用模糊规则解读特定领域中的对比嵌入

Javier Fumanal-Idocin、Mohammadreza Jamalifard、Javier Andreu-Perez埃塞克斯大学计算机科学与电子工程学院科尔切斯特，英国{j.fumanal-idocin、jm23525、j.andreu-perez}@essex.ac.uk

arXiv:2603.12227v1 [cs.SC] 12 Mar 2026

摘要—**自由格式文本仍然是现实环境中记录数据的常见方式之一，例如法律程序和医疗记录。正因如此，自然语言处理领域一直在投入大量努力，将这些文本转换为结构化格式，以便标准机器学习方法加以利用。将文本嵌入为向量表示的最流行方法之一是对比语言-图像预训练模型（CLIP），该模型同时使用图像和文本进行训练。尽管 CLIP 计算得到的表示在零样本和少样本学习问题中取得了非常成功的效果，但将其应用于特定领域时仍然存在问题。在本研究中，我们将基于模糊规则的分类系统与一些标准文本处理技术结合起来，把部分感兴趣的特征映射到 CLIP模型创建的空间中。随后，我们讨论所获得的规则和关联，以及每个所考虑特征的重要性。我们在两个不同的数据领域——临床报告和影评——中应用这一方法，并比较分别获得的结果以及同时考虑两者时的结果。最后，我们讨论这一方法的局限性，以及如何进一步改进。**

索引词—**模糊规则、二型模糊集、对比嵌入、中风康复、可解释人工智能**

I. 引言

如今，非结构化数据表示是数字信息最重要的来源之一 [1]。无论是视频 [2], 、文本 [3] 还是音频 [4], ，现有的大多数深度学习研究都旨在利用这些信息，完成各种类型的任务，并适应不同程度的监督 [5]。

大多数最先进的深度学习模型都使用注意力模块，该模块已证明能够很好地扩展到大量数据。其中一种模型是对比语言-图像预训练模型（CLIP） [6],，它将文本和图像编码到同一向量空间中。CLIP嵌入已用于零样本分类和图像检索 [7], ，并成功编码了来自不同领域的文本和图像 [8]。

然而，CLIP嵌入在特定领域应用中的表现并不总是令人满意，而且可能存在众所周知的偏差 [9], [10]。由于目标类别以自然语言指定，CLIP的性能也会受到类别指定方式的影响 [6]。为缓解这些问题，有必要进一步处理CLIP特征或对模型进行微调。对CLIP模型进行微调在数据和计算成本方面都可能代价高昂，而且文献中已提出多种缓解这一问题的方法 [11]。此外，对原始模型进行任何重新训练都可能导致模型性能显著下降 [11]。因此，人们针对特定应用开发了相应策略，例如艺术品分类[12]和临床记录分析 [13]。

对 CLIP特征最常见的处理方法之一是使用 T-SNE 或 PCA 进行可视化 [14]。人们还使用了许多其他方法来理解这些模型生成的嵌入，这对于可解释人工智能（XAI）以及在更大的机器学习流水线中利用这些知识都很有帮助 [15]。模糊规则可以作为可解释人工智能的实用工具 [16], ，但此前尚未与 CLIP嵌入结合用于这一目的。

本文提出使用模糊规则实现可解释人工智能（XAI），并利用这些规则理解 CLIP模型如何构建通用特征嵌入。随后，我们对原始数据进行特征提取，以寻找在语义上相关的变量。最后，将这些变量中的值映射到嵌入空间。通过这种方式，我们可以研究问题中的相关变量如何影响嵌入空间中的数据结构。我们使用了来自两个不同数据领域的数据：中风患者关于康复训练的报告，以及 IMDB 影评。我们比较了这两个数据领域中获得的嵌入如何体现所需的特性。

