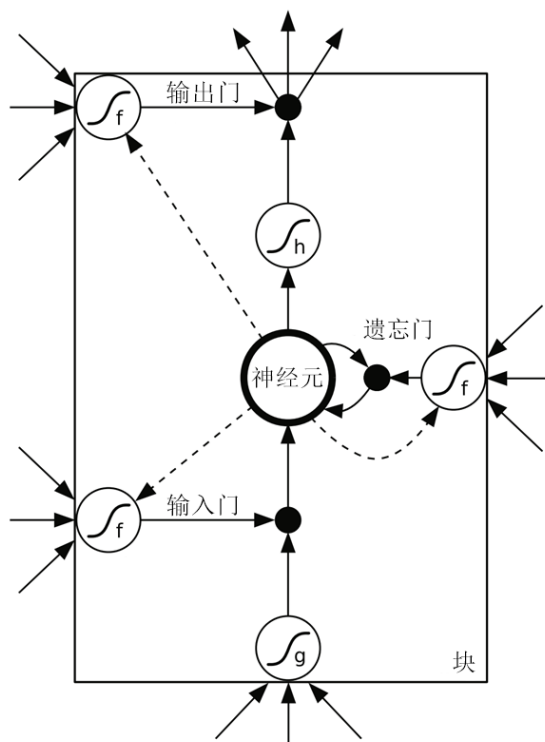


图 4 LSTM 神经元结构图 (Graves (2012))

2.5 评价指标

本文采取了“水平”和“方向”两个方面作为预测模型性能的评估标准. 水平预测精度评价指标包括平均绝对误差 (MAE)、均方根误差 (RMSE)、平均绝对百分比误差 (MAPE). 其中, MAE、RMSE 和 MAPE 的值越小, 代表预测的误差越小, 即预测效果越好. 水平预测精度评价指标公式如下:

$$\text{MAE} = \frac{\sum_{i=1}^m |y_i - \hat{y}_i|}{m}, \quad (7)$$

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^m (y_i - \hat{y}_i)^2}{m}}, \quad (8)$$

$$\text{MAPE} = \frac{100\%}{m} \sum_{i=1}^m \left| \frac{y_i - \hat{y}_i}{y_i} \right|. \quad (9)$$

至于方向预测精度评价指标, 本文选取了方向对称性 (DS). DS 值越大, 代表预测方向越准确, 即预测效果越好. 方向预测精度评价指标公式如下:

$$\text{DS} = \frac{100\%}{m} \sum_{i=1}^m d_i, \quad \text{where } D_i = \begin{cases} 1, & \text{if } (y_i - y_{i-1})(\hat{y}_i - y_{i-1}) \geq 0, \\ 0, & \text{otherwise.} \end{cases} \quad (10)$$

其中, m 为样本数量, y 为真实值, $\hat{y}$ 为预测值.

随后, 本文选取 DM 统计量来进一步验证模型的有效性. DM 检验假设两个模型对序列波动的预测能力相同, 因此本文使用 DM 统计量来检测不同模型的水平预测能力 (Chen et al. (2015), 李艳丽等 (2016)).

3 实证研究

本章通过实证研究, 探究以下 4 个问题: 1) 验证舆情信息能否预测汇率的趋势; 2) 探究三类舆情信息中 (专业舆情、社会舆情、大众舆情), 哪一类舆情信息对汇率的影响更大, 对汇率的预测效果更好; 3) 验证相对单一舆情信息, 是否采用多种舆情信息对汇率的预测效果更好; 4) 验证所提出预测模型的预测效果.

为展示其预测效果, 本文选取了 LSTM 为主要预测模型, 并根据以往研究中预测模型的表现, 选取了其他三种预测表现优良的模型作为基准模型, 包括 ARIMA、SVR、ELM, 并采用 MAE、RMSE、MAPE 和 DS 作为预测误差的评价指标对模型的预测效果进行对比和评估.

3.1 数据来源与处理

对于数据选取, 本文基于舆情信息对人民币汇率进行预测, 汇率数据选取 2020 年 11 月 2 日到 2022 年 9 月 20 日的美元兑人民币的汇率日度数据, 数据来源于 Wind 数据库. 同时本文还选取了三种舆情数据, 专业舆情信息来自东方财富网的宏观研报板块, 通过 Python 语言, 自动化测试库 Selenium 及爬虫框架 Scrapy, 共爬取了 15042 篇宏观研报的摘要; 社会舆情数据来自于 Wind 数据库, 宏观舆情主表中汇率相关的新闻, 共收集了 1945 篇新闻的标题. 大众舆情数据选取了关键词“人民币”“人民币汇率”的百度搜索指数, 及“RMB”“Chinese Yuan”谷歌趋势的搜索量. 具体数据来源如表 5 所示.

