软件学报 ISSN 1000-9825, CODEN RUXUEW
2025,36(5):2064−2078 [doi: 10.13328/j.cnki.jos.007180] [CSTR: 32375.14.jos.007180]

E-mail: jos@iscas.ac.cn
http://www.jos.org.cn
Tel: +86-10-62562563

基于高斯混合多层自编码器的情感漂移检测模型*

张文跃[1], 李 旸[2], 王素格[3], 廖 健[3]

[1](山西财经大学 信息学院, 山西 太原 030006)

[2](山西财经大学 金融学院, 山西 太原 030006)

[3](山西大学 计算机与信息技术学院, 山西 太原 030006)

通信作者: 王素格, E-mail: wsg@sxu.edu.cn

摘 要: 社交网络情感数据最为显著的特征是其动态性. 针对群体文本情感漂移分析任务, 提出一种高斯混合多层自编码器 (GHVAE) 用于情感漂移检测. GHVAE 将高斯混合分布作为潜在分布的假设先验, 对应潜在分布的多中心性质从而提高模型性能. 此外, 还对原始 HVAE 模型内建的漂移度量算法进行改进, 改善了高漂移值之间过于接近导致分类性能下降的问题. 采用多项对照实验和消融实验用于验证 GHVAE 的性能, 实验结果显示新模型的创新点为其漂移检测表现带来了提升.

关键词: 情感漂移; 层次变分自编码器; 情感元分布; 漂移度量; 高斯混合

中图法分类号: TP18

中文引用格式: 张文跃, 李旸, 王素格, 廖健. 基于高斯混合多层自编码器的情感漂移检测模型. 软件学报, 2025, 36(5): 2064–2078. http://www.jos.org.cn/1000-9825/7180.htm

英文引用格式: Zhang WY, Li Y, Wang SG, Liao J. Gaussian Mixture Based Hierarchical Auto-encoder Model for Sentiment Drift Detection. Ruan Jian Xue Bao/Journal of Software, 2025, 36(5): 2064–2078 (in Chinese). http://www.jos.org.cn/1000-9825/7180.htm

Gaussian Mixture Based Hierarchical Auto-encoder Model for Sentiment Drift Detection

ZHANG Wen-Yue[1], LI Yang[2], WANG Su-Ge[3], LIAO Jian[3]

[1](School of Information, Shanxi University of Finance and Economic, Taiyuan 030006, China)

[2](School of Finance, Shanxi University of Finance and Economics, Taiyuan 030006, China)

[3](School of Computer and Information Technology, Shanxi University, Taiyuan 030006, China)

Abstract: The most significant feature of social network sentiment data is its dynamic nature. Tackling public sentiment drift analysis, this study proposes a Gaussian mixture based hierarchical variational auto-encoder (GHVAE) model for detecting sentiment drifts. Specifically, the GHVAE applies Gaussian mixture distribution as the prior assumption of latent distribution, which corresponds to the multi-center of the latent distribution property to improve model performances. Moreover, the built-in drift measurement algorithm in the original HVAE model is revised to enlarge the distances among big drift scores and improve the classification performance. Several contrast and ablation experiments are implemented to validate the performance of the GHVAE. The results indicate the novelties of the GHVAE bring improvement in sentiment drift detection.

Key words: sentiment drift; hierarchical variational auto-encoder (HVAE); sentiment meta-distribution; drift measure; Gaussian mixture

随着互联网 Web 2.0 的深入发展, 社交网络用户在网上表达他们对不同事件的看法变得轻松而又频繁, 由此产生了海量的群体文本情感数据亟待合适工具进行处理. 群体情感的一个重要特征是其动态性, 即群体情感分布随着时间推移不断变化. 例如, 公众对灾难救援行动的情感是动态变化的, 当出现“救援难度增大”“遇难者人数增

* 基金项目: 国家自然科学基金 (62376143, 62106130); 山西省基础研究计划青年科学研究项目 (202203021212495, 20210302124084); 山西省高等学校科技创新项目 (2021L283, 2021L284); 山西省基础研究计划 (202303021211021); CCF-智谱 AI 大模型基金 (CCF-Zhipu202310)
收稿时间: 2023-12-09; 修改时间: 2024-01-23; 采用时间: 2024-03-03; jos 在线出版时间: 2024-08-21
CNKI 网络首发时间: 2024-08-22

加”等消息时正面情感比率会下降, 而当得知“多方救援物资供应充足”“被困人员安然无恙”等消息时则会上升. 检测群体情感漂移对社会各界 (如政府机构、公司、新闻机构, 甚至个人) 均具有重要实用价值, 根据检测结论相关方可以采取积极行动, 例如: 公众的情感漂移可能会影响股市[1], 投资者可以利用这些关键漂移信息来决定或调整投资策略; 或者, 如果一家企业能够快速捕捉用户对其产品的情感漂移, 便可以实现及时改进产品设计或调整销售策略; 此外, 各国政府也希望监测公众对他们制定的政策是否发生情感上的漂移.

通常群体情感信息隐藏在时间顺序的文档中, 如推文、微博、论坛和新闻网站的评论等. 这些有序数据需要根据它们到达的具体时间段 (例如, 按照天、小时或分钟等) 被分成互不重叠的集合, 其中每个时间段内包含多个文档. 通过比较不同时间段之间的情感差异实现自动检测群体情感漂移. 群体情感漂移检测任务的核心问题是如何提高检测的准确性, 又可以具体分为两个方面的挑战, 即如何设计更适合的漂移检测模型以及与其配合的漂移度量.

近年来, 针对群体文本情感漂移检测, Zhang 等人[2]提出了 HVAE 模型. 它将原有的 VAE 模型扩展为一个由输入层、潜在分布层和元分布层组成的 3 层结构, 分别对应于文档级、时段级和历史窗口这 3 种数据粒度. HVAE 模型假设潜在分布及其元分布均为高斯分布. 由于历史窗口保存的时段数量有限, 难以满足中心极限定理的成立条件, 所以使用单一中心的高斯分布相比多中心的高斯混合分布缺少对历史信息更精准的拟合能力. 此外, HVAE 模型采用的漂移度量算法 ADD2 具有高灵敏度, 当情感频繁波动即各时段的度量值在高区间 (接近 1) 时过于相近, 从而模糊了漂移判断边界而导致性能下降, 因此需要增加高区间度量值的区分度.

针对上述两个方面的挑战, 本文提出了一种用于情感漂移检测的基于高斯混合的分层变分自动编码器 (GHVAE) 模型, 具体贡献总结如下.

(1) 提出了 GHVAE 模型, 该模型包含 3 层结构, 并且在顶层使用高斯混合作为元分布, 有助于提高模型在时段窗口较小时的性能.

(2) 本文提出的 GHVAE 模型使用指数变换改进了 ADD2 算法[2], 提出 Ex_ADD 度量算法以解决漂移边界模糊问题.

(3) 采用多种实验以验证新的 GHVAE 模型在不同数据集 (包括人工数据和真实数据) 中的有效性. 此外, 增加两个案例研究用于定性评估. 实验结果证明了 GHVAE 模型的各项创新确实提高了群体情感漂移检测性能.

本文第 1 节介绍情感漂移检测的相关工作. 第 2 节中详细介绍本文提出的 GHVAE 模型. 第 3 节介绍一系列实验设计和实现方案. 第 4 节展示所有实验结果并对其进行分析. 最后在第 5 节对全文工作进行总结以及对未来改进方向进行展望.

1 相关工作

一个完整的情感漂移检测过程由 3 个关键子任务组成, 即情感建模、漂移检测和漂移自适应. 情感建模的主要目的是从给定的输入序列文档中提取情感分布信息. 漂移检测专注于计算历史数据和新数据之间的分布差异, 然后将其用于判断漂移是否发生. 最后, 当漂移发生时, 漂移自适应将根据新数据更新模型.

1.1 情感建模

情感建模的主要目的在于提取或者表现文本情感信息. 已有情感建模工作常见于文档级别的情感极性分类任务[3], 即自动检测给定文档的情感极性. 早期工作多采用基于统计的方法, 首先计算情感词及其上下文的统计信息, 然后决定该文档是属于正类 (positive) 还是负类 (negative). 此外, 也有一些工作建立模型将情感信息用于对话场景[4]以及摘要生成[5]任务中.