本文其余部分安排如下：第二节讨论文本处理、CLIP模型和基于模糊规则的分类的相关概念与文献，这些内容对于理解本文其余部分至关重要。第三节介绍我们使用的数据以及所应用的模糊推理系统。第四节讨论所进行的实验及其结果。最后，第五节总结本文并阐述未来的研究方向。

II. 背景

A. 文本处理与情感分析

文本处理是指对原始文本数据进行操作、分析以及信息提取。该过程中的两个基本步骤是分词和词形还原。分词是初始步骤，包括将一段文本拆分为更小的单位，称为词元，这些词元可以是

单词、单词的部分或标点符号。另一方面，词形还原是一种将单词还原到其词根意义（即词元）的技术。自然语言处理中的大多数任务都是在执行这两个步骤之后进行计算的 [17]。

情感分析是一种非常流行的文本处理方式。它关注文本的极性，也可以研究特定的情绪（愤怒、快乐、悲伤等）和意图（感兴趣或不感兴趣）。情感分析技术通常使用为每个词预先设定取值的词典。这些取值可以直接使用，也可以通过其他算法进一步利用 [18], [19]。此外，还可以为分类定义自定义特征 [18]。深度学习模型已在许多自然语言处理任务中达到最先进水平 [20]。不过，传统技术仍应用于许多现实场景，例如客户反馈分析、社交媒体监测和品牌声誉管理 [21]。

*B. CLIP*特征

CLIP特征嵌入最初是为通用零样本图像分类问题提出的 [6]。CLIP模型能够将不同目标类别的文本表示与待分类图像编码到同一向量空间中。预测类别是图像与该向量空间之间余弦距离最小的类别。

除这一最初的应用外，后续研究还测试了对比嵌入在其他相关领域中的有效性，例如图像描述 [22], 图像检索 [7] 和视频转文本 [23]。除了使用按原始方式训练的 CLIP模型 [6], ，也可以重新训练或微调它，使其适用于特定领域。这是一个热门的研究方向，因为原始 CLIP模型在特定或小众任务中的表现不佳 [10], ，而且从头开始重新训练基础模型成本高昂，需要大量数据，而某些特定应用可能根本没有这类数据 [8]。

另一种可能性是不经微调，直接利用 CLIP嵌入空间，并将这些特征用作正则化元素 [24], [25] ，或作为其他机器学习流水线的输入 [26]。这种方法受原始 CLIP特征表达能力的限制，但如果额外的机器学习方法计算速度很快，它可能比微调 CLIP模型更高效。现有研究表明，即使是线性模型，也足以实现领域自适应 [27]。

C. 基于模糊规则的分类

基于模糊规则的分类是指使用遵循以下结构 [28]的规则，将观测样本判别为不同类别：

$$\text{IF } \mathbf{x}_1 \text{ is } \mathbf{a}_{j1} \ldots \mathbf{x}_n \text{ is } \mathbf{a}_{jn} \text{ THEN class } j \text{ for } j = 1, \ldots, C \tag{1}$$

.

通常，模糊规则分类器（FRBC）会计算输入与所有规则的匹配度，并选择匹配度最高的规则。需要注意的是，可以采用不同的标准来衡量匹配度。也可以考虑多条规则，而不仅仅是匹配度最高的那条规则 [29]。不过，这会降低系统的可解释性，因为最终决策现在涉及多条规则。

D. 相关工作

模糊逻辑已成为许多不同应用中可解释数据分类的重要工具，为可解释人工智能领域的重要贡献奠定了基础。具体而言，一项研究 [30] 提出了一种专为情感分析设计的基于模糊的系统，并特别关注产品评论。他们专门研究了模糊限制语和语言修饰语在产品评论中的作用。在此背景下，另一项值得注意的贡献是 [31],，其中将模糊推理系统与其他机器学习方法进行比较，模糊系统在可解释性和准确率方面均取得了有利结果。模糊规则也被用于多模态数据， [32]。在该研究中，作者在更大的机器学习流水线中使用模糊规则，从而使模糊分类器利用从文本和音频中获得的特征做出最终决策。