在获取所有数据后, 以美元兑人民币的历史汇率数据为标准, 将舆情信息数据中周六、周日以及节假日的数据加总到休息日后工作日的第一天, 使舆情数据与汇率数据时间一致.

在三种舆情数据中, 专业舆情数据研报及社会舆情数据新闻皆为文本数据. 本文对文本数据采用第三章中所述的预处理及情感分析方法, 将研报及新闻文本数据转化为研报情感指数 $\text{Ln}(\text{Report emotion})$ 和新闻情感指数 $\text{Ln}(\text{News emotion})$. 对于大众舆情数据, 本文将关键词“人民币”“人民币汇率”的百度搜索指数, 及“RMB”“Chinese Yuan”谷歌趋势的搜索热度, 四个关键词的搜索热度分别归一化后相加取自然对数, 获得搜索引擎指数 $\text{Ln}(\text{Indicator})$.

表 5 数据来源

舆情分类	数据来源	时间	数量
人民币汇率	Wind 金融终端 (USD/CNY)	2020/11/02–2022/09/20	461 天
专业舆情	东方财富网-宏观研报	2020/11/02–2022/09/20	15042 篇/ 550 天
社会舆情	Wind 新闻库-宏观舆情主表-汇率相关新闻	2020/11/02–2022/09/20	1945 篇/541 天
大众舆情	百度搜索指数 “人民币” “人民币汇率” 谷歌趋势 “RMB” “Chinese Yuan”	2020/11/02–2022/09/20	/

emotion) 所代表的大众舆情数据. 对于汇率数据, 本文对美元兑人民币的汇率取自然对数, 获得用于预测的历史汇率数据 $\ln(\text{USD/CNY})$.

3.2 数据描述性统计与检验

预测前, 须对数据进行检验. 数据的描述性统计见表 6. 表 6 中所有数据集都表现出过度的峰度, 这表明所有数据集都不遵循正态分布. Jarque-Bera 检验进一步证实了这一结果, 该检验在 10% 和 1% 的显著性水平上拒绝了正态分布的零假设. 本文通过 Dickey and Fuller (1979) 设计的 Dickey-Fuller (ADF) 单位根检验验证数据的平稳性, 结果如表 7 所示. 经检验发现, $\ln(\text{Report emotion})$ 和 $\ln(\text{News emotion})$ 通过 1% 的显著性水平的检验, 为平稳的时间序列. 而 $\ln(\text{USD/CNY})$ 和 $\ln(\text{Indicator emotion})$ 不平稳, 在对这两个序列进行一阶差分后, 再进行检验, 显示两个序列皆通过 1% 的显著性水平的检验, 为平稳的时间序列.

为检验舆情数据对汇率进行预测是否可行, 本文分别对汇率数据与三种舆情数据进行了 Johansen 协整检验, 检验结果如表 8 所示. 协整检验结果显示将所有数据进行一阶差分后, $D(\ln(\text{USD/CNY}))$ 与 $D(\ln(\text{Report emotion}))$, $D(\ln(\text{USD/CNY}))$ 与 $D(\ln(\text{News emotion}))$, $D(\ln(\text{USD/CNY}))$ 与 $D(\ln(\text{Indicator emotion}))$ 之间都是协整的, 表明人民币汇率与三种舆情数据之间存在长期协整关系. 因此, 从计量经济学的角度来看, 使用三种舆情数据对人民币汇率进行预测是可行的.

表 6 描述性统计

	$\ln(\text{USD/CNY})$	$\ln(\text{Report emotion})$	$\ln(\text{News emotion})$	$\ln(\text{Indicator emotion})$
均值	1.874	0.904	1.304	0.314
中位数	1.867	0.86	1.285	0.292
最大值	1.948	6.506	6.889	0.717
最小值	1.842	-4.673	-5.402	0.018
标准差	0.023	1.506	2.232	0.109
偏度	1.049	0.276	-0.231	0.377
峰度	3.473	5.314	2.832	3.977
Jarque-Bera	88.822***	108.701***	4.637*	29.268***
样本数	461	461	461	461

注: *** 表示在 1% 的显著性水平上拒绝原假设, * 表示在 10% 的显著性水平上拒绝原假设.

表 7 ADF 单位根检验

变量	Augmented Dickey-Fuller t-Statistic		Critical values (1% level)
	Level	1st difference	
Ln(USD/CNY)	0.851	-20.364 ***	-2.57
Ln(Report emotion)	-6.179 ***		-2.57
Ln(News emotion)	-3.164 ***		-2.57
Ln(Indicator emotion)	-1.176	-15.284 ***	-2.57

注: *** 表示在 1% 的显著性水平上拒绝原假设.

