

财经情感词典面临的挑战与机遇

韩俊杰

重庆城市职业学院 重庆永川 402160

DOI:10.12238/jief.v6i5.7851

[摘要] 随着经济的快速发展，财经领域信息量的爆炸式增长为从海量文本中提取有价值信息带来了挑战。作为解读市场情绪和预测趋势的新工具，财经情感词典提供了市场心理理解的新视角。本文探讨了财经情感词典在构建和应用过程中面临的数据质量与规模、专业术语识别、情感极性歧义和实时性需求等挑战，以及信息科技的发展、信息消费的升级以及财经市场的复杂性等因素，为财经情感词典的应用和发展提供的重要机遇。

[关键词] 财经情感词典、情感分析、挑战与机遇

Challenges and opportunities facing the financial emotion dictionary

Han Junjie

Chongqing City Vocational College, Chongqing Yongchuan 402160

[Abstract] With the rapid development of economy, the explosive growth of information in the financial and economic field has brought a challenge to extracting valuable information from massive texts. As a new tool to interpret market sentiment and predict trends, the Financial Emotion Dictionary provides a new perspective on the understanding of market psychology. This paper discusses the financial and emotional dictionary in the process of building and application of data quality and scale, professional term identification, emotional polarity ambiguity and real-time challenges, as well as the development of information technology, the upgrade of information consumption and the complexity of the financial market, for the application of financial emotional dictionary and development of the important opportunities.

[Key words] the dictionary of financial emotions, emotion analysis, challenges and opportunities

一、引言

随着我国经济的快速发展，财经领域的信息量呈现出爆炸式增长，如何从海量财经文本中快速、准确地获取有价值的信息，成为了当前研究的热点问题。情感分析作为一种新兴技术，正逐渐成为投资者和分析师解读市场情绪、预测市场趋势的重要工具。财经情感词典作为情感分析的关键组成部分，通过量化财经文本中的情感色彩，为理解市场心理提供了新的视角。然而，这一领域的研究与应用仍处于发展阶段，面临着一系列挑战与机遇。财经情感词典的构建不仅需要大量的情感标注数据和专业知识，而且财经领域的文本往往包含了大量的专业术语和复杂情感，这使得情感词典的构建和应用变得更加困难。然而，随着信息科技的发展，财经情感词典在处理大规模数据集和复杂情感分析方面的能力将得到显著增强。

二、财经情感词典的构建

（一）数据来源

数据来源是财经情感词典构建的基础。高质量的数据集可以为词典提供丰富的语料，进而提高情感分析的准确度。在数据收集过程中，通常会选择财经新闻、社交媒体、股票论坛、年度报告等渠道，以获取与财经市场相关的文本数据。为了保证数据质量，需要对数据进行严格筛选，确保其准确性和多样性。数据来源的多样性有助于构建更加全面和准确的财经情感词典。

（二）情感标注

情感标注是财经情感词典构建的关键步骤。情感标注主要包括正向情感、负向情感和中性情感三种，对标注者的专业知识和语言理解能力提出了较高要求。为了实现精确标注，可以采用人工标注、机器学习算法和深度学习模型等方法。人工标注是由领域专家根据经验和直觉进行的，通常会采用多人标注

和交叉验证的方法，以确保标注的准确性和专业性。机器学习算法和深度学习模型可以自动从大量文本数据中学习情感倾向，能提高标注的效率和准确性。情感标注的准确性直接影响到财经情感词典的质量。

（三）词典构建

词典构建是在情感标注基础上进行的，将标注好的数据进行整理和归纳，通过词频统计、词性分析、语法分析等方法，构建包含情感极性的词汇库，即财经情感词典。财经情感词典应包含关键词、短语等多种粒度，以满足不同层次的财经情感分析需求。财经情感词典的构建需要综合考虑词汇的情感极性、情感强度、情感类别等多个维度。

三、财经情感词典面临的挑战

（一）数据质量与规模

数据质量与规模是财经情感词典构建的首要挑战。高质量的数据是确保情感分析准确性的前提。随着财经领域数字化的快速发展，财经文本数据量呈现出爆炸式增长。然而，财经文本数据虽然数据规模庞大，但往往存在噪声多、信息不完整等问题，需要对数据进行严格的筛选和清洗，以提高数据质量。同时，数据规模是影响分析结果泛化能力的关键因素。大规模的数据集能够提供更全面的市場情绪视图，从而增强分析结果的代表性和泛化能力，但同时也带来了存储和处理上的挑战。大规模数据的收集、存储和处理需要相应的技术支持和资源投入。如何获取并处理大规模、高质量的财经文本数据，是构建财经情感词典需要解决的首要问题。