模糊逻辑的一些优势也已应用于医疗领域，因为专家知识可以表示为规则，而且基于规则的系统也更容易让机器学习系统的最终用户理解。在 [33] 中，作者修订了从医学知识构建知识库的一些方法。

III. 方法

A. 使用的数据

我们研究了来自不同应用领域的两个数据集：一组成年人在中风后进行活动能力锻炼的报告 [34], ，以及一套知名的 IMDB 影评集合。下文中，我们将前者称为“临床”数据集，将后者称为“Film”数据集 [35]。

- 临床数据集由一系列对曾中风的老年患者进行的访谈转录文本组成。在访谈中，受访者描述了他们对康复训练及其中涉及的技术（在线课程、可穿戴设备、应用程序等）的看法。该数据集包含来自 33 位不同患者的 51 条评论。

- Film 数据集由 IMDB 网站上 50,000 条热门电影评论组成。

B. 模糊推理系统

为获得可视为具有可解释性的模糊规则分类器，我们将规则的最大数量设为15，前件的最大数量设为3。不过，实际使用的规则数量会根据质量指标从该上限中进一步筛减。我们还为模糊划分选择了3个易于理解的语言标签：低、中和高，这些标签最直观易懂。

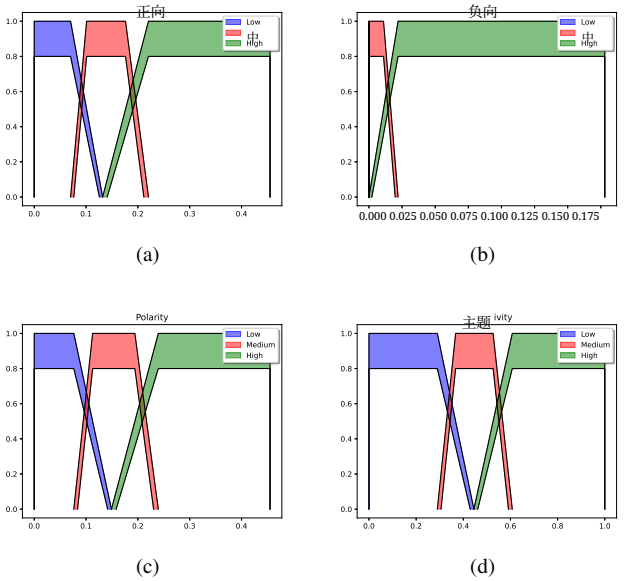


图 1：使用区间二型模糊集对临床数据集中的各变量进行模糊划分。

这些划分是根据可用数据进行经验计算得出的。低划分从最小值到0.2分位数的范围内具有最大隶属度值，并逐渐下降，直到0.5分位数。中划分从0.2分位数开始，在0.5分位数处达到最大值，并逐渐下降，直到0.8分位数。高划分对0.5、0.8和1分位数范围采用相同方式。使用区间二型模糊集时，下隶属度的最大程度限制为0。8。（见图1）。

对于每条规则的匹配度，我们计算每条规则 r [36]，[37] 的优势得分 ds_r。该指标是以下两项的乘积：

- 规则触发的先验概率， s_r：

$$s_r = \frac{\sum_{\mathbf{x} \in Cons_r} w_r(\mathbf{x})}{|\mathbf{x} \in Cons_r|}. \tag{2}$$

- 规则激活强度相对于其他规则的大小， c_r：

$$c_r = \frac{\sum_{\mathbf{x} \in Cons_r} w_r(\mathbf{x})}{\sum_{\mathbf{x} \in Cons_r} \sum_{r\prime=1}^{R} w_{r\prime}(\mathbf{x})}. \tag{3}$$

其中， $w_r(\mathbf{x})$ 是规则 r 对样本 $\mathbf{x}$ 的激活强度，$\mathbf{x} \in Cons_r$ 是真实类别与规则类后件相对应的观测样本集合， R 是所有规则的集合。激活强度由规则所有前件的真值度相乘得到。