除基于统计方法外, 还有一些工作采用概率图方法提取情感信息. 例如, Xiong 等人[6]通过增加一个表示情感的潜在变量来扩展 LDA, 并提出了用于产品评论的词对 (word-pair) 情感-主题 (sentiment-topic) 模型. Catal 等人[7]提出了一种情感分类的集成方法, 该方法集成了多个分类器, 包括朴素贝叶斯、SVM 和 Bagging 等. 近年来, 深度学习由于其显著的拟合能力在该研究领域变得愈加流行. Yang 等人[8]实现了一种具有注意力机制的神经网络, 用

于提取方面级情感分类的跨域信息. Chen 等人[9]提出了一种深度学习框架, 该框架利用递归图卷积网络来识别情感极性. 赵传君等人[10]对跨语言情感分类方法进行了阶段性总结.

深度学习类的模型通常属于“黑盒模型”缺少可解释性, 且训练成本往往比较高. 对此, 人们提出了 VAE 类型的模型, 它将神经网络与概率图相结合, 从而产生相对轻量级和可解释性较好的模型. 原始 VAE 是由 Kingma 等人[11]提出的一种无监督模型, 它首先应用编码模块 (*Encoder*) 学习潜在变量, 然后使用解码模块 (*Decoder*) 拟合输入数据分布. 已经有很多工作提出了多种 VAE 类型的模型用于应对各种情感分析任务, 例如, 分类[12−14], 对话生成[15−17], 自动意见摘要[18,19]等. VAE 类型的模型一般被认为具有更好的泛化性能, 因为它可以避免训练数据的过度拟合[20]. 目前标准 VAE 模型已扩展到两层结构模型[21−24], 这些模型两层潜在参数之间采用“一对一”形式的配对. 在群体情感漂移分析任务中, 由于存在“文档-时段-历史窗口” 3 层结构, 其潜在参数更像“多对一”模式 (如多个情感文档对应于一个时间段), 因此已有的模型不太适合本文的任务.

本文提出的情感建模与上述已有工作不同, 在本文工作中增加了由多个文档组成的时间段层以及多个时间段组成的窗口层, 形成“金字塔型”的“多对一”层次结构. 新模型不会聚焦于研究文档级别的情感分类, 而是学习情感分布信息, 即属于同一时间段/窗口的多个文档的正负极性的统计参数.

1.2 漂移检测

漂移检测任务需要先计算新数据和历史数据之间的差异 (包括分布差异[25]), 然后将其量化作为指标用于判断变化是否足够显著以宣布漂移已经发生. 漂移检测方法可以分为以下 3 类: 1) 序列分析 (SA), 主要出现在一些早期的研究工作中[26,27], 该类方法预先设置了一个固定阈值, 当新到达的数据和历史数据之间的差异度大于阈值时触发漂移警报, 虽然 SA 方法简单有效但必须设置合理的固定阈值, 这一点限制了它的应用范围; 2) 统计过程控制 (SPC), 通过积累漂移指示的统计信息 (如均值和标准差等) 建立检测机制. 在此之后有很多研究工作[28−30]采用了 SPC 检测机制, 与 SA 相比, SPC 通常更加灵活; 3) 双窗口比较 (CTW) 模型, 它分配一个固定大小的窗口来保存历史信息, 而另一个窗口则在新到达的数据上滑动, 通过比较两个窗口之间的数据分布来判断漂移是否发生, 例如 Bifet 等人[31]提出了 ADWIN 方法, 该方法利用 Hoeffding 边界来确定漂移位置; 在 Yu 等人[32]的工作中, 构建了一个用于检测数据漂移的两层结构模型, 其第 1 层是 CTW 方法, 此外还有 Nguyen 等人[33]将 ADWIN 与变分推断相结合, 实现了一种新的在线分类系统.

1.3 漂移自适应

漂移自适应机制在漂移发生时用新数据更新模型, 通常包含“盲目 (Blind)”和“通知 (Informed)”两种策略. 为了实现自适应, 许多方法采用了“盲目策略”, 即一旦新数据到达不进行漂移检测直接更新模型. 由于需要不断更新模型, 因而这种策略只适用于简单、轻量级、更新成本不高的模型. 例如, Krawczyk 等人[34]根据数据的到达顺序对所有数据进行加权, 并强化那些新到达的数据权重, 这些处理后数据都被组装在一起进行模型训练. Krawczyk 与他人的另一项研究工作[35]提出了一种能够处理漂移引起的突然变化的更新机制, 在其他方法[36−38]中也提出了一些类似的方法. 第 2 种自适应策略是“通知策略”, 它通常包含一个触发器并根据漂移检测结果决定是否触发重新训练或更新模型操作[39−41]. Iosifidis 等人[42]将这种通知策略应用于流情感分类.

2 群体情感漂移检测模型

本文提出用于检测群体情感漂移的模型是基于高斯混合的分层变分自动编码器 (GHVAE), 主要由 3 部分组成: GHVAE 模型、漂移检测模块、漂移自适应机制. 具体而言, GHVAE 模型从历史数据中学习情感分布, 采用拟合历史数据的方法进行验证, 然后重新到数据中学习新的情感分布. 在漂移检测模块, 定量计算历史情感分布和新情感分布之间的分布差异, 并在差异程度过大时认定发生情感漂移. 最后, 漂移自适应机制使用新到数据更新历史数据分布.

2.1 GHVAE 模型

关于情感漂移, 要得到更稳定和合理的检测结果, 其关键条件是将检测粒度设置在多文档级别, 因为个人情感

是不稳定的, 更容易受到偶然和不可追踪的因素的影响, 例如最近的经历、个人情绪等. 这种不稳定性导致单一用户、单文档级别的情感不适合指导漂移检测分析. 相反, 由于随机影响因素被中和, 基于多文档进行漂移检测能够揭示群体在特定时间段内潜在的相对稳定的态度.

本文根据情感数据的时间戳对其进行分割并建立如图 1 所示的多层结构. 在这种结构中, 一组情感数据属于一个时间段, 几个时间段组成一个窗口, 其中时序情感文档 (表示为 s) 根据它们所在时间段被分割成块 (表示为 z'), z 表示几个时间段的集合用于表示一段时间窗口内的历史数据.

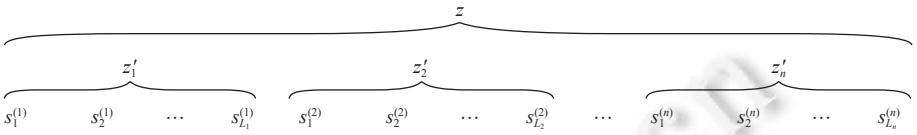
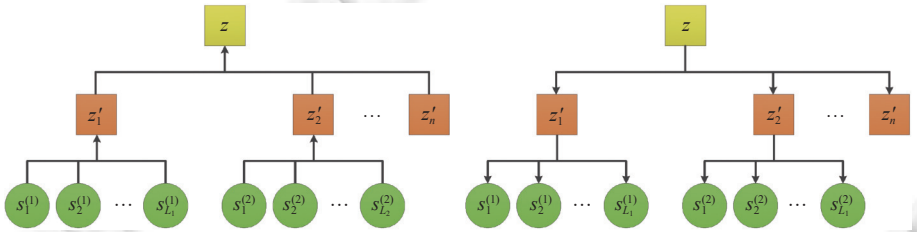


图 1　依据时间将时序情感文档组织为层次结构

相应地, GHVAE 采用 3 级层次结构来对应序列情感文档的多层潜在参数. 图 2 展示了 GHVAE 模型的结构, 其中 3 层分别对应于不同粒度的信息. 所有输入的群体情感都保存于最底层, 所有时段的潜在分布位于中间层并用 z' 表示, n 个历史时段的情感保存于长度为 n 滑动窗口 W, 这些历史情感的元分布位于顶层并用 z 表示. 从下到上的 3 层分别命名为文档层、时段层和窗口层. 具体而言, 窗口中第 i 个时间段的情感表示为 $s_{1:L_i}^{(i)}$, 而 $s_{1:L_{\text{new}}}$ 是新到达的文档的集合. 模型中的所有标注都列在表 1 中.



