在金融学期刊的找一些文章，这些文章是用机器学习模型来做的金融/ESG问题

国外A+至B等级，近3年

* 论坛上搜/提问：哪些金融期刊发表ML的文章多
* 看看老师们的主页，他们发在哪了
* 别人的参考文献

Expert Systems with Applications

1. 题目1
2. 题目2

Journal of Credit Risk

1. 题目1
2. 题目2

通过对相关文献的梳理和投资基金的调研，现有的ESG评级指标体系难以赢得证券市场投资者信任，引导市场资源在可持续领域的有效配置，其原因可以主要归纳为以下几个方面**：1）主观性过强，缺乏独立性，导致结果产生偏差。**评级机构获取的数据与信息越来越依赖于企业，企业更倾向于披露对自己有利的信息，即偏重于宣传业绩，对负面指标鲜少涉及，并不能保证真实地表达企业的情况（叶陈刚等，2015）。特别是，在环境信息的获取上，公司披露的信息及公开可用的信息并不能保证真实地表达企业的情况。此外，一些评级机构在评估和评级之外，还为公司提供咨询服务，例如财务分析、风险评估等，潜在的利益冲突影响了评级机构的独立性，降低评级信息的可信度。**2）透明度不足，深度披露不够，导致公信力降低**。目前，一些主要ESG评估机构几乎都认为评级方法的具体内容、具体的评级指标属于商业机密，公布信息有限。例如，对信息收集方法、假设、计算、比重、阈值和分析等均很少披露或披露程度不足。因此，使用者和利益相关者无法对其评估和评分结果做出完整的评价。目前，部分评级机构开始逐渐公开评级过程的少量初步细节，在一定程度上增加了透明度，但核心的指标和评估过程的具体过程并未披露。3）**时效性差，不能及时反应企业的重大变化，消弱其市场风向标的能力**。目前，大部分ESG评级指标都是半年或者一年一公布。这主要因为评级机构所采用的数据主要来自企业财务报表（如季报、年报）和独立报告（如社会责任报告、可持续发展报告）。这些报告通常以季度、半年度、年度为报告周期，导致评估的企业ESG评分更新频率低，显然难以满足投资者的日常投资需求（Giese等，2019）。

表 1. 国内外主流ESG评级体系主要情况

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 指标名称 | 主要数据来源 | 评级方法 | 评级结果 | 评级周期 |
| MSCI KLD | 年报、企业披露数据、监管数据 | 对最重要的若干个  ESG问题进行评分 | -11 到 +19 | 年度 |
| MSCI IVA | 宏观数据、企业访谈、企业披露数据 | 确定关键问题及  关键问题比重 | AAA到CCC | 年度 |
| Asset4 | 年报、官网、CSR报告 | 重要性矩阵、  ESG争议事件得分 | 百分制 | 年度 |
| Bloomberg | 年报、CSR报告、新闻、  公司治理报告 | 基于SASB标准框架  指定指数标准 | 百分制 | 年度 |
| Sustainalytics | CSR报告、ESG问题、  ESG事件 | “风险敞口”和“风险管理”二维评级方法 | 百分制 | 年度 |
| RobecoSAM | 问卷数据、监管数据、  少量媒体数据 | 同行业对比打分 | 百分制 | 年度 |
| 商道融绿 | 年报、CSR报告、  政府信息 | 根据行业实证因子加权计算，重视负面事件 | A+ 到 D | 年度 |
| 社投盟 | 金融机构、上市公司、学术研究、政策研究等信息 | 评级“利益并举”，遵循“负面剔除、综合打分” | AAA 到 D | 半年度 |
| 华证 | CSR报告、政府和监管部门数据、新闻数据等 | 剔除不合理指标、加入特色指标，构建行业权重 | AAA到C | 季度 |