表 8 协整检验

Cointegration between D(Ln(USD/CNY)) and D(Ln(Report Emotion))				
	Eigenvalue	Statistic	Critical Value	Prob.**
None *	0.24736	157.2913	12.3209	0
At most 1 *	0.06255	29.13303	4.12991	0
Cointegration between D(Ln(USD/CNY)) and D(Ln(News Emotion))				
	Eigenvalue	Statistic	Critical Value	Prob.**
None *	0.22891	149.1997	12.3209	0
At most 1 *	0.06774	31.70484	4.12991	0
Cointegration between D(Ln(USD/CNY)) and D(Ln(Indicator Emotion))				
	Eigenvalue	Statistic	Critical Value	Prob.**
None *	0.25791	227.2567	12.3209	0
At most 1 *	0.18134	91.23955	4.12991	0

注: * 表示在 5% 的显著性水平上拒绝原假设;
**MacKinnon-Haug-Michelis (1999) *p*-values.

3.3 预测模型性能比较

本文的开发环境为 Matlab2021a, 操作系统为 Windows10. 为了验证本文选取的 LSTM 模型的预测性能, 本文以 ARIMA、SVR、ELM 作为基准模型 (预测模型参数设置如表 9 所示), 分别对汇率数据进行提前 1 天、提前 3 天和提前 6 天的预测, 并记录代表预测性能的指标数值. 表 10 展示了不同模型对测试集的预测性能. 根据表 4 预测结果的性能指标 MAE、RMSE、MAPE、DS, 我们有如下发现:

- 1) LSTM 模型的预测能力优于其余三种模型. 关注本文选取的 LSTM 模型, 其在提前 1 天、提前 3 天和提前 6 天的预测结果的 MAE、RMSE、MAPE 指标都小于同组的其他三种预测模型, DS 指标大于其他模型. 在提前 1 天预测组, ARIMA 的 MAE 值是 LSTM 的 18 倍, ELM 模型的 MAE 值是 LSTM 模型的 1.5 倍; 在提前 3 天组, 这一差异被放大, ARIMA 模型的 MAE 是 LSTM 模型的 24 倍, ELM 模型的 MAE 是 LSTM 模型的 2.2 倍. 在提前 1 天、3 天、6 天组中, 可以发现 LSTM 的 DS 值都高于其余三个模型.
- 2) 传统计量经济学线性预测模型 ARIMA 模型的预测精度表现最差. ARIMA 模型的水平预测精度评价指标与其他三种模型的相关指标有数量级上的差异, 可以认为线性预测模型对于汇率的预测能力很差, 难以刻画汇率的波动特征.

表 9 预测模型参数设置

模型	参数	决定方式	取值
ARIMA	AR	预设	1/3/6
	<i>I</i>	ADF 检测	2
	MA	自相关函数	2
SVR	正则化系数 <i>c</i>	网格搜索	[1, 300]
	内核参数 <i>g</i>	网格搜索	$[2^{-8}, 2^8]$
ELM	输入	预设	-
	隐藏层节点数	试错法	8
	输出	预设	1
	最大迭代次数	预设	200
LSTM	输入	预设	-
	隐藏层节点数	试错法	200
	输出	预设	1
	最大迭代次数	预设	200
	优化器	试错法	Sgdm
	Batch size	预设	30
	学习率	试错法	0.005

表 10 不同模型预测性能比较 (测试集)

	性能指标	ARIMA	SVR	ELM	LSTM
提前 1 天预测	MAE	0.059792	0.003604	0.005041	0.003271
	RMSE	0.066405	0.003863	0.006652	0.004167
	MAPE	3.119187	0.188012	0.26216	0.171026
	DS	57.6087	75.65217	67.56522	82.91304
提前 3 天预测	MAE	0.084725	0.004123	0.0079	0.003459
	RMSE	0.094863	0.005099	0.009972	0.004653
	MAPE	4.418776	0.21503	0.411235	0.180267
	DS	57.6087	61.08696	60	76.73913
提前 6 天预测	MAE	0.043527	0.015405	0.013544	0.003634
	RMSE	0.047336	0.019046	0.015771	0.004743
	MAPE	2.272112	0.801676	0.705673	0.189471
	DS	57.6087	48.56522	61.91304	68.91304

3) SVR 模型的预测精度极不稳定, 精度随着提前预测天数的增加大幅下降. 关注提前 1 天预测组, 可以发现 SVR 模型的各项指标仅次于 LSTM 模型, SVR 模型的 MAE 值仅比 LSTM 模型大 10%, 说明 SVR 模型在提前 1 天预测组展现出较好的预测性能. 而提前 6 天预测组中, SVR 模型的预测精度严重下降, 其 MAE 值比 LSTM 模型大 424%, 说明 SVR 模型仅在超短期预测时拟合效果较好, 在提前预测时间延长后, SVR 模型的预测精度大幅下降.