(a) 模型编码过程　　(b) 模型解码过程

图 2　GHVAE 模型结构

表 1　本文标注

表示符号	含义
s、S、S_{new}	s表示文本情感文档, 其上标表示该文档到达时段序号, S 代表窗口内的文档集合, S_{new} 是新时段内的文档集合
L_i	表示序号为 i 的时段内的文档数量
z'、z	分别表示对应时段和窗口的隐藏变量
W	长度为 n 的窗口, 保存着一些时段组的文档 (对应于 z)
θ、φ	分别表示解码器和编码器的参数
$(\mu'_{\phi_i},\sigma'_{\phi_i})$、$(M_\phi,\Sigma_\phi,\pi_\phi)$	两组参数分别表示由编码器生成的隐藏分布和元分布参数, 它们分别来自于 i 号时段和窗口 W
$(\mu'_{\theta_i},\sigma'_{\theta_i})$、$(\mu_\theta,\sigma_\theta)$	两组参数分别表示由解码器生成的隐藏分布和元分布参数, 它们分别来自于 i 号时段和窗口 W
ε、ε'	分别针对 z 和 z' 使用的随机采样噪声
α、λ	漂移检测的调整参数

新模型包含两个过程分别由两个对应组件实现, 即编码器 (*Encoder*) 和解码器 (*Decoder*). 在编码过程中, 首先由模型文档层接收一个时段的情感文档. 然后, 时段级的潜在分布信息被编码为表示向量 z', 且 $z' \sim z' \mid S$ 如公式 (1) 和公式 (2) 所示. 潜在分布的所有参数都是根据它们对应的输入情感生成的. 选择高斯作为潜在分布是因为输入数据的独立性假设符合中心极限定理.

$$(\mu'_{\phi_i},\sigma'_{\phi_i}) = Encoder_{\phi}(s^{(i)}_{1:L_i}) \tag{1}$$

$$z'_i \mid s^{(i)}_{1:L_i} \sim N(\mu'_{\phi_i},\sigma'^2_{\phi_i}) \tag{2}$$

最后, 窗口层提取窗口中所有时段对应分布的潜在元分布, 用 z 表示, 即 $z \sim z \mid z'_{1:n}$. 窗口级的编码过程详见公式 (3) 和公式 (4). 根据时段潜在分布表示 $z'_{1:n}$ 计算元分布的参数. GHVAE 和原始 HVAE 之间最大的不同在于元分布的设计, 本文采用高斯混合作为 $z \mid z'_{1:n}$ 分布. 由于窗口的长度相对较小, 其元分布更可能是多峰值形式而不是如同单个高斯分布一样只有一个中心, 因此采用高斯混合分布更适合作为窗口级情感分布. 所有潜在分布都由参数为 ϕ 的编码器生成, 新模型使用神经网络作为编码器. 此外, 元分布的参数 $\pi_{\phi} = \{\pi_1,\pi_2,\ldots,\pi_n\}$ 是一个长度 n 的归一化向量, 并从时段的集合中通过公式 (5) 计算而来, 其中函数 $f(\cdot)$ 是 MLP.

$$(M_{\phi},\Sigma_{\phi}) = Encoder_{\phi}(z'_{1:n}) \tag{3}$$

$$z \mid z'_{1:n} \sim MN(M_{\phi},\Sigma_{\phi},\pi_{\phi}) \tag{4}$$

$$\pi_{\phi} = Softmax(f(z'_{1:n})) \tag{5}$$

另一方面, 解码器模块使用编码器学习而来的隐藏分布拟合输入数据再进行拟合效果评估. 通常而言, 学到更好的潜在分布和元分布应该有助于更好地拟合输入数据, 即通过随机梯度下降训练缩小先验分布和后验分布之间的差距. 在解码器模块中, 如公式 (6)–公式 (8) 所示, 输入情感和时段的条件分布是高斯分布, 解码器分别根据其对应的时段和窗口来评估其参数. 本文采用神经网络作为解码器, 其中 θ 表示网络的参数.

$$(\mu_{\theta},\sigma_{\theta}) = Decoder_{\theta}(z) \tag{6}$$

$$z'_i \mid z \sim N(\mu_{\theta},\sigma^2_{\theta}) \tag{7}$$

$$(\mu'_{\theta_i},\sigma'_{\theta_i}) = Decoder_{\theta}(z'_i) \tag{8}$$

$$s^{(i)}_j \mid z'_i \sim N(\mu'_{\theta_i},\sigma'^2_{\theta_i}) \tag{9}$$

综合编码和解码模块, 模型的训练目标是最大限度地提高输入情感的对数似然, 以获得更好的隐藏分布参数, 如公式 (10) 所示:

$$\log p_{\theta}(S) = \log\left(\frac{p_{\theta}(S,z'_{1:n},z)}{p_{\theta}(z'_{1:n},z \mid S)}\right) \tag{10}$$

$S = \left\{s^{(1)}_{1:L_1}, s^{(2)}_{1:L_2},\ldots,s^{(n)}_{1:L_n}\right\}$ 是窗口中的历史数据, $\log p_{\theta}(S)$ 是拟合输入的对数似然. 由于 $p_{\theta}(z'_{1:n},z \mid S)$ 是不可解的, 所以使用变分推断来解决这一问题, 从而将对数似然公式 (10) 的右部分改为公式 (11) 后经整理得到公式 (12).

$$\log\left(\frac{p_{\theta}(S,z'_{1:n},z)}{p_{\theta}(z'_{1:n},z \mid S)}\right) = E_{q_{\phi}(z'_{1:n},z|S)}\left[\log\left(\frac{p_{\theta}(S,z'_{1:n},z)q_{\phi}(z'_{1:n},z \mid S)}{q_{\phi}(z'_{1:n},z \mid S)p_{\theta}(z'_{1:n},z \mid S)}\right)\right] \tag{11}$$

$$E_{q_{\phi}(z'_{1:n},z|S)}\left[\log\left(\frac{p_{\theta}(S \mid z'_{1:n})p_{\theta}(z'_{1:n} \mid z)p_{\theta}(z)}{q_{\phi}(z'_{1:n},z \mid S)}\right)\right] + D_{\mathrm{KL}}\left[q_{\phi}(z'_{1:n},z \mid S)\|p_{\theta}(z'_{1:n},z \mid S)\right] \tag{12}$$

公式 (12) 中的第 2 项是 $q_{\phi}(z'_{1:n},z \mid S)$ 和 $p_{\theta}(z'_{1:n} \mid z)$ 之间的 Kullback-Leibler (KL) 散度. 由于 KL 散度非负, GHVAE 模型的目标函数被转换为证据下界 (evidence lower bound, ELBO) 形式:

$$\log p(S) \geqslant E_{q_{\phi}(z'_{1:n},z|S)}\left[\begin{array}{l}\overbrace{\log p(z)+\log p_{\theta}(z'_{1:n} \mid z)+\log p_{\theta}(S \mid z'_{1:n})}^{\text{Decode}} \\ \underbrace{-\log q_{\phi}(z \mid z'_{1:n})-\log q_{\phi}(z'_{1:n} \mid S)}_{\text{Encode}}\end{array}\right] \tag{13}$$

2.2 分布差异度量

漂移度量算法是一种探测历史数据和新到达数据之间差异的策略. 经过训练的 GHVAE 模型能够揭示历史数据分布信息, 这些信息被分配于目标函数的不同部分. 本文选择两个子模块, 即 $z' \mid S$ 和 $z' \mid z$ 用于计算漂移度量. 首先, 根据公式 (1) 提取新到达时段 S_{new} 中的数据及其潜在分布参数 z'_{new}, 命名为 μ_{new} 和 σ_{new}. 然后计算历史窗口下的条件分布参数, μ 和 σ, 这两个参数是通过公式 (6) 产生而来. 分布 $N(\mu_{\mathrm{new}},\sigma^2_{\mathrm{new}})$ 和 $N(\mu,\sigma^2)$ 之间的量化差异, 即

为漂移指标用 p 表示, p 值越大说明漂移越显著.