最后，我们使用 $w_r(\mathbf{x})$ 和 ds_r 计算匹配度 $as_r(x)$：

$$as_r(x) = w_r(\mathbf{x}) * ds_r. \tag{4}$$

根据与每个样本关联度最高的规则的后件类别对样本进行分类。

我们还使用支配度分数来剪除未达到最低值的规则。

在实验中，我们使用遗传算法，基于预计算划分优化每条规则的前件和后件。优化指标为马修斯相关系数（MCC）：

$$MCC = \frac{(TP \times TN) - (FP \times FN)}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \tag{5}$$

其中，TP为真正例，TN为真负例，FP为假正例，FN为假负例。

我们还尝试了另一种损失函数，在MCC中加入两个正则项。这些正则项根据每条规则的前件数量（l_1）以及优势得分大于最小阈值（l_2）的规则数量，对规则库的规模进行惩罚：

$$l_1 = \frac{1}{R} \sum_{i=1}^{R} \frac{nAnts(r_i)}{\text{maxAnts}}, \tag{6}$$

$$l_2 = \frac{1}{R} \sum_{i=1}^{R} ds(r_i) > h, \tag{7}$$

其中， R 是规则库中的规则数量， h 是优势得分的最小阈值，maxAnts是规则允许的最大前件数量，而 $nAnts$（r）是规则 r的前件数量。

最终损失是MCC与 l_1 和 l_2的加权和的凸组合：

$$l = 0.95 * MCC + 0.05 * (0.50 * l_1 + 0.50 * l_2). \tag{8}$$

C. 利用模糊规则发掘*CLIP*嵌入

这一过程分为以下不同步骤：

1）从文本中获取 CLIP嵌入：由于 CLIP嵌入最多由 77 个字符构成，当待嵌入文本超过这一长度时，我们会将文本拆分为相应大小的部分，并计算每个部分所得嵌入的平均值。

2）从原始文本中提取相关特征。在本例中，我们关注文本的情感特征：正向性、负向性、中性和极性 [38]。对于患者报告，这些特征代表患者对其康复情况的看法；而对于影评，它们代表影评中对 Film 的看法。

3）我们在嵌入空间中寻找聚类结构。这一点很有用，因为在某些情况下并不需要所有维度；与其试图理解嵌入空间中的全部 512 个特征，不如理解这些群组更为合理。我们使用 k均值算法构建这些簇。