4) 随着提前预测天数的增加, 几乎所有模型的预测精度都呈现下降趋势. 以 LSTM 为例, 从提前 1 天预测到提前 3 天预测, 其 MAE 值增加了 5.75%; 从提前 3 天预测到提前

66 天预测, 其 MAE 值增加了 5.06%。SVR 的 MAE 对应增加值分别为 14.4% 和 273.64%; ELM 的 MAE 对应增加值分别为 56.71% 和 71.44%。值得注意的是, ARIMA 模型的提前 6 天水平预测精度指标比提前 1 天、提前 3 天都要小, 说明其提前 6 天预测性能最优。通过计算本文发现最优的 ARIMA 系数配置落在提前 6 天组, 说明这一结果是可靠的。

通过与三种基准模型 ARIMA、SVR、ELM 的对比, 本文发现 LSTM 模型的预测性能在横向与纵向角度都是最优的。因此本文用 LSTM 模型预测基于舆情信息的人们人民币汇率。

3.4 基于舆情信息对人民币汇率的预测

由于 LSTM 模型的高预测精度, 本文选用 LSTM 模型提前 1 天、3 天、6 天对 8 组预测组合进行预测。表 11~13 展现了提前 1 天、3 天、6 天的预测结果, 我们得到了如下结论:

1) 总的来说提前 1 天的预测效果优于提前 3 天, 优于提前 6 天。纵向比较表 11~13, 可以发现每个预测组合内存在随着提前预测时间增加, 预测精度减小的现象。以汇率为例, 提前 3 天的 MAE 值比提前 1 天的 MAE 值高 5.81%, 提前 6 天的 MAE 值比提前 3 天的 MAE 值高 4.91%; 汇率 + 指数组, 对应 MAE 提高量分别为 7.37% 和 7.46%; 汇率 + 研报 + 新闻 + 指数组, 对应 MAE 提高量分别为 4.73% 和 6.94%。关注提前 1 天、3 天、6 天组的 DS 值, 可以发现不同组之间的方向预测精度也存在明显的梯度关系, 随着预测天数增加, 方向预测精度逐步减小。

2) 舆情信息可以协助预测汇率趋势, 即舆情信息的加入优化了汇率预测的精度。以表 11 中提前 1 天预测结果为例, 汇率组预测的 MAPE 值在中最大, 为 0.17103, 而含舆情信息组别的 MAPE 值都小于这一数值。在提前 3 天预测的结果中, 这一现象依旧明显, 纯汇率预测组的 RMSE 值最大, 高于第二位的汇率 + 指数组 4.73%。在提前 6 天的预测结果中结论依旧, 但含有舆情信息与否对预测精度带来的影响逐渐缩小。可以发现纯汇率预测组和汇率 + 研报、汇率 + 新闻、汇率 + 指数组的 MAE、RMSE、MAPE 数值较为接近。

3) 社会舆情最能提高汇率预测精度, 专业舆情其次, 大众舆情对汇率精度的提高影响较小。以提前 1 天预测结果为例, 对比汇率 + 研报、汇率 + 新闻、汇率 + 指数组之间的预测精度, 可以发现汇率 + 研报、汇率 + 新闻组的水平预测精度指标更小, 汇率 + 指数组偏大, 且其 MAE 值大于汇率 + 新闻组 3.31%。在提前 3 天和 6 天的预测结果中也可以发现类似结论。值得注意的是, 汇率 + 研报组与汇率 + 新闻组的预测精度没有明显的顺序, 说明专业舆情和社会舆情对汇率预测影响程度随提前预测天数而不同。我们认为社会舆情最能提高汇率的预测精度的原因在于, 有关人民币汇率的新闻信息将更为广泛地影响市场参与者, 从而广泛地影响人们的投资或者换汇决策, 进而影响汇率波动。

而专业舆情可以更深入地影响专业投资者的决策, 进而影响大额的汇率波动。社会舆情和专业舆情都可以主动影响汇率波动, 反观大众舆情, 是一种对汇率搜索量的被动反应, 从而对汇率预测精度影响不如前两种舆情信息。

4) 相对单一舆情信息, 采用多种舆情信息更有助于提高汇率预测的精度, 且采用的舆情信息越多, 预测精度往往更高。以提前 1 天预测结果为例, 可以发现随着舆情信息种类的增加, 汇率预测精度有阶梯式的提升。汇率的 MAPE 值为 0.17103; 增加了一项舆情信息的汇率 + 研报、汇率 + 新闻、汇率 + 指数组的 MAPE 值最高为 0.16283, 小于纯汇率预测; 增