前面提到的 HVAE 模型使用 ADD 算法 (累积分布差), 该算法用到了两种分布的均值和变量. ADD 中的漂移度量为公式 (14), 用图 3 表示两个分布之间的 ADD 差异.

$$p' = \frac{1}{2}\int \left| N(x;\mu_{\text{new}},\sigma^2_{\text{new}}) - N(x;\mu_\theta,\sigma^2_\theta) \right| \mathrm{d}x \tag{14}$$

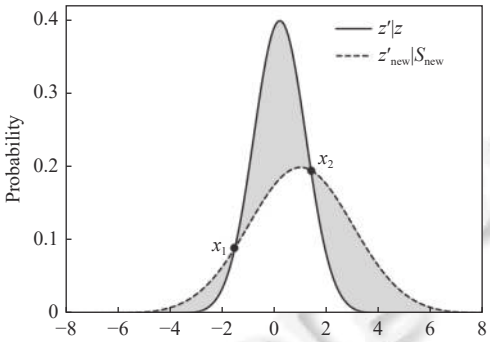


图 3 两个分布之间 ADD 差异

公式 (15) 的积分不可解因而无法直接计算, 解决这个问题的方法是分段. 首先, 根据求解方程找到两条分布曲线的交点. 由于两个分布的联立方程是二次方程所以最多有两个交点, 分别命名为 x_1 和 x_2 (当只有一个交点时, $x_1 = x_2$, 并在元素上令 $x_1 \leqslant x_2$). 交点将曲线分割成 3 段 (或者 2 段), 对其累积概率密度差进行求和, 然后归一化到 $[0,1]$, 如公式 (15).

$$p' = \left| \int_{-\infty}^{x_1} \left[N(x;\mu_{\text{new}},\sigma^2_{\text{new}}) - N(x;\mu_\theta,\sigma^2_\theta) \right] \mathrm{d}x \right| - \left| \int_{-\infty}^{x_2} \left[N(x;\mu_{\text{new}},\sigma^2_{\text{new}}) - N(x;\mu_\theta,\sigma^2_\theta) \right] \mathrm{d}x \right| \tag{15}$$

尽管 ADD 指标 p' 反映了两种分布之间的差异, 但不建议作为最终指标应用于漂移检测 (即令 $p = p'$). 根据一些实验结果显示, 当一段时期输入的情感波动频繁时 (即舆论分化严重时期), 该段时期内的所有 ADD 得分 p' 会非常趋近于 1. 因此各时刻漂移指标 p 过于相似, 类 (漂移与否) 边界的差距过小, 导致漂移检测性能下降. 为了缓解这一问题, HVAE 模型将 ADD 指标 p' 的平方作为最终差异指标 (即 $p = p'^2$), 扩大了接近 1 的 ADD 分数之间的差距, 并将其命名为“ADD2”. 为了进一步改进漂移度量, 在得到两个情感分布之间的 ADD 分数后, 本文采用指数函数来改进原始的漂移测量算法命名为“Ex_ADD”, 具体如公式 (16) 所示. 公式 (16) 中的 p 是通过 Ex_ ADD 测量的最终漂移分数.

$$p = \exp[\lambda(p' - 1)] \tag{16}$$

后文图 4 展示了 ADD 和 Ex_ADD 的直观对比, 其中曲线表示与 ADD 分数 (p') 相对应的最终漂移分数 (p) 所有漂移指标都限制在 [0, 1] 中. 当 p' 接近 1 时, λ 大于 2 的 Ex_ ADD 曲线的斜率明显大于 ADD2, 从而扩大了高区间分数之间的差距, 使类边界 (漂移与否) 更容易区分. 此外, 参数 λ 能够调整放大的效果, λ 越大, 分数接近 1 的曲线斜率越来越大, 因而效果越显著. 这种调整的作用效果不是一直有效的, 因为随着 λ 的增长, 低分区的斜率将如图 4 所示变小因而低分 p 之间的差异也将缩小. 此外, 当 λ 提高到一定程度后, 高分区 p 将足以区分彼此, 再继续增加 λ 值类边界也不会变化因而不会再影响检测结果.

2.3 漂移自适应

当发生突发事件或公众舆论开始改变时, 顺序到达的文本情感可能会显著波动, 在这种情况下, 当前模型不再适合新的数据. 因此, 当漂移累积达到一定水平时, 应该对模型进行重新训练. 本文的漂移自适应策略基于“通知”方法, 即当满足条件时激活触发器, 模型开始更新. 新模型采用基于 SPC[43]的漂移检测方法, 漂移自适应采用一种基于比较窗口 (CTW) 的记忆策略, 该策略决定哪些数据可以处理, 哪些数据可以放弃.

漂移自适应策略设置一个 n 大小的滑动窗口 W_k, 用于保存历史信息, 其下标 k 表示与最近一次漂移发生的时

间段间隔. 根据 Ex_ADD 度量算法, 计算在窗口中每个时间段的所有情感漂移度量 p. 是否应该用窗口中的数据更新模型可以被视为一次伯努利实验, 其参数为 $\bar{p}_k = mean(\{p_1^{(k)}, p_2^{(k)}, \ldots, p_n^{(k)}\})$ 和偏差 $\sigma_k = \sqrt{\bar{p}_k(1-\bar{p}_k)/k}$. p 参数是由公式 (17) 得来的漂移度, k 表示它们来自哪个窗口. 当新的周期 p_i 到达时, 窗口移动同时计算新的参数. 漂移自适应在算法 1 中有详细说明.

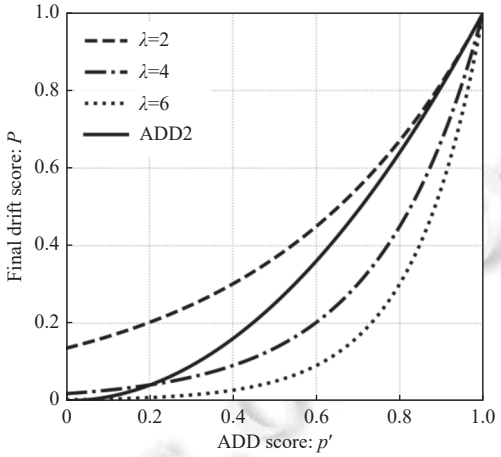


图 4 ADD 与 Ex_ADD 对比

算法 1. GHVAE 漂移自适应.

输入: 窗口和新时段之间的漂移度 p_i; 适应参数 $p_{\min}$, $\sigma_{\min}$;
输出: 发生情感漂移的时段序列.

$\bar{p}_k = mean(W_k)$;
$\sigma_k = \sqrt{\bar{p}_k(1-\bar{p}_k)/k}$;
IF $\bar{p}_k + \sigma_k > p_{\min} + \alpha\sigma_{\min}$ THEN
  $W_k = \{p_i, p_{i+1}, \ldots, p_{i+n}\}$;
  使用 W_k 中的时段数据重新训练模型;
  Set $k = 1$;
ELSE
  $k+ = 1$;
IF $\bar{p}_k + \sigma_k < p_{\min} + \sigma_{\min}$ THEN
  $p_{\min} = \bar{p}_k$;
  $\sigma_{\min} = \sigma_k$;

3 实验设计

为了验证情感漂移分析模型的有效性, 本文使用各种数据和指标进行有针对性的实验. 具体来说, 本文采用两个人工数据集来展示在多种分布和漂移模式场景中的模型性能. 平衡标注的“真实世界 (real-world)”语料库被用于测试实际场景下模型漂移检测的准确性. 此外, 还设置了两个案例研究用于定性评估. 本节分 3 个部分详细介绍了模型验证实验的设计: 数据集、对照模型、实验设置.

3.1 数据集

本文使用了 3 类数据集: 人工数据集、推特情感 140 语料库 (缩写为 S140) 和两个特定事件相关的数据集. 人

工数据集是按照文献 [44] 的方式生成的数据集, 即 GAUSS 和 CIRCLES. 剩余两类数据集都来自推特, 其中 S140 由非特定事件的推文组成, 另一类数据集的推文内容分别关于一场成功的救援行动和一部知名电视剧. 所有数据集的详细信息如下.

GAUSS: 突发 (abrupt) 漂移模式, 带有噪声的数据. 正类和负类数据从两个高斯分布中采样而来, 其参数分别为 $(\mu^+ = (1,1), \sigma^+ = (1,1))$ 和 $(\mu^- = (3,3), \sigma^- = (4,4))$. 每个正 (或负) 数据块由 10 个时间段 (Period) 组成, 每个时间段包含 50 个二维向量. 来自两个类别的数据块彼此交替, 最终构成一个包含 50 个块 (Block) 的数据集. 如图 5 所示.

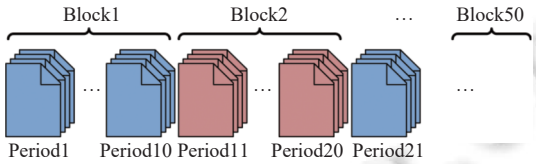


图 5 GAUSS 数据集的结构

CIRCLES: 渐变 (gradual) 漂移模式, 无噪声数据. CIRCLES 数据生成方法遵循已有论文[44], 它从均匀分布中采样, 其参数为 $x \in [0, 1.2]$ 和 $y \in [0, 1]$. 数据有 4 个类别, 它们的类边界均为圆形, 具体见表 2.

表 2 CIRCLES 数据集 4 种类边界描述

圆心	(0.2, 0.5)	(0.4, 0.5)	(0.6, 0.5)	(0.8, 0.5)
半径	0.15	0.2	0.25	0.3

来自 4 个类别的数据相互交替以形成数据块, 其中每个类别有 10 个时间段, 每个时间段由 100 个向量构成. CIRCLES 数据集共包含 40 个块. 如图 6 所示.