4) 我们使用第 III-B 节提出的模糊规则分类器，将特征映射到 CLIP 空间中的簇。

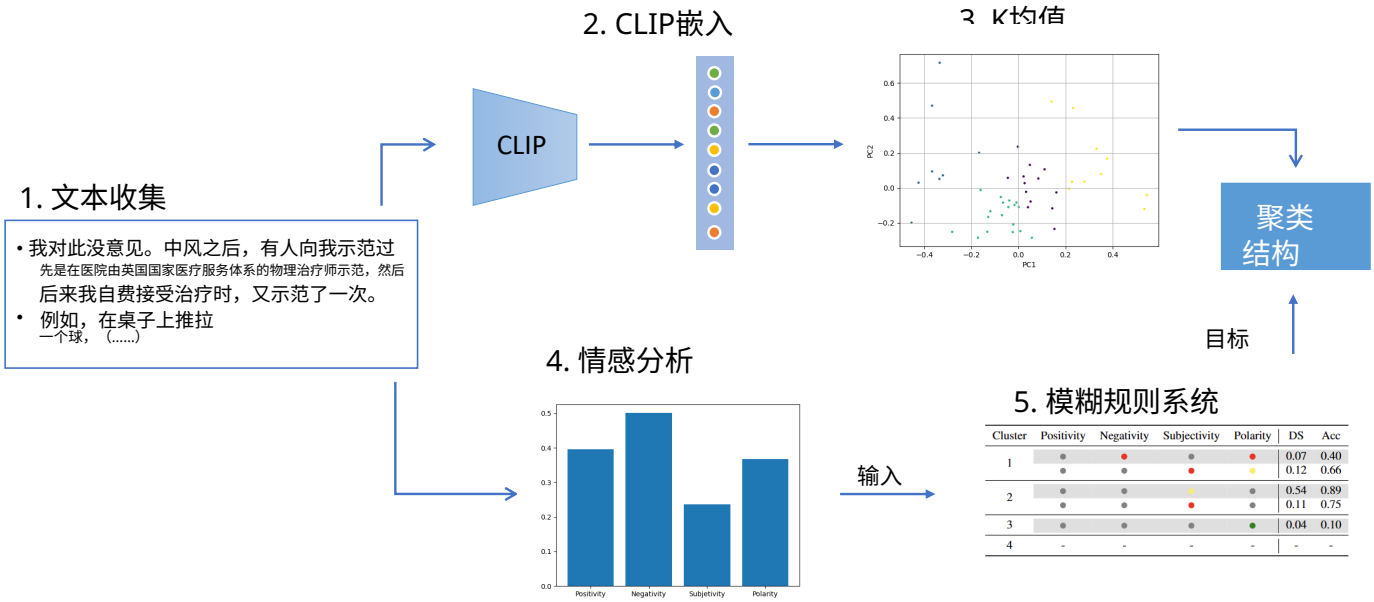


图2：**利用模糊规则分类器（FRBC）结合 CLIP特征和情感分析的拟议方法**。1. 待研究的文本集合。2. 计算所有文本的 CLIP嵌入。3. 使用 K均值聚类算法，获取在该空间中形成的结构。4. 对每段文本，提取情感分析特征。5. 模糊规则分类器（FRBC）：它以每段文本的情感分析特征作为输入，并采用图1所示的模糊划分。随后，它将第3步检测到的簇作为目标。这样，FRBC 就能将我们感兴趣的特征映射到这些特征在嵌入空间中投影到的空间区域。

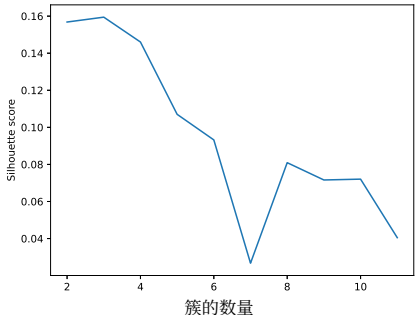


图3：基于所有患者报告获得的 CLIP特征进行 K均值聚类的轮廓系数。

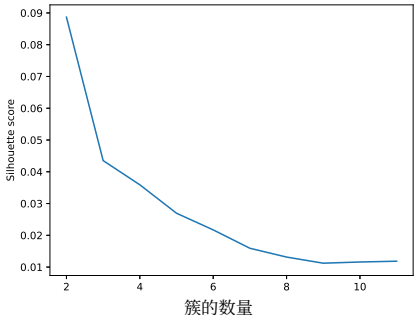


图4：对电影数据集获得的 CLIP特征进行 K均值聚类的轮廓系数。

我们还报告了模糊规则分类器的准确率和马修斯相关系数，以及每条单独规则的支配度分数和准确率。图2直观展示了这一过程。

IV. 实验

在实验中，我们使用 CLIP模型将数据集中的所有文本嵌入到同一特征空间中。随后，使用 K均值在该表示中寻找簇。为了确定最佳簇数，我们针对不同的簇数进行了轮廓系数分析。我们发现，对于这些数据集，3 或 4 个簇都可能是合适的，因为随着簇数增加，轮廓系数会急剧下降（图 3 和图 4）。我们还考虑了其他聚类算法，但结果非常相似。