（二）财经术语的识别与理解

财经领域的术语识别与理解，是构建财经情感词典时面临的一项挑战。如何精确地识别和理解这些术语，是构建词典过程中必须解决的关键问题。财经领域文本中充斥着众多专业术语和复杂表达，它们往往具有多重含义和丰富的感情内涵，其感情色彩亦随上下文变化而变化。财经情感词典的构建者必须具备对语境的敏感度，能够捕捉术语在不同情境下的感情色彩。由于财经术语的多样性和复杂性，词典的构建者还需要具备跨学科的知识，以全面理解和解释术语。因此，财经情感词典的构建需要具备深入的语境理解能力，以便在特定情境下准确识别并解释这些术语的感情倾向。为了提高财经情感词典的准确性和实用性，可采用自然语言处理技术，如情感分析和语义分析，来辅助识别和解释术语。通过机器学习算法，可训练模型识别术语的感情倾向，并根据上下文进行调整。

（三）感情极性的歧义问题

在构建财经领域的情感词典时，感情极性的歧义性构成了一个根本性的挑战，这一挑战的根源在于自然语言的多义性和对上下文的依赖。这种语言特性导致相同词汇在不同的财经语境中可能呈现出截然不同的感情倾向，从而为感情分析的任务增添了额外的复杂性。例如，词语“涨”在财经语境中就表现

出感情的复杂性：在股票市场中，它通常与正向感情相关联，象征着投资者财富的增长；然而，在谈论物价时，它却往往带有负向感情，因为它意味着生活成本的提升。这种一个词汇同时蕴含正面和负面感情色彩的现象，不仅增加了感情分析的难度，而且对构建准确反映词汇感情倾向的词典提出了挑战。此外，财经领域的专业术语和表达方式也增加了感情极性歧义性的复杂性。许多财经词汇具有特定的感情色彩，对这些词汇的理解和把握需要具备相关的专业知识。这就要求我们在构建财经情感词典时，不仅要关注词汇的多义性和上下文依赖性，还要深入了解财经领域的专业知识。

四、财经情感词典的机遇

（一）信息科技的发展

信息科技的迅猛发展为财经情感词典的构建与应用注入了新的活力。在大数据、云计算、人工智能等前沿技术的推动下，财经情感词典得以在处理和数据分析数据集的广度与深度上实现显著提升，极大地扩展了财经情感词典的能力边界，使其在准确性与效率方面均有了质的飞跃。

大数据技术的引入，为财经情感词典提供了从海量财经文本中挖掘有价值信息的可能。财经文本数据的体量庞大、类型多样，传统的分析方法难以应对其复杂性与多变性。而大数据技术的应用，使得财经情感词典能够高效地处理和分析这些数据，提取出反映市场情绪的关键信息。金融市场的动态变化要求分析工具能够快速响应，提供即时的市场情绪反馈。信息科技的进步，特别是流数据处理技术的应用，使得财经情感词典能够实现对实时数据的快速分析，为投资者和分析师提供了更为丰富和深入的市场洞察。

人工智能技术，尤其是机器学习和深度学习算法的发展，为财经情感词典的感情分析能力带来了革命性的提升。这些算法能够自动识别和学习文本中的感情模式，从而提高对复杂感情倾向的识别能力。深度学习算法，特别是自然语言处理领域的创新，使得财经情感词典在理解语境、识别多义词以及处理歧义表达方面更加精准，极大地提高了感情分析的深度和广度。

（二）信息消费升级

在信息消费升级的浪潮中，财经情感词典正迎来前所未有的发展机遇。随着投资者对财经信息需求的不断增长，他们对信息的质量和深度提出了更高的要求。财经情感词典以其独特的市场情绪分析能力，为投资者提供了深入洞察市场趋势和风险的新途径，不仅丰富了投资者的信息来源，也提高了其对市场动态的理解和应对能力。信息消费的升级为财经情感词典的发展提供了多维度的机遇。财经情感词典需要不断适应信息消费升级的趋势，优化自身的分析能力和服务模式，以满足投资者和企业对高质量财经信息的需求。

信息消费的升级，意味着投资者对财经信息的需求不再局

限于基本的数字和事实，而是更加注重信息背后的情感色彩和市场心理。财经情感词典通过对财经文本的情感倾向进行分析，揭示了市场情绪的微妙变化，为投资者提供了一种全新的视角，有助于投资者识别市场的潜在风险和机会，从而做出更为明智的投资决策。随着投资者对个性化和定制化信息需求的增加，财经情感词典可根据投资者的特定需求，提供定制化的情感分析服务，如投资者可以根据自己的投资偏好，选择关注特定行业或市场的情感分析，从而获得更加精准和有针对性的市场信息。

移动设备和社交媒体的普及，进一步推动了信息消费的升级。投资者可以随时随地通过智能手机、平板电脑等移动设备获取财经信息，同时利用社交媒体平台分享和讨论市场观点。这种即时性和互动性，为财经情感词典的应用提供了便利条件。投资者可以在第一时间获取到经过情感分析的财经信息，及时调整自己的投资策略。随着投资者对财经信息的需求日益多样化，财经情感词典需要不断优化其分析模型，提高分析的准确性和深度。同时，财经情感词典也需要加强与移动设备和社交媒体的融合，提升用户体验，满足投资者对便捷性和互动性的需求。