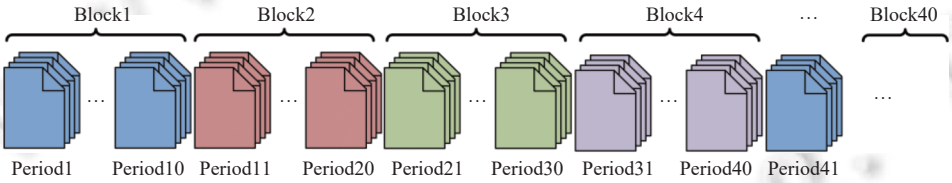


图 6 CIRCLES 数据集的结构

人工数据涵盖了突变和渐变两种漂移模式, 此外还包含了正态和均匀两种数据分布以及有噪声和无噪声两种情况. 处理人工数据可以揭示模型在不同场景下的性能. 除了理想场景外, 本文还使用如下实际数据进行验证.

Sentiment 140 (http://help.sentiment140.com/). 带标签的真实世界数据, 简称 S140. 该数据来自 Twitter Sentiment, 其中包含从 2009-04-06 到 2009-06-25 发布的 1 600 000 条推文. 语料库具有平衡的情感类别, 类别标签被转化为 one-hot 向量作为情感表示. 语料库按小时划分时段, 共有 593 个时段.

ThaiCaveRescue (https://github.com/AlexisZWY/ThaiCaveResuce). 案例研究数据. 对 2018-06-27 至 2018-07-16 期间发布的有关“泰国美人洞救援”相关的推文按照日期进行了爬取, 这些推文包含“#ThaiCaveRescue”“#caverescue thailand”“#ThaiCave”和“#Thamluang”标签. 限制每个日期爬取的推文数量不超过 1 000 条, 最后总共得到 14 348 条推文. 所有推文的情感极性由多项式朴素贝叶斯分类器 (MNB) 标记, 分类器采用 S140 语料库进行训练. 最终 ThaiCaveRescue 数据集包含 20 个时段.

GoTh8 (https://kaggle.com/monogenea/game-of-thrones-twitter). 语料库从与电视剧相关的推文中提取而来, 即《权力的游戏》第 8 季 (GoTh8). 经过预处理后, 共有 755 759 条推文, 发布时间为 2019-04-07 至 2019-05-28, 按日期分为 52 个时段. 采用与 ThaiCaveRescue 相同的分类器标记文档的情感极性.

3.2 对照模型

将已有 state-of-the-art 模型和本文模型进行对照实验, 这些基准模型在表 3 中进行介绍.

表 3 基准模型

模型	描述
HVAE[2]	基于层次结构的VAE模型采用新的漂移度量检测群体情感漂移
VIGO[33]	使用变分编码器学习潜在参数并配合一个内置方法实现漂移检测
IEWMA[30]	改进了具有异方差的传统EWMA方法 (exponentially weighted moving average)[29]
IAMNB[42]	该模型通过监视分类器性能变化实现检测情感流数据漂移

GHVAE 包含两个组件: 编码器和解码器, 它们在模型中扮演不同角色. 为了揭示每个组件对模型性能的影响, 本文设计了 GHAVE 模型的几种变体: No_D 不包含解码器模块, 而 No_E 不包含编码器模块, Plain 只有一层结构以及一个隐藏分布. 消融模型具体内容见表 4.

表 4 消融模型

模型	描述
No_D	无解码器, 即 $p_\theta(S \mid z')$ 和 $p_\theta(z' \mid z)$ 两部分被删去. 在编码过程之后, 使用ADD计算 $z'_{\text{new}} \mid S_{\text{new}}$ 和 $z \mid z'_{1:n}$ 之间的漂移度
No_E	无编码器, 即 $q_\phi(z' \mid S)$ 和 $q_\phi(z \mid z'_{1:n})$ 不经过MLP计算而是直接分别采用输入和隐藏变量的后验
Plain	只有一层元分布 z 的结构. 使用窗口中的所有数据生成 z 后验分布, 即 $z \mid S_{1:n}$, 并将其与新时段的隐藏分布 $z'_{\text{new}} \mid S_{\text{new}}$ 进行对比用于计算漂移度

3.3 实验设置

漂移检测的目的是标记时序数据中发生漂移的时段, 其最终结果是一条带有标记的序列. 因此漂移检测可以被视为一种序列分段 (segmentation) 任务, 且每个分段内部的数据应当是彼此相似的. 假设分段结果越好, 说明来自相同分段数据的同质性就越高. 由此, 在每个分段内用一部分数据训练分类器并预测段内剩余数据, 如果分类器全局准确性高说明序列分段结果较好, 即漂移检测结果较好. 基于该假设本文设计了带标签数据的实验.

GHVAE 的实验设置参考了已有方法[42], 选择累积多项式朴素贝叶斯 (AMNB) 作为情感分类器, 然后根据先验评估 (prequential evaluation) 的原则以“重建 (rebuild)”方式进行漂移自适应[45]: 如果新时段到达后没有发生漂移, 则预测新时段内的文档标签, 然后将其附加到训练集并重新训练分类器; 否则, 将放弃当前累积的训练集, 并使用新窗口数据作为新的训练集重新训练分类器. 在所有实验中, 算法 1 的调整参数 α 为 1. 最后, 将全局准确度用于评估分类任务的结果. 上述实验设置被应用于 CIRCLES 和 S140 数据集的分类实验中.

GAUSS 数据集具有特殊性. 从不同分布 (类别) 中采样的数值可能相似, 甚至可能相同, 使得类别边界非常模糊, 以至于诸如准确率等常规分类任务评估指标不可行. 为了量化模型实验表现, 本文选择曲线下面积 (AUC) 作为度量用于消融实验 (同应用于 CIRCLES 数据集).

一般来讲, 当窗口尺寸较小时会增加时间成本但可能减少漏报增加错报, 当窗口较大时则会相反. 综合考虑各种实验数据的分段数量以及小规模预实验结果, 将人工数据、S140 和定性评估的窗口长度分别设置为 5、20、2 (定性实验数据量较少). 由于输入的情感数据维度数量很小, 神经网络隐层维度设置不需要设置很高, 同时为了控制时间成本保证处理速度将神经网络中隐藏层的维度设置为 100, 在重新参数化中每个潜在参数的样本量为 100. GAUSS 的调整参数 α 为 2, 因为 GAUSS 数据集的漂移很容易被检测到, 而在训练其他数据集时设置为 1. 经过预实验, 当 λ 为 4 时, 可以获得最佳结果. 使用随机梯度下降方法 (SGD) 更新所有参数, 并根据小规模预训练的经验将训练迭代数设置为 50.

4 实验结果与分析

本文使用多个数据集并实现基线模型来验证新模型的优势, 本节分别对新模型和对照模型在各数据集上的实

验结果进行介绍并分析. 实验结果证明新模型结合新情感漂移度量算法可以得到更好的情感漂移检测表现.

4.1 人工数据实验表现

基于人工数据的实验有助于揭示理想场景下的模型性能, 理想场景具有有限的噪声, 涵盖了各种漂移模式和数据分布. 具体而言, GAUSS 数据集具有有限的噪声, 存在突然式漂移, 并从高斯分布中采样, 而 CIRCLES 是无噪声的逐渐漂移, 并从均匀分布中采样. 因此, 涵盖足够消融情况的两个人工数据集的性能表明了 GHVAE 的有效性. 实验结果在表 5 和表 6 中介绍.

表 5 各模型在 GAUSS 数据集上的 AUC 表现

模型		漂移度量		
		ADD	ADD2	Ex_ADD
GHVAE		0.992	0.989	**0.994**
HVAE		0.989	0.993	0.992
Ablations	Plain	0.992	0.990	0.990
	No_D	0.748	0.754	0.797
	No_E	0.984	0.981	0.988

表 6 各模型在 CIRCLES 数据集上的 AUC 表现

模型		漂移度量		
		ADD	ADD2	Ex_ADD
GHVAE		0.639	0.658	**0.662**
HVAE		0.615	0.656	0.660
Ablations	Plain	0.591	0.652	0.657
	No_D	0.513	0.503	0.504
	No_E	0.608	0.651	0.656

正如结果所示, 所有模型使用 GAUSS 数据相比使用 CIRCLES 数据表现明显更好, 因为编码过程和人工数据的采样是互逆的操作, 同时也意味着如果 GAUSS 上的实验结果突出, 表明模型有能力更准确地从输入中提取信息. 此外, GHVAE 不仅在 GAUSS 上的漂移检测表现突出, 在 CIRCLES 上也取得了最好结果 (如表 5 和表 6 中加粗显示), 这证明新模型能够拟合输入并同时表现出良好的鲁棒性.