表 11 提前 1 天预测结果 (测试集)

	MAE	RMSE	MAPE	DS
汇率	0.00327	0.00417	0.17103	82.91304
汇率 + 研报	0.00308	0.00414	0.16087	83.47826
汇率 + 新闻	0.00302	0.00416	0.15761	87.82609
汇率 + 指数	0.00312	0.00431	0.16283	85.65217
汇率 + 研报 + 新闻	0.00278	0.00388	0.14525	88.91304
汇率 + 研报 + 指数	0.00294	0.00396	0.15362	87.82609
汇率 + 新闻 + 指数	0.00293	0.00409	0.15274	88.91304
汇率 + 研报 + 新闻 + 指数	0.00275	0.00391	0.14386	86.73913

表 12 提前 3 天预测结果 (测试集)

	MAE	RMSE	MAPE	DS
汇率	0.00346	0.00465	0.18027	76.73913
汇率 + 研报	0.00306	0.00416	0.15939	76.73913
汇率 + 新闻	0.00322	0.00436	0.1678	75.65217
汇率 + 指数	0.00335	0.00444	0.17442	73.47826
汇率 + 研报 + 新闻	0.00289	0.00398	0.15109	77.82609
汇率 + 研报 + 指数	0.00296	0.00412	0.15459	76.73913
汇率 + 新闻 + 指数	0.00298	0.00413	0.15528	74.56522
汇率 + 研报 + 新闻 + 指数	0.00288	0.0039	0.15015	75.65217

表 13 提前 6 天预测结果 (测试集)

	MAE	RMSE	MAPE	DS
汇率	0.00363	0.00474	0.18947	68.91304
汇率 + 研报	0.00351	0.00467	0.18315	67.82609
汇率 + 新闻	0.00346	0.00473	0.18025	64.56522
汇率 + 指数	0.0036	0.00489	0.18759	70
汇率 + 研报 + 新闻	0.00346	0.00462	0.18024	67.82609
汇率 + 研报 + 指数	0.00315	0.00431	0.16405	67.82609
汇率 + 新闻 + 指数	0.00332	0.00427	0.1734	88.91304
汇率 + 研报 + 新闻 + 指数	0.00308	0.00427	0.16055	68.56522

加了两项舆情信息的汇率 + 研报 + 新闻、汇率 + 研报 + 指数、汇率 + 新闻 + 指数组的 MAPE 值最高为 0.15362, 小于增加了一项舆情信息组的最小值; 汇率 + 研报 + 新闻 + 指数组的 MAPE 值最小. 为了更直观了解增加舆情项目数对汇率预测精度的提升程度, 我们对纯汇率预测、增加一项、两项、三项舆情信息的 MAPE 值分别取组内平均值, 得到值分别为 0.17103、0.16044、0.15054、0.14386. 在提前 3 天和 6 天的预测结果中可以发现结论依旧明显.

3.5 DM 检验

在本节中, 我们对三种基准模型 ARIMA、ELM、SVR 和我们选用的 LSTM 模型的预测结果进行了 DM 检验, 本文使用了均方误差 MSE 作为 DM 统计量的损失函数, 并分别对提前 1 天、3 天、6 天进行 DM 检验. DM 检验的结果展示在表 14 中. 我们发现如下结论:

1) 本文所选用的 LSTM 模型与其他三种基准模型的 DM 检验结果均小于 -1.03709 , 所对应的 p 值绝大多数都小于 0.01 , 只有 LSTM 与 SVR 项小于 0.05 . 这说明本文采用的预测模型在 95% 的置信水平下好于另外三种基准模型. 说明本文选取 LSTM 模型探究舆情信息对汇率预测精度的提高是有效的.

2) 关注提前 1 天、3 天、6 天组中, 当 ARIMA 作为检验模型时, 其他三个模型 LSTM、ELM、SVR 的 DM 检验结果都远小于 -12.5 , 且 p 值小于 0.01 . 这一现象说明 ARIMA 模型的预测性能远差于其他三种模型, 究其原因, 可能是 ARIMA 的线性预测特性难以适应人民币汇率的波动特性和非线性特征.