得到聚类后，我们按照第 III-B 节所述，使用遗传算法拟合模糊规则分类器。我们使用每代 30 名受试者，对其进行 300 代优化。为了确保结果具有统计显著性，我们将实验重复了 30 次。

图5a展示了一个用于临床报告的模糊规则分类器（FRBC），其损失函数仅使用马修斯相关系数（MCC）。该规则库的准确率为 0.63，MCC 为 0.47。在这种情况下，我们共得到 5 条规则，其中“主观性”属性出现在最多的规则中，前件的平均数量为 2.5。使用标准模糊集时，表现最佳的模糊规则分类器的 MCC 为 0.56，准确率为 0.81（该模糊规则分类器见图5b）。最后，我们展示使用所示损失函数时的结果

聚类	正向性	负向性	主观性	极性	DS	Acc
1	●	●	●	●	0.07	0.40
	●	●	●	●	0.12	0.66
2	●	●	●	●	0.54	0.89
	●	●	●	●	0.11	0.75
3	●	●	●	●	0.04	0.10
4	-	-	-	-	-	-

(a) 使用区间值模糊集的MCC损失

聚类	正向性	负向性	主观性	极性	DS	Acc
1	●	●	●	●	0.07	0.75
2	●	●	●	●	0.17	0.85
	●	●	●	●	0.42	0.89
	●	●	●	●	0.06	1.00
3	●	●	●	●	0.06	0.50
	●	●	●	●	0.02	1.00
4	-	-	-	-	-	-

(b) 使用标准模糊集的MCC损失

聚类	正向性	负向性	主观性	极性	DS	Acc
1	●	●	●	●	0.06	1.00
2	●	●	●	●	0.16	0.75
	●	●	●	●	0.06	1.00
	●	●	●	●	0.33	0.94
3	●	●	●	●	0.08	0.16
4	●	●	●	●	0.04	0.10

（c）使用区间值模糊集时，公式（8）中的损失

图5：针对患者报告获得的规则结果，这些规则将情感指标映射到 CLIP 空间中的簇。→ 低，→ 中，→高，→无关。DS 表示支配度分数，Acc. 表示每条规则在其触发的样本中获得的准确率。

后件	正向	负向	极性	主观性	DS	Acc
1	●	●	●	●	0.02	0.40
	●	●	●	●	0.06	0.70
	●	●	●	●	0.10	0.83
	●	●	●	●	0.29	0.72
	●	●	●	●	0.03	1.00
2	●	●	●	●	0.07	0.39
	●	●	●	●	0.03	0.24
3	●	●	●	●	0.23	0.40

（a）使用区间值模糊集的MCC损失

后件	正向	负向	极性	主观性	DS	Acc
1	●	●	●	●	0.02	0.50
	●	●	●	●	0.08	0.37
2	●	●	●	●	0.06	0.44
3	●	●	●	●	0.15	0.56
	●	●	●	●	0.14	0.71
	●	●	●	●	0.09	0.55

(b) 使用标准模糊集的MCC损失

后件	正向	负向	极性	主观性	DS	Acc
1	●	●	●	●	0.01	0.16
2	●	●	●	●	0.01	0.14
	●	●	●	●	0.15	0.32
	●	●	●	●	0.03	0.19
3	●	●	●	●	0.47	0.83

（c）使用区间值模糊集计算式（8）中的损失

图 6：影评中将情感指标映射到 CLIP 空间中各簇的规则所得结果。→低，→中，→高，→无关。DS 表示优势得分，Acc. 表示每条规则在其触发的样本中的准确率。