在所有模型中, No_E 表现最差, 说明编码过程在模型中发挥着更加关键的作用. 在 No_E 中用于比较的分布 (即 $z \mid z'$ 和 $z' \mid S$) 彼此之间有很大不同. 显著的差异使得漂移测度过大且不稳定, 从而导致模型有效性降低. 相反, Plain 方法将一个窗口中的所有数据混合在一起, 并计算总体分布的参数. 因此, Plain 的性能比 No_E 模型好得多, 因为生成的窗口分布与其中的时间段分布相对接近 (相比 No_E). 在所有度量之间的比较中, 本文提出的 Ex_ADD 取得了更高的准确度, 说明指数放大机制是有效的. 根据结果, GHVAE+Ex_ADD 在 CIRCLES 和 GAUSS 数据集上都实现了相对更好的表现. 特别是在 CIRCLES 的实验中, 新模型取得了更高的 AUC 证明了 GHVAE 的有效性.

前面提到的其他 3 个对照模型没有被应用于 AUC 测试, 因为这些模型与评估指标 AUC 不兼容. 为了将这些模型加入对比, 对 CIRCLES 数据额外进行分类实验, 因为 CIRCLES 是带标记的语料库. 表 7 展示了所有模型的分类性能, 其中 GHVAE+Ex_ADD 的分类准确度最高. 模型 IAMNB[42]不支持数值型数据所以不参与对照实验. VIGO[33]和 IEWMA[30]模型的度量算法采用内建式度量算法, 不支持 ADD 类方法, 所以表 7 中记录的是这两种模型原始方法的实验结果. GHVAE 的准确度高于 HVAE 模型, 表明 Ex_ADD 有助于提高漂移检测准度. 此外, GHVAE 和消融模型之间的对比结果体现了结构完整性的重要性, 即编码器和解码器的合作是 GHVAE 的关键.

表 7 各模型在 CIRCLES 数据上准确度性能

模型		漂移度量		
		ADD	ADD2	Ex_ADD
GHVAE		0.317	0.358	**0.367**
HVAE		0.315	0.350	0.357
VIGO[33]			0.245	
IEWMA[30]			0.251	
Ablations	Plain	0.265	0.310	0.323
	No_D	0.262	0.270	0.285
	No_E	0.207	0.235	0.251

4.2 S140 数据实验表现

与人工数据不同, S140 语料库中的文档是来自真实用户发布的推文. 此外, 语料库中每个文档都有情感极性

标记, 因此方便对新模型和对照模型进行定量评估.

所有模型和度量的分类性能如表 8 所示. 此外, 添加了一个增量训练分类器, 以验证在没有漂移检测的情况下分类的表现 (即表 8 中“No drift detection”). 根据结果, GHVAE+Ex_ADD 获得了更好的分类准确性, 比在 DS1[42]和其他对比模型中使用同一语料库的最佳结果高出约 10%. 此外, 有漂移检测的模型的性能优于无漂移检测的模型, 这也说明实验的分段假设是有效的.

为了揭示调整参数 λ 的影响, 实验记录了不同 λ 取值下的分类准确度. 在图 7 中, 当 $\lambda=4$ 时达到最高精度, 小于 4 时显示出显著的性能退化, 一定程度下增加高区间度量值区分度有助于更好地检测漂移. $\lambda>4$ 时结果彼此相似, 因为此时再扩大类 (是否漂移) 的差距不再有助于提高性能. 此外, 由于低区间的漂移分数越来越彼此接近, 使得 $\lambda>4$ 时反而出现精度下降的情况.

表 8 各模型在 S140 数据集上的准确度性能

模型		漂移度量		
		ADD	ADD2	Ex_ADD
GHVAE		0.827	0.831	**0.842**
HVAE		0.827	0.833	0.838
VIGO[33]			0.793	
IEWMA[30]			0.789	
IAMNB[42]			0.745	
Ablations	Plain	0.798	0.769	0.795
	No_D	0.801	0.800	0.804
	No_E	0.826	0.829	0.840
No drift detection			0.742	

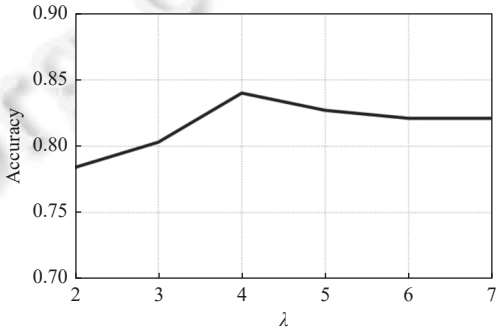


图 7 λ 的取值对准确度的影响

4.3 案例研究: ThaiCaveRescue

在青少年足球队被困洞穴的新闻事件中, 公众情感随着救援行动的推进而剧烈波动. 当救援行动有了重大收获时, 公众情感会高涨, 相反, 当收到坏消息时, 正面情感比例会急剧下降. 因此, 使用救援关键事件和检测到的漂移点之间的相关性对新模型进行定性评估.

在前期, 即 2018 年 6 月 30 日之前, 群体情感仍处于负面 (正面比率保持在 50% 以下), 因为随着男孩们被困在洞穴中的时间不断增长而救援仍然没有令人鼓舞的进展. 随着公众关注度大幅提高, 许多相关人员来到了现场, 如国际救援队、物资支持和媒体等. 2018 年 7 月 1 日, 救援人员发现了 3 号区域并将其作为潜水员的行动基地. 第 2 天, 两名英国潜水员发现了被困男孩, 他们状态都很好. 这一令人兴奋的消息将公众的积极情感提升到了顶峰, 但一个艰巨的挑战出现了, 即如何安全地救出男孩们? 在接下来的几天里, 几项先前拟议的计划但都被否决了, 在此期间正面情感率有所回落. 正如在 2018 年 7 月 6 日出现的低谷一样, 最低沉的公众情感是由一名泰国海豹突击救援人员的死亡新闻引发的. 公众期待已久的好消息出现在 2018 年 7 月 8 日, 救援队成功救出了 4 名男孩, 这将情感曲线提升到了一个积极的水平. 直到 2018 年 7 月 10 日, 当最后 4 个男孩和他们的教练安全脱困时, 情感曲线达到了最高点. 关键事件描述 (根据“Tham Luang 洞穴救援”维基百科页面 (https://wikipedia.org/wiki/Tham_Luangcave_rescue)) 和检测到的漂移时间段的关键词 (基于 TF-IDF 评分) 记录在后文表 9 中.

4.4 案例研究: GoTh8

作为一部非常受欢迎的电视剧, 《权力的游戏》有大量的观众, 自然观众们对它的观点数据是海量的. 图 8 展示的是第 8 季的播出过程中公众的情感随着最新情节的播出而不断变化, 其中横坐标表示播出的剧集编号 (1–6 集), 纵坐标表示烂番茄用户对每一集的所有评价中正面评价的比例. 与 ThaiCaveRescue 案例类似, 本文通过验证检测到的漂移与播出时间之间的相关性定性分析漂移检测结果的合理性.

表 9 漂移节点的关键事件

漂移日期	事件描述	关键词
2018-06-29	Prime Minister Prayut Chan-o-cha visited the search site and told the families of the boys not to give up hope	hope, missing, update, pm, drilling
2018-07-01	The rescuers made some progress, they reached alarge cavern serve as a key base for the divers	missing, predictions, press, clairvoyant, con
2018-07-06	Saman Kunan, a 37-year-old former Thai NavySEAL, died of asphyxiation during the rescue	seal, navy, oxygen, dies, died
2018-07-08	Four boys was reported to have exited	soccer, prayers, praying, news, hope
2018-07-10	The remaining four boys and their coach wererescued	people, amazing, involved, news, saman

模型检测出的漂移时段和新剧集的播出日期之间的同步性通过表 10 表示, 其中还记录了每集的正面评价比例以及漂移日期相对播出日期的远近关系. 通过结果表格可以发现一种模式, 即公众对电视剧的积极情感在新一集播出后会显著下降. 根据表 10, 检测到的 4 个漂移点被证明是播出日期的第 2 天, 由于每一集都会在晚上 8 点左右首播 10 点左右结束, 这个延迟是可以接受的. 此外, 新模型检测到了两个漂移点出现在播出当天而不在普遍出现漂移的第 2 天, 这表明这两个日期群体情感的变化比其他漂移点更快速更显著, 其中一集是剧情转折点而另外一集是整个剧的大结局. 总而言之, 检测到的漂移日期和播出日期之间的相关性表明, 新模型能够有效发现群体情感的变化.