表 2. ESG有效性特征集

|  |  |  |
| --- | --- | --- |
| 维度 | 特征变量 | 子特征 |
| 环境 | 气候变化  （FE1） | 温室气体排放（FE11）、碳排放（FE12）、全球变暖（FE13）、  环境脆弱性（FE14） |
| 自然资源  （FE2） | 水资源压力（FE21）、矿物资源（FE22）、生态多样性（FE23）、  土地使用（FE24）、耗水量（FE25） |
| 污染和废弃物（FE3） | 水污染（FE31）、大气污染（FE32）、有毒物质排放（FE33）、  电子废弃物（FE34）、包装材料使用（FE35）、  环境处罚（FE36）、环保投入（FE37） |
| 能源消耗及管理（FE4） | 清洁能源（FE41）、节能减排（FE42）、可再生能源（FE43）、  风能（FE44）、太阳能耗电量（FE45）、耗电量（FE46） |
| 社会 | 人力资本  （FS1） | 人力资源管理（FS11）、员工健康（FS12）、员工福利（FS13）、  员工满意度（FS14）、员工流失率（FS15）、员工培训（FS16） |
| 产品可靠性  （FS2） | 产品质检（FS21）、产品召回（FS22）、产品质量与安全（FS23）、隐私数据安全（FS24）、服务质量评价（FS25）、网络数据安全（FS26） |
| 消费者和社区关系  （FS3） | 消费者权益（FS31）、社会责任（FS32）、慈善公益（FS33）、  社区贡献项目（FS34） |
| 人权  （FS4） | 招聘歧视（FS41）、工作机会平等（FS42）、使用童工（FS43）、  强制劳动（FS44）、保证结社自由（FS45）、性别比例（FS46） |
| 治理 | 公司治理  （FG1） | 董事会成员（FG11）、管理结构调整（FG12）、公司制度（FG13）、  股东权利（FG14）、会计准则（FG15） |
| 公司行为  （FG2） | 商业伦理（FG21）、腐败与不稳定性（FG22）、  反垄断行动（FG23）、税收透明度（FG24） |
| 商业和产品策略  （FG3） | 产品合格率（FG31）、售后服务（FG32）、物流及时率（FG33）、  生产循环效率（FG34） |
| 风险管理  （FG4） | 未来战略目标（FG41）、不确定因素结果控制（FG42）、  协调风险容量（FG43）、改善资本调配（FG44） |
| 供应链管理  （FG5） | 物流控制（FG51）、信息流控制（FG52）、资金流控制（FG53）、  集成业务流程（FG54）、组织间协作（FG55） |
| 激励薪酬  （FG6） | 高管薪酬（FG61）、员工薪酬（FG62）、薪酬构成（FG63）、  股权形式（FG64）、薪酬与业绩关系（FG65）、  集权与分权（FG66）、货币化职务消费（FG67） |
| 信息披露  （FG7） | 招股说明书（FG71）、上市公告书（FG72）、定期报告（FG73）、  责任人违法违纪披露（FG74）、行政处罚披露（FG75）、  临时报告（FG76）、审计意见披露（FG77） |

**（2）基于深度学习模型的ESG新闻内容重要程度研究**

现有的ESG评级指标体系透明度不足，深度披露不够，导致公信力下降。目前，一些主要ESG评估机构几乎都认为具体的评级方法、评级指标属于商业机密，公布信息有限。例如，对评级指标的计算、权重、阈值和分析等，均很少披露或披露程度不足，因此评级的**科学性、客观性**无法得到证实，使用者和利益相关者难以对其评分结果做出完整的评价。因此，这部分研究将基于深度学习模型，设计合理的注意力机制，动态追踪ESG新闻不同内容（词语、段落、文章等）对证券市场波动的影响，致力于为ESG新闻中不同内容的重要程度提供合理的解释，来增强利益相关者对企业可持续发展的科学判断。