数据集	Loss	模糊集	准确率	规则	前件
临床	公式 8	T1	0.64 ± 0.09	4.60 ± 1.01	11.40 ± 2.57
	公式 8	T2	0.64 ± 0.07	4.40 ± 1.01	10.00 ± 2.68
	MCC	T1	0.66 ± 0.08	5.50 ± 0.92	13.00 ± 2.14
	MCC	T2	0.70 ± 0.06	5.80 ± 1.72	14.10 ± 3.75
Film	公式 8	T1	0.46 ± 0.96	4.83 ± 1.06	15.54 ± 0.10
	公式 8	T2	0.47 ± 0.01	4.10 ± 1.00	17.60 ± 3.44
	MCC	T1	0.44 ± 0.01	7.00 ± 1.00	11.58 ± 1.02
	MCC	T2	0.46 ± 0.22	6.91 ± 1.05	14.00 ± 2.53

表 I：不同模糊规则分类器在所测试的不同数据配置、损失和模糊集下的性能报告。结果为 30 次试验的平均值。

见公式 (8)。在这种情况下，准确率为 0.69，MCC 为 0.33。该模糊规则分类器学得 6 条规则，前件的平均数量为 1.66。

在使用 Film 数据集重复相同实验时，采用标准模糊集和 MCC 损失，我们得到的准确率为 0.36。使用梯度提升分类器（被认为是表格数据分类领域最先进的方法），得到的结果也相近：准确率为 0.40，MCC 为 0.10。

该数据集中的词汇在各样本间的差异更大，因此更难处理 [39]。因此，我们感兴趣的特征对该空间中所形成结构的影响要小得多。在这种情况下，我们的做法是研究一次局部交互：随机选择一个样本，然后根据欧几里得距离，在 CLIP 空间中选出与其最接近的 1000 个样本（使用余弦距离会得到相近的结果）。这种方法类似于其他可解释人工智能研究，这些研究会考察变量之间的局部交互 [40]。

图 6 展示了此案例的规则库。所得规则库的规模与临床数据集中得到的规则库相近。我们发现，大多数情况下，一个簇能够被清晰识别，但其余簇则需要更复杂的模式。我们还可以观察到，大多数规则基本忽略了主观性特征，并且多数规则都考虑了极性。

如表 I 所示，最佳配置采用标准模糊集和式 8 中的损失函数，所得平均准确率为 0.49。我们还可以观察到，与 MCC 损失相比，规则数量有所减少，但前件数量并非如此。

A. 讨论

我们的结果表明，使用模糊规则和情感分析特征，有助于理解利用 CLIP模型在特定数据集上形成的结构。CLIP模型对文本进行嵌入，兼顾文本的所有特征，并使用固定数量的维度；与向量表示相比，这是一个优势。与 IMDB 数据集相比，该模糊规则分类器在临床数据集上的效果显著更好；在 IMDB数据集中，得到的规则集包含更多规则，效果也较差。我们认为，这是因为临床数据集的样本数量要少得多，而且文本内容的变化范围也较小。两种情况下的结果似乎都表明，增加变量有助于提升规则库性能。为减少额外变量的数量，我们可以利用特征工程，对最有利于模糊规则分类器的变量进行人工构造。将这类变量与用于可解释人工智能的可解释变量结合使用，是一个活跃的研究课题 [40]。

我们尝试了两种不同的损失函数进行优化：一种只关注分类性能，另一种还考虑了规则库大小。第二种损失函数在优化过程中的作用使规则库变得更小，但精确率也有所下降。这种准确率与可解释性之间的权衡是模糊规则分类器中众所周知的问题 [41]。我们还发现，采用 MCC 损失时，区间二型模糊集比标准模糊集取得了更好的结果。性能提升可归因于更灵活的隶属函数，以及这些模糊集划分输入空间的方式更为高效 [42]。然而，使用尺寸感知损失函数时，这种正向效果便会消失。

最后，我们还发现，在某些规则中，优势得分与准确率可能相差很大。这种情况在临床数据集的规则库中尤为明显（见图5）。优势得分用于衡量一条规则的优劣，依据的是其模式的普遍性以及相对于模糊规则分类器中其他规则的独特性。准确率与优势得分之所以不一致，是因为这些规则捕捉到了簇中的局部模式。可以基于这些规则构建全局解释 [40], ，但这需要专门构建一个模糊规则分类器来寻找这些解释。