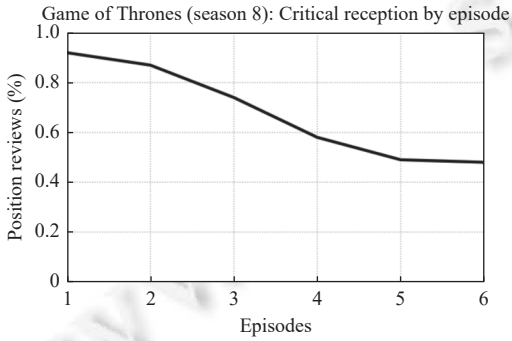


图 8 烂番茄网站评论中正面评价的比率

表 10 检测到的观众情感漂移时段

剧集	播出日期	漂移日期	正面比例	相对播出日
1	2019-04-14	2019-04-15	0.92	次日
2	2019-04-21	2019-04-22	0.87	次日
3	2019-04-28	2019-04-28	0.74	当天
4	2019-05-05	2019-05-06	0.58	次日
5	2019-05-12	2019-05-13	0.49	次日
6	2019-05-19	2019-05-19	0.48	当天

5 结论与展望

作为一项具有挑战性的任务, 群体情感漂移检测需要情感建模和漂移处理. 为了应对这一交叉任务, 本文提出了基于高斯混合的分层变分自编码器 (GHVAE) 模型, 该模型使用了 Ex_ADD 漂移测量算法. 在 GHVAE 部分, 通过时段层学习输入情感的潜在分布, 并应用高斯混合来提取顺序到达的数据之间的元分布. 在使用 GHVAE 学习历史输入信息后, 使用 Ex_ADD 来测量过去和新情感之间的差异. 为了验证模型的有效性, 本文进行了大量的实验. 使用人工和带标签的推特数据进行定量评估, 并使用两个案例研究进行定性评估. 所有实验结果表明, GHVAE 的性能优于对照和消融模型.

在未来的工作中, 该模型可以进一步扩展. 尽管本文中的模型在工程上有一些优势, 但是因为它与情感表示之间缺少非常强的联系, 所以我们仍然希望有新的方法在文本情感表示和漂移分析之间建立更紧密的联系. 这项工作中处理的信息是情感表示, 今后可以用其他类型的信息来代替, 例如主题、上下文等.

References:

[1] Li Q, Wang TJ, Li P, Liu L, Gong QX, Chen YZ. The effect of news and public mood on stock movements. Information Sciences, 2014, 278: 826–840. [doi: 10.1016/j.ins.2014.03.096]

[2] Zhang WY, Li XL, Li Y, Wang SG, Li DY, Liao J, Zheng JX. Public sentiment drift analysis based on hierarchical variational auto-

encoder. In: Proc. of the 2020 Conf. on Empirical Methods in Natural Language Processing (EMNLP). Association for Computational Linguistics, 2020. 3762–3767. [doi: 10.18653/v1/2020.emnlp-main.307]

[3] Ye J, Xiang L, Zong CQ. Aspect-level sentiment classification combining aspect modeling and curriculum learning. Ruan Jian Xue Bao/Journal of Software, 2024, 35(9): 4377–4389 (in Chinese with English abstract). http://www.jos.org.cn/1000-9825/6963.htm [doi: 10.13328/j.cnki.jos.006963]

[4] Zhao YY, Lu X, Zhao WX, Tian YJ, Qin B. Survey on emotional dialogue techniques. Ruan Jian Xue Bao/Journal of Software, 2024, 35(3): 1377–1402 (in Chinese with English abstract). http://www.jos.org.cn/1000-9825/6807.htm [doi: 10.13328/j.cnki.jos.006807]

[5] Liang MY, Li DY, Wang SG, Liao J, Zheng JX, Chen Q. Senti-PG-MMR: Research on generation method of sentimental summary of multi-document travel notes. Journal of Chinese Information Processing, 2022, 36(3): 128–135 (in Chinese with English abstract). [doi: 10.3969/j.issn.1003-0077.2022.03.015]

[6] Xiong SF, Wang KY, Ji DH, Wang BK. A short text sentiment-topic model for product reviews. Neurocomputing, 2018, 297: 94–102. [doi: 10.1016/j.neucom.2018.02.034]

[7] Catal C, Nangir M. A sentiment classification model based on multiple classifiers. Applied Soft Computing, 2017, 50: 135–141. [doi: 10.1016/j.asoc.2016.11.022]

[8] Yang M, Yin WP, Qu Q, Tu WT, Shen Y, Chen XJ. Neural attentive network for cross-domain aspect-level sentiment classification. IEEE Trans. on Affective Computing, 2021, 12(3): 761–775. [doi: 10.1109/TAFFC.2019.2897093]

[9] Chen JJ, Hou HX, Gao J, Ji YT, Bai TG. RGCN: Recurrent graph convolutional networks for target-dependent sentiment analysis. In: Proc. of the 12th Int'l Conf. on Knowledge Science, Engineering and Management. Athens: Springer, 2019. 667–675. [doi: 10.1007/978-3-030-29551-6_59]

[10] Zhao CJ, Wang SG, Li DY. Research progress on cross-domain text sentiment classification. Ruan Jian Xue Bao/Journal of Software, 2020, 31(6): 1723–1746 (in Chinese with English abstract). http://www.jos.org.cn/1000-9825/6029.htm [doi: 10.13328/j.cnki.jos.006029]

[11] Kingma DP, Welling M. Auto-encoding variational Bayes. arXiv:1312.6114, 2014.

[12] Wu CH, Wu FZ, Wu SX, Yuan ZG, Liu JX, Huang YF. Semi-supervised dimensional sentiment analysis with variational autoencoder. Knowledge-based Systems, 2019, 165: 30–39. [doi: 10.1016/j.knosys.2018.11.018]

[13] Fu XH, Wei YZ, Xu F, Wang T, Lu Y, Li JQ, Huang JZ. Semi-supervised aspect-level sentiment classification model based on variational autoencoder. Knowledge-based Systems, 2019, 171: 81–92. [doi: 10.1016/j.knosys.2019.02.008]

[14] Wang YY, Meng XJ, Liu XM. Differentially private recurrent variational autoencoder for text privacy preservation. Mobile Networks and Applications, 2023, 28: 1565–1580. [doi: 10.1007/s11036-023-02096-9]

[15] Kong X, Li BH, Neubig G, Hovy E, Yang YM. An adversarial approach to high-quality, sentiment-controlled neural dialogue generation. arXiv:1901.07129, 2019.

[16] Liang XW, Du JC, Niu TY, Zhou LJ, Xu RF. Knowledge interpolated conditional variational auto-encoder for knowledge grounded dialogues. Applied Sciences, 2023, 13(15): 8707. [doi: 10.3390/APP13158707]

[17] Yang Z, Chen ZH, Cai TC, Wang YF, Liao XW. A survey of deep learning based emotional dialogue response. Chinese Journal of Computers, 2023, 46(12): 2489–2519 (in Chinese with English abstract). [doi: 10.11897/SP.J.1016.2023.02489]

[18] Hoang T, Le H, Quan T. Towards autoencoding variational inference for aspect-based opinion summary. Applied Artificial Intelligence, 2019, 33(9): 796–816. [doi: 10.1080/08839514.2019.1630148]

[19] Xiong Y, Yan MH, Hu X, Ren CH, Tian H. An unsupervised opinion summarization model fused joint attention and dictionary learning. The Journal of Supercomputing, 2023, 79(16): 17759–17783. [doi: 10.1007/s11227-023-05316-x]

[20] Zhao SJ, Ren HY, Yuan A, Song JM, Goodman N, Ermon S. Bias and generalization in deep generative models: An empirical study. In: Proc. of the 32nd Int'l Conf. on Neural Information Processing Systems. Montréal: Curran Associates Inc., 2018. 10815–10824.

[21] Kingma DP, Salimans T, Jozefowicz R, Chen X, Sutskever T, Welling M. Improved variational inference with inverse autoregressive flow. In: Proc. of the 30th Int'l Conf. on Neural Information Processing Systems. Barcelona: Curran Associates Inc., 2016. 4743–4751.