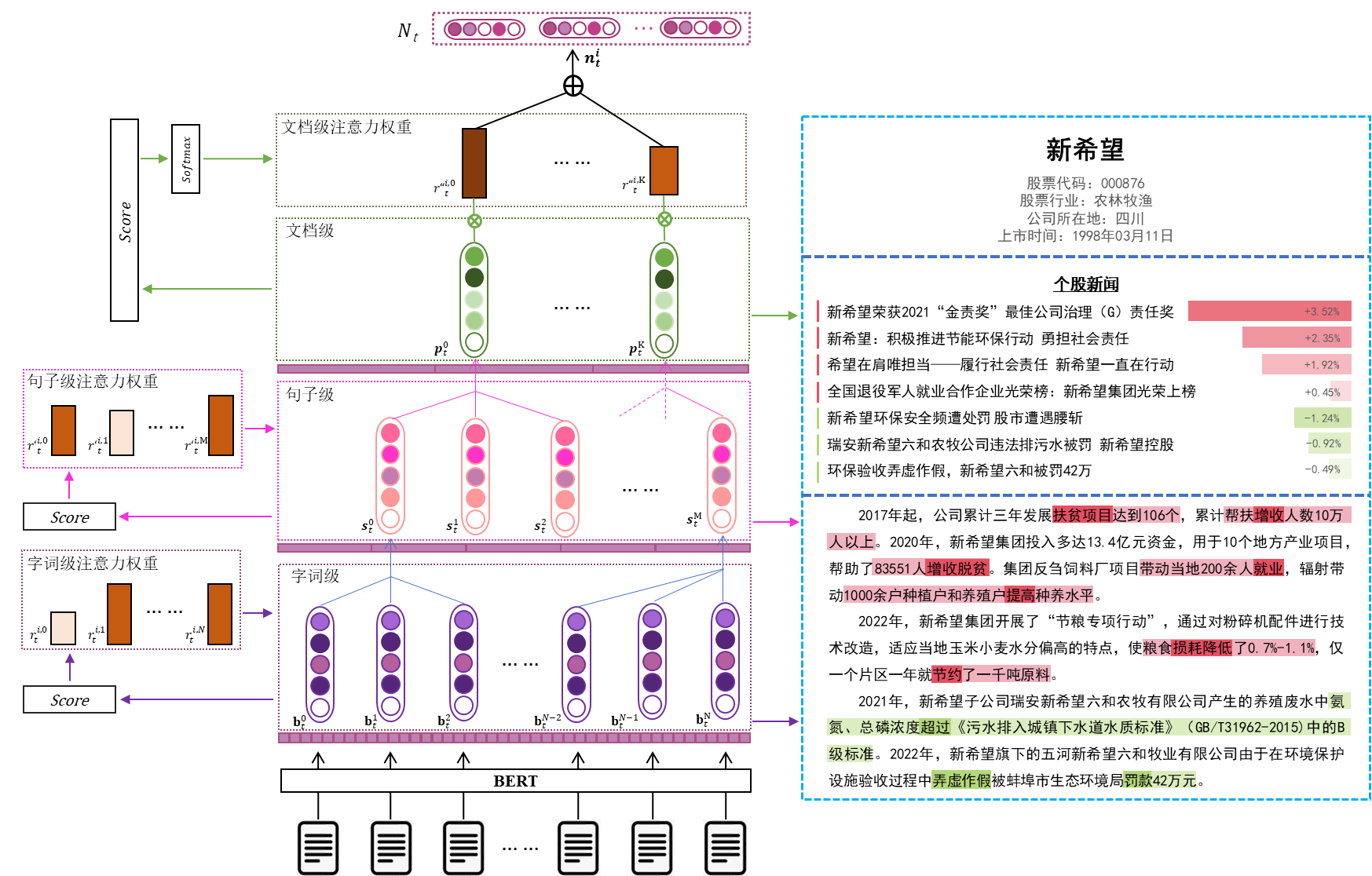


图 7. 多层次注意力机制的ESG新闻内容重要程度可视化

可信人工智能是现代AI技术一个重要的突破点。为了避免传统机器学习“黑盒子”的缺陷，研究人员致力于为机器学习预测到的结构，提供合理的解释。在本研究中，我们创新性地在深度学习框架中，引入多层注意力机制（Mutil- attention Mechansim）来应对非线性因果推导挑战，可以显示 时刻，不同ESG新闻的重要性，并且同一ESG新闻内容中，不同内容（字，词，句）的重要性，即对证券市场波动的影响重要性，从而为更加科学合理地测度和应用ESG新闻所蕴含的企业可持续信息提供可能。

具体而言，首先，采用基于Transformer架构的预训练模型BERT对媒体文本信息进行向量化表示（Devlin，2018）。模型是将每一个字映射到一个数值型的高维度空间向量中。通过该种方式的嵌入，可以保证在这个映射空间中，两个相似含义的字可以有更加近的空间距离。比如，在新闻文字中，“涨”和“升”之间的距离会比“涨”和“降”之间的距离更近。如图7所示，在本申报项目中，我们采用多层注意力机制来挖掘不同文本内容对证券波动的影响。通过BERT模型对将每一个字映射为一个高维向量后，将字聚集为句，引入注意力评分机制，衡量每个句的重要性。然后，将句聚集文档，引入注意力评分机制，衡量每篇文章的重要性，最后，使用方程来对其进行归一化，来获取相对重要性表示。通过训练后，我们就可以利用每一层的Score 评分来确定字、词、句、文档的重要性，通过不同颜色的深浅来显示其对最后预测结果的影响力度。

CSMAR数据库下载

1. 现有ESG评级存在“透明度不足，深度披露不够，导致公信力降低。”缺陷。
2. 专家凭借主观经验给出的权重
3. 评级机构不公开评级的过程
4. 新闻能够弥补这个缺陷，具体思路
5. 用多层注意力机制得到了每篇文章影响力的大小，文章影响力权重。（高中低）
6. 通过聚类算法，发现“高”的里面，有10%的新闻都提到了污染问题。
7. 高60%，中30%，低10%

60%\*10%=0.06污染的权重→ESG评分

高（词语层面的权重）

1. 污染10%
2. 土地使用 9%
3. 员工健康 8.5%
4. 产品质检7%

100%

|  |  |
| --- | --- |
| 正 | 负 |
| 高return>2%  中1%<return<2 %  低0<return<1% | 高return <-2%  中-2%<return<-1 %  低-1 % < return<0 |

E类新闻

|  |  |
| --- | --- |
| 正 | 负 |
| 高return>2%  中1%<return<2 %  低0<return<1% | 高return <-2%  中-2%<return<-1 %  低-1 % < return<0 |

S类

|  |  |
| --- | --- |
| 正 | 负 |
| 高return>2%  中1%<return<2 %  低0<return<1% | 高return <-2%  中-2%<return<-1 %  低-1 % < return<0 |

G类

|  |  |
| --- | --- |
| 正 | 负 |
| 高return>2%  中1%<return<2 %  低0<return<1% | 高return <-2%  中-2%<return<-1 %  低-1 % < return<0 |

60%\*10%=0.06污染的权重→ESG评分

2018开始重视ESG

1. 前后评分的提高，说明企业在重视，企业被教育了
2. 前后影响力的提高，说明投资者在重视起来了，投资者被教育了
3. 政策市

中港对比：

1. 中2023的影响力小，港的2000影响力大。投资者被教育了