(三) 财经市场的复杂性

财经市场以其固有的复杂性而闻名，这种复杂性源自于宏观经济的波动、政策环境的变迁以及市场参与者心理的多样性。正是这些因素的交织，为财经情感词典的应用提供了一个充满机遇的广阔领域。财经情感词典通过分析文本中的情感色彩，揭示了市场情绪的深层次动态，为投资者提供了一种全新的市场分析工具。

在宏观经济层面，财经情感词典能够捕捉到经济指标变化对市场情绪的即时影响，如利率的调整往往会在投资者中引起一系列情绪反应。财经情感词典通过对这些反应的分析，能够帮助投资者理解市场情绪如何随着宏观经济环境的变化而变化。政策变化是影响市场情绪的另一个关键因素。政府的财政政策、货币政策或是行业监管政策的调整，都会对市场产生深远的影响。财经情感词典能够通过对政策相关文本的分析，评估政策变动对市场情绪的潜在影响，为投资者提供政策影响的预判。市场心理的复杂性则为财经情感词典的应用提供了更为微妙的机遇。市场心理包括投资者的预期、信心以及风险偏好等，这些心理因素在市场波动中起着至关重要的作用。财经情感词典通过分析投资者在论坛、社交媒体等平台上的讨论，可以洞察市场心理的微妙变化，从而为投资者提供更为精准的市场情绪画像。

五、结论

财经情感词典的构建与应用，作为财经领域的创新实践，正站在信息技术与财经市场深度融合的前沿。财经情感词典，作为情感分析的核心组件，通过量化分析文本中的情感色彩，

为揭示市场心理提供了新的视角。然而，财经情感词典的发展之路并非一帆风顺。它面临着数据质量与规模的挑战，需处理大量且质量参差不齐的财经文本数据；财经术语的识别与理解，以及情感极性的歧义问题，增加了构建的复杂性。尽管存在挑战，财经情感词典的发展机遇同样明显。信息科技的发展，特别是大数据、云计算和人工智能技术的突破，为财经情感词典提供了强大的技术支持。信息消费的升级，投资者对财经信息质量和深度要求的提升，为财经情感词典提供了更广阔的市场空间。财经市场的复杂性，为情感分析的应用提供了更深层次的需求，使其在风险管理和合规性检查方面展现出独特的价值。

展望未来，财经情感词典将在信息科技的推动下，不断提升其在市场情绪分析、风险评估和投资决策中的作用。它将更好地适应信息消费升级的趋势，满足投资者和企业对高质量财经信息的需求。随着技术的不断演进和市场对财经情感分析工具需求的增加，财经情感词典有望成为财经市场分析中的关键工具，为投资者和分析师提供更加深入和精准的市场洞察。

[参考文献]

- [1]陈宋生,王朝晖,朱乐琪,等. 大数据、学科交叉与会计研究: 机遇与挑战[J]. 财会通讯, 2023, (20): 3-10, 55.
- [2]邢玲,程兵. 基于结巴分词的领域自适应分词方法研究[J]. 计算机仿真, 2023, 40(04): 310-316, 503.
- [3]毛淑珍,陈冉,谢飞. 会计文本语调研究文献综述及展望[J]. 2023, 44(08): 61-68.
- [4]李合龙,任昌松,柳欣茹,等. 金融市场文本情绪研究综述[J]. 数据分析与知识发现, 2023, 7(12): 22-39.
- [5]姚加权,张银澎,罗平. 金融学文本大数据挖掘方法与研究进展[J]. 经济学动态, 2020, (04): 143-158.
- [6]赵又霖,徐竟楠,陆颖隽. 面向实时财经信息的领域情感歧义搭配词典构建研究[J]. 数据分析与知识发现, 2023, 7(07): 100-110.
- [7]鲁万波,曾攀,亢晶浩,等. 管理层的讨论与分析语调对企业未来业绩表现的影响研究[J]. 数理统计与管理, 2023, 42(03): 391-402.
- [8]李守皓,丁立贵. 公众情感与经济波动的实现: 以财经新闻为例[J]. 贵州财经大学学报, 2021, (05): 30-41.

作者简介：韩俊杰（1986—），男，山东菏泽人，重庆城市职业学院商学院高级会计师。

基金项目：本文系重庆城市职业学院智慧零售协同创新中心项目成果，平台编号：KYPT202200003、重庆市教育委员会科学技术研究计划 2023 年度青年项目资助项目（项目编号：KJQN202303916）和 2023 年重庆市教育科学规划课题教改专项“基于 RPA 技术的高职院校会计专业人机协作技能培养研究”（项目编号：K23ZG3160193）资助项目的阶段性成果。