表 14 DM 检验

	模型	LSTM	ELM	SVR
提前 1 天预测	ELM	-5.09250(0.0000)		
	SVR	-3.63327(0.0302)	-1.03709 (0.0005)	
	ARIMA	-13.1491(0.0000)	-13.07850 (0.0000)	-13.05790 (0.0000)
提前 3 天预测	ELM	-5.83605(0.0000)		
	SVR	-6.15748(0.0000)	-4.98757 (0.0000)	
	ARIMA	-12.5947(0.0000)	-12.54490 (0.0000)	-12.53010 (0.0000)
提前 6 天预测	ELM	-8.02387(0.0000)		
	SVR	-6.42664(0.0000)	3.72494(0.0003)	
	ARIMA	-14.6176(0.0000)	-13.76670 (0.0000)	-13.51910 (0.0000)

4 结论

本文提出了一种基于多种舆情信息预测人民币汇率趋的研究方法. 首先, 本文收集了有关人民币汇率的研报和新闻的舆情信息, 并使用 NLP 技术及基于词典的文本数据情感分析方法, 将文本数据转化为可计算的专业舆情和社会舆情情感指数. 其次, 本文综合了两大搜索引擎对人民币汇率搜索的指数, 构建了大众舆情指数. 接着, 本文通过提前 1 天、3 天、6 天, 分别对比了四种预测汇率常用模型的预测精度, 筛选出最适合预测人民币汇率趋势的模型 LSTM. 最后, 本文构建了 8 组“汇率-舆情”组合, 使用 LSTM 模型提前 1 天、3 天、6 天预测人民币汇率趋势, 并对比舆情信息对预测精度提高的程度.

根据实证检验的结果我们得到以下结论: 1) 在提前 1 天、3 天、6 天组, 我们选用的 LSTM 模型对测试集的水平预测精度指标和方向预测精度指标都是最优的. 原因在于这一深度学习的网络结构可以更精准地刻画长期数据趋势, 优于传统计量经济学线性预测模型和基础的机器学习算法, 体现出 LSTM 模型对刻画汇率波动的优势. 2) 有强烈的证据表明, 舆情信息的加入可以显著提高汇率预测精度. 在提前 1 天、3 天、6 天组内, 都存在纯汇率预测精度最低这一结果. 但是随着提前预测天数的增加, 舆情信息对预测精度的提高程度逐渐缩

小. 其原因在于提前预测天数的增加导致了舆情信息中部分内容没有被纳入模型, 从而降低了预测精度. 3) 在我们构建的 8 组“汇率-舆情”预测组合中, 舆情信息项目越多, 汇率预测精度越高. 值得注意的是, 在提前 1 天、3 天、6 天组中, 随着舆情信息项目的增加, 预测精度有阶梯式的提升, 最能提高汇率预测精度的组别是“汇率 + 研报 + 新闻 + 指数”组. 其原因在于综合了专业、社会、大众舆情, 可以更全面且系统地刻画汇率波动的主动和被动因素, 从而提高汇率预测精度. 4) 对于提高预测精度的作用, 我们认为社会舆情大于专业舆情, 专业舆情大于大众舆情. 其内在原因在于, 社会舆情可以更广泛地影响社会公民, 普遍的导致群体换汇等一系列决策的发生; 而专业舆情可以深入地影响金融机构或者专业投资者的决策, 一定程度上也会主动影响汇率的波动; 大众舆情往往是汇率波动的被动反应, 体现在群众看到汇率波动后的互联网搜索上, 因此造成了三种舆情对汇率预测精度影响程度不同的局面. 5) 随着提前预测天数的增加, 汇率预测的精准程度逐渐减弱. 我们发现汇率预测精度往往在提前 1 天最为准确, 提前 3 天到提前 6 天, 汇率预测的精度呈现下降趋势. 背后的原因在于市场对汇率波动信息的反应往往是及时和迅速的, 随着时间的增加, 相关信息对汇率的影响作用开始消散; 除此之外, 在预测模型训练时, 具有时效性的信息没有被纳入训练集, 从而导致预测效力减弱.

本文使用了研报、财经新闻、搜索指数作为舆情信息, 对于微博, 贴吧评论等时效高、量级大社交舆情信息本文并未引入. 如果可以将类似的社交信息作为舆情信息的一种纳入汇率预测当中, 预测精度可能会增强, 可以作为未来的研究方向. 总的来说, 本文通过纳入多方位的舆情信息, 为提高汇率预测精度提供了一种新的视角和思路, 对帮助市场更好地理解外汇市场价格的形成机制, 对相关政策制定者和市场参与者做决策有着重要的意义.