[22] Sønderby CK, Raiko T, Maaløe L, Sønderby SK, Winther O. How to train deep variational autoencoders and probabilistic ladder networks. arXiv:1602.02282, 2016.

[23] Pu YC, Gan Z, Henao R, Yuan X, Li CY, Stevens A, Carin L. Variational autoencoder for deep learning of images, labels and captions. In: Proc. of the 30th Int'l Conf. on Neural Information Processing Systems. Barcelona: Curran Associates Inc., 2016. 2360–2368.

[24] Bai RN, Huang RZ, Qin YB, Chen YP, Lin C. HVAE: A deep generative model via hierarchical variational auto-encoder for multi-view document modeling. Information Sciences, 2023, 623: 40–55. [doi: 10.1016/j.ins.2022.10.052]

[25] Wen YM, Liu S, Miao YQ, Yi XH, Liu CJ. Survey on semi-supervised classification of data streams with concept drifts. Ruan Jian Xue Bao/Journal of Software, 2022, 33(4): 1287–1314 (in Chinese with English abstract). http://www.jos.org.cn/1000-9825/6476.htm [doi: 10.

13328/j.cnki.jos.006476]
[26] Tartakovsky AG, Moustakides GV. State-of-the-art in Bayesian changepoint detection. Sequential Analysis, 2010, 29(2): 125–145. [doi: 10.1080/07474941003740997]
[27] Ikonomovska E, Gama J, Džeroski S. Learning model trees from evolving data streams. Data Mining and Knowledge Discovery, 2011, 23(1): 128–168. [doi: 10.1007/s10618-010-0201-y]
[28] Ross GJ, Adams NM, Tasoulis DK, Hand DJ. Exponentially weighted moving average charts for detecting concept drift. Pattern Recognition Letters, 2012, 33(2): 191–198. [doi: 10.1016/j.patrec.2011.08.019]
[29] Raza H, Prasad G, Li YH. EWMA model based shift-detection methods for detecting covariate shifts in non-stationary environments. Pattern Recognition, 2015, 48(3): 659–669. [doi: 10.1016/j.patcog.2014.07.028]
[30] Zhou D, Liu L, Lai X. The improved EWMA chart for heteroscedasticity process. Annals of Data Science, 2018, 5(1): 21–27. [doi: 10.1007/s40745-017-0133-0]
[31] Bifet A, Gavalda R. Learning from time-changing data with adaptive windowing. In: Proc. of the 7th SIAM Int'l Conf. on Data Mining. Minneapolis: SIAM, 2007. 443–448. [doi: 10.1137/1.9781611972771.4]
[32] Yu SJ, Wang XY, Príncipe JC. Request-and-reverify: Hierarchical hypothesis testing for concept drift detection with expensive labels. In: Proc. of the 27th Int'l Joint Conf. on Artificial Intelligence. Stockholm: IJCAI, 2018. 3033–3039. [doi: 10.24963/ijcai.2018/421]
[33] Nguyen TTT, Nguyen TT, Liew AWC, Wang SL. Variational inference based Bayes online classifiers with concept drift adaptation. Pattern Recognition, 2018, 81: 280–293. [doi: 10.1016/j.patcog.2018.04.007]
[34] Krawczyk B, Woźniak M. One-class classifiers with incremental learning and forgetting for data streams with concept drift. Soft Computing, 2015, 19(12): 3387–3400. [doi: 10.1007/s00500-014-1492-5]
[35] Krawczyk B, Minku LL, Gama J, Stefanowski J, Woźniak M. Ensemble learning for data stream analysis: A survey. Information Fusion, 2017, 37: 132–156. [doi: 10.1016/j.inffus.2017.02.004]
[36] Wang S, Minku LL, Yao X. Resampling-based ensemble methods for online class imbalance learning. IEEE Trans. on Knowledge and Data Engineering, 2015, 27(5): 1356–1368. [doi: 10.1109/TKDE.2014.2345380]
[37] Song G, Ye YM, Zhang HJ, Xu XF, Lau RYK, Liu F. Dynamic clustering forest: An ensemble framework to efficiently classify textual data stream with concept drift. Information Sciences, 2016, 357: 125–143. [doi: 10.1016/j.ins.2016.03.043]
[38] Brzezinski D, Stefanowski J. Reacting to different types of concept drift: The accuracy updated ensemble algorithm. IEEE Trans. on Neural Networks and Learning Systems, 2014, 25(1): 81–94. [doi: 10.1109/TNNLS.2013.2251352]
[39] Lu N, Lu J, Zhang GQ, De Mantaras RL. A concept drift-tolerant case-base editing technique. Artificial Intelligence, 2016, 230: 108–133. [doi: 10.1016/j.artint.2015.09.009]
[40] Bu L, Alippi C, Zhao DB. A pdf-free change detection test based on density difference estimation. IEEE Trans. on Neural Networks and Learning Systems, 2018, 29(2): 324–334. [doi: 10.1109/TNNLS.2016.2619909]
[41] Bu L, Zhao DB, Alippi C. An incremental change detection test based on density difference estimation. IEEE Trans. on Systems, Man, and Cybernetics: Systems, 2017, 47(10): 2714–2726. [doi: 10.1109/TSMC.2017.2682502]
[42] Iosifidis V, Oelschlager A, Ntoutsi E. Sentiment classification over opinionated data streams through informed model adaptation. In: Proc. of the 21st Int'l Conf. on Theory and Practice of Digital Libraries. Thessaloniki: Springer, 2017. 369–381. [doi: 10.1007/978-3-319-67008-9_29]
[43] Bouchachia A. Fuzzy classification in dynamic environments. Soft Computing, 2011, 15(5): 1009–1022. [doi: 10.1007/s00500-010-0657-0]
[44] Gama J, Medas P, Castillo G, Rodrigues P. Learning with drift detection. In: Proc. of the 17th Brazilian Symp. on Artificial Intelligence. Sao Luis: Springer, 2004. 286–295. [doi: 10.1007/978-3-540-28645-5_29]
[45] Bifet A, De Francisci Morales G, Read J, Holmes G, Pfahringer B. Efficient online evaluation of big data stream classifiers. In: Proc. of the 21st ACM SIGKDD Int'l Conf. on Knowledge Discovery and Data Mining. Sydney: ACM, 2015. 59–68. [doi: 10.1145/2783258.2783372]

附中文参考文献:

[3] 叶静, 向露, 宗成庆. 属性建模与课程学习相结合的属性级情感分类方法. 软件学报, 2024, 35(9): 4377–4389. http://www.jos.org.cn/1000-9825/6963.htm [doi: 10.13328/j.cnki.jos.006963]
[4] 赵妍妍, 陆鑫, 赵伟翔, 田一间, 秦兵. 情感对话技术综述. 软件学报, 2024, 35(3): 1377–1402. http://www.jos.org.cn/1000-9825/6807.htm [doi: 10.13328/j.cnki.jos.006807]
[5] 梁梦英, 李德玉, 王素格, 廖健, 郑建兴, 陈千. Senti-PG-MMR: 多文档游记情感摘要生成方法. 中文信息学报, 2022, 36(3): 128–135.

[doi: 10.3969/j.issn.1003-0077.2022.03.015]
[10] 赵传君, 王素格, 李德玉. 跨领域文本情感分类研究进展. 软件学报, 2020, 31(6): 1723–1746. http://www.jos.org.cn/1000-9825/6029.htm [doi: 10.13328/j.cnki.jos.006029]
[17] 杨州, 陈志豪, 蔡铁城, 王宇峰, 廖祥文. 基于深度学习的情感对话响应综述. 计算机学报, 2023, 46(12): 2489–2519. [doi: 10.11897/SP.J.1016.2023.02489]
[25] 文益民, 刘帅, 缪裕青, 易新河, 刘长杰. 概念漂移数据流半监督分类综述. 软件学报, 2022, 33(4): 1287–1314. http://www.jos.org.cn/1000-9825/6476.htm [doi: 10.13328/j.cnki.jos.006476]

张文跃(1989－), 男, 博士, 讲师, 主要研究领域为自然语言处理, 情感分析, 人工智能.

王素格(1964－), 女, 博士, 博士生导师, CCF 专业会员, 主要研究领域为自然语言处理, 文本挖掘, 情感分析.

李旸(1988－), 女, 博士, 副教授, 主要研究领域为自然语言处理, 智能金融, 金融文本情感分析.

廖健(1990－), 男, 博士, 副教授, CCF 专业会员, 主要研究领域为自然语言处理, 情感分析, 知识图谱.