参 考 文 献

- 成舟, 余峥, 过弋, 王志宏, (2019). 基于外汇舆情的人民币汇率波动预测研究 [J]. 计算机科学, 46(S2): 143–148.
- Cheng Z, Yu Z, Guo Y, Wang Z W, (2019). Research on Volatility Forecasting of RMB Exchange Rate Based on Public Opinion[J]. Computer Science, 46(S2): 143–148.
- 黄丽明, 陈维政, 闫宏飞, 陈翀, (2019). 基于循环神经网络和深度学习的股票预测方法 [J]. 广西师范大学学报 (自然科学版), 37(1): 13–22.
- Huang L M, Chen W Z, Yan H F, Chen C, (2019). A Stock Prediction Method Based on Recurrent Neural Network and Deep Learning[J]. Journal of Guangxi Normal University (Natural Science Edition), 37(1): 13–22.
- 李佳, 黄之豪, 陈冬兰, (2019). 基于 GRU 神经网络的欧元兑美元汇率预测研究 [J]. 浙江金融, (3): 12–19.
- Li J, Huang Z H, Chen D L, (2019). Research on Euro/USD Exchange Rate Forecast Based on GRU Neural Network[J]. Zhejiang Finance, (3): 12–19.
- 李艳丽, 邓贵川, 李辰阳, (2016). 人民币汇率波动的预测 —— 基于损失函数和 DM 检验的比较分析 [J]. 国际金融研究, (2): 84–96.
- Li Y L, Deng G C, Li C Y, (2016). Explorations of Financial Products as Research on the Predictability of RM Exchange Rate Volatlty — Anaysis Based on Loss Function and DM Test the Medium of Exchange[J]. Studies of International Finance, (2): 84–96.

- 倪铮, 梁萍, (2018). 基于 LSTM 深度神经网络的精细化气温预报初探 [J]. 计算机应用与软件, 35(11): 233–236.
- Ni Z, Liang P, (2018). Fine Temperature Forecast Based on LSTM Deep Neural Network[J]. Computer Applications and Software, 35(11): 233–236.
- 饶东宁, 邓福栋, 蒋志华, (2017). 基于多信息源的股价趋势预测 [J]. 计算机科学, 44(10): 193–202.
- Rao D N, Deng F D, Jiang Z H, (2017). Stock Price Movements Prediction Based on Multisources[J]. Computer Science, 44(10): 193–202.
- 任仙玲, 邓磊, (2019). 网络舆情对人民币汇率的冲击效应 —— 基于中美贸易摩擦事件 [J]. 管理科学, 32(6): 46–56.
- Ren X L, Deng L, (2019). Effect of Network Public Opinion Shocks on Exchange Rate: Evidence from the China-US Trade Friction[J]. Journal of Management Science, 32(6): 46–56.
- 孙少龙, 魏云捷, 黎建强, (2022). 基于在线外汇新闻情感挖掘的汇率预测研究 [J]. 计量经济学报, 2(2): 441–464.
- Sun S L, Wei Y J, Li J Q, (2022). Exchange Rate Forecasting with Online Forex News Sentiment Mining[J]. China Journal of Econometrics, 2(2): 441–464.
- 王吉祥, 过弋, 戚天梅, 王志宏, 李真, 等, (2019). 嵌入互联网舆情强度的人民币汇率预测 [J]. 计算机应用与软件, 39(11): 3403–3408.
- Wang J X, Guo Y, Qi T M, Wang Z H, Li Z, et al. (2019). RMB Exchange Rate Forecast Embedded with Internet Public Opinionintensity[J]. Journal of Computer Applications, 39(11): 3403–3408.
- 吴武清, 赵越, 闫嘉文, 汪寿阳, (2020). 分析师文本语调会影响股价同步性吗? —— 基于利益相关者行为的中介效应检验 [J]. 管理科学学报, 23(9): 108–126.
- Wu W Q, Zhao Y, Yan J W, Wang S Y, (2020). Does Textual Tone in Analyst Reports Affect Stock Price Synchronicity? Ananalysis Based on Mediating Effects of Stakeholders' Behavior[J]. Journal of Management Sciences in China, 23(9): 108–126.
- 许启发, 伯仲璞, 蒋翠侠, (2017). 基于分位数 Granger 因果的网络情绪与股市收益关系研究 [J]. 管理科学, 30(3): 147–160.
- Xu Q F, Bo Z P, Jiang C X, (2017). Exploring the Relationship between Internet Sentiment and Stock Market Returns Based on Quantile Granger Causality Analysis[J]. Journal of Management Science, 30(3): 147–160.
- 杨超, 姜昊, 雷峥嵘, (2019). 基于文本挖掘和百度指数的汇率预测 [J]. 统计与决策, 35(13): 85–87.
- Yang C, Jiang H, Lei Z R, (2019). Exchange Rate Forecast Based on Text Mining and Baidu Index[J]. Statistics and Decision, 35(13): 85–87.
- 姚加权, 冯绪, 王赞钧, 纪荣嵘, 张维, (2021). 语调、情绪及市场影响: 基于金融情绪词典 [J]. 管理科学学报, 24(5): 26–46.
- Yao J Q, Feng X, Wang Z J, Ji R R, Zhang W, (2021). Tone, Sentiment and Market Impacts: The Construction of Chinese Sentiment Dictionary in Finance[J]. Journal of Management Sciences in China, 24(5): 26–46.
- 尹海员, 吴兴颖, (2019). 投资者高频情绪对股票日内收益率的预测作用 [J]. 中国工业经济, (8): 80–98.
- Yin H Y, Wu X Y, (2019). Predictive Effect of High-frequency Investor Sentiment on the Intraday Stocks Return[J]. China Industrial Economics, (8): 80–98.
- 张杰, 张永卿, 翟东升, (2021). 融合财经新闻信息的汇率波动预测 [J]. 系统工程, 39(3): 121–131.
- Zhang J, Zhang Y Q, Zhai D S, (2021). The Prediction of Exchange Rate Fluctuation Based on Financial News Information[J]. Systems Engineering, 39(3): 121–131.

- Bulut L, (2018). Google Trends and the Forecasting Performance of Exchange Rate Models[J]. *Journal of Forecasting*, 37(3): 303–315.
- Chen X Y, Chau K W, Busari A O, (2015). A Comparative Study of Population-based Optimization Algorithms for Downstream River Flow Forecasting by a Hybrid Neural Network Model[J]. *Engineering Applications of Artificial Intelligence*, 46: 258–268.
- Colonescu C, (2018). The Effects of Donald Trump's Tweets on US Financial and Foreign Exchange Markets[J]. *Athens Journal of Business & Economics*, 4(4): 375–388.
- Crone S F, Koeppel C, (2014). Predicting Exchange Rates with Sentiment Indicators: An Empirical Evaluation Using Text Mining and Multilayer Perceptrons[C]// 2014 IEEE Conference on Computational Intelligence for Financial Engineering & Economics (CIFEr). IEEE: 114–121.
- Dickey D A, Fuller W A, (1979). Distribution of the Estimators for Autoregressive Time Series with a Unit Root[J]. *Journal of the American Statistical Association*, 74(366a): 427–431.
- Feuerriegel S, Gordon J, (2018). Long-Term Stock Index Forecasting Based on Text Mining of Regulatory Disclosures[J]. *Decision Support Systems*, 112: 88–97.
- Graves A, (2012). Supervised Sequence Labelling with Recurrent Neural Networks[C]// Supervised Sequence Labelling with Recurrent Neural Networks: 1–141. Doi: 10.1007/978-3-642-24797-2.
- Hochreiter S, Schmidhuber J, (1997). Long Short-term Memory[J]. *Neural Computation*, 9(8): 1735–1780.
- Huang A Z, Zhang Y L, Peng J P, Chen H, (2022). Application of Informetrics on Financial Network Text Mining Based on Affective Computing[J]. *Information Processing & Management*, 59(2): 102822.
- Jabeen A, Yasir M, Ansari Y, Yasmin S, Moon J, et al. (2022). An Empirical Study of Macroeconomic Factors and Stock Returns in the Context of Economic Uncertainty News Sentiment Using Machine Learning[J]. *Complexity*: 4646733.
- Oliveira N, Cortez P, Areal N, (2017). The Impact of Microblogging Data for Stock Market Prediction: Using Twitter to Predict Returns, Volatility, Trading Volume and Survey Sentiment Indices[J]. *Expert Systems with Applications*, 73: 125–144.
- Rao T, Srivastava S, (2013). Modeling Movements in Oil, Gold, Forex and Market Indices Using Search Volume Index and Twitter Sentiments[C]// Proceedings of the 5th Annual ACM Web Science Conference: 336–345.
- Semiromi H N, Lessmann S, Peters W, (2020). News will Tell: Forecasting Foreign Exchange Rates Based on News Story Events in the Economy Calendar[J]. *The North American Journal of Economics and Finance*, 52: 101181.
- Wilcoxson J, Follett L, Severe S, (2020). Forecasting Foreign Exchange Markets Using Google Trends: Prediction Performance of Competing Models[J]. *Journal of Behavioral Finance*, 21(4): 412–422.
- Yasar H, Kilimci Z H, (2020). US dollar/Turkish Lira Exchange Rate Forecasting Model Based on Deep Learning Methodologies and Time Series Analysis[J]. *Symmetry*, 12(9): 1553.
- Zhang D, Simoff S J, Debenham J, (2005). Exchange Rate Modelling Using News Articles and Economic Data[C]// AI 2005: Advances in Artificial Intelligence: 18th Australian Joint Conference on Artificial Intelligence, Sydney, Australia, December 5–9, 2005. Proceedings 18. Springer Berlin Heidelberg: 467–476.