V. 结论与未来研究方向

本文提出了一种使用模糊规则解读CLIP模型嵌入空间的方法。首先，我们解释了通过情感分析从原始文本中提取的特征。然后，为了在每个数据集的嵌入空间中寻找相关结构，我们使用了K均值。最后，我们将这些特征的表达程度映射到CLIP数据集中的各个簇。我们展示了采用不同模糊集和损失函数对模糊规则分类器规模的影响。我们将该方法应用于一个包含患者访谈记录的数据集，这些记录讨论了他们在运动康复训练中的体验；此外，还应用于IMDB电影评论数据集。最后，我们还比较了分别使用各个数据集时模糊规则分类器的差异。

我们的结果表明，当文本主题之间的差异足够小时，这种映射能够捕捉CLIP嵌入所形成的结构与目标特征之间的关系。然而，这种映射并不完美，要获得更准确的映射还需要更多信息。区间二型模糊集的效果优于标准模糊集。我们还发现，对规则数量和规则规模进行惩罚，会显著降低模糊规则分类器最终的马修斯相关系数。因此，性能与可解释性之间的权衡仍有待未来研究探索。

未来的研究方向应探索其他临床数据的嵌入空间，以获得能够捕捉患者状态并与其他患者进行比较的表征。我们还打算将模糊规则分类器与其他局部和模型无关可解释性方法结合使用。

VI. 致谢

Javier Fumanal-Idocin 的研究获得了欧盟玛丽·斯克沃多夫斯卡-居里 YUFE4 博士后项目的资助。

参考文献

[1] S. Jain、A. de Buitl´eir 和 E. Fallon，“非结构化数据分析与解析方法综述”，载于*2020*年新兴智能计算与信息学国际会议（*ESCI*）。电气电子工程师学会，2020年，第164–169页。[2] D. Liu、Y. Li、J. Lin、H. Li 和 F. Wu，“基于深度学习的视频编码：综述与案例研究”，《ACM计算调查》，第53卷，第1期，第1–35页，2020年。[3] J. Devlin、M.-W. Chang、K. Lee 和 K. Toutanova，“BERT：用于语言理解的深度双向变换器预训练”，*arXiv*，2018年10月。[4] D. Michelsanti、Z.-H. Tan、S.-X. Zhang、Y. Xu、M. Yu、D. Yu 和 J. Jensen，“基于深度学习的视听语音增强与分离概述”，《IEEE/ACM音频、语音与语言处理汇刊》，第29卷，第1368–1396页，2021年。[5] L. Ericsson、H. Gouk 和 T. M. Hospedales，“自监督模型的迁移效果如何？”，载于*IEEE/CVF*计算机视觉与模式识别会议论文集，2021年，第5414–5423页。[6] A. Radford、J. W. Kim、C. Hallacy、A. Ramesh、G. Goh、S. Agarwal、G. Sastry、A. Askell、P. Mishkin、J. Clark 等，“通过自然语言监督学习可迁移视觉模型”，载于国际机器学习会议。机器学习研究论文集，2021年，第8748–8763页。[7] A. Baldrati、M. Bertini、T. Uricchio 和 A. Del Bimbo，“结合基于CLIP的特征进行有效的条件式与组合图像检索”，载于*IEEE/CVF*计算机视觉与模式识别会议论文集，2022年，第21 466–21 474页。[8] A. J. Thirunavukarasu、D. S. J. Ting、K. Elangovan、L. Gutierrez、T. F. Tan 和 D. S. W. Ting，“医学领域的大语言模型”，《自然·医学》，第29卷，第8期，第1930–1940页，2023年。[9] S. Agarwal、G. Krueger、J. Clark、A. Radford、J. W. Kim 和 M. Brundage，“评估CLIP：探索其更广泛的能力及下游影响”，*arXiv*预印本 *arXiv:2108.02818*，2021年。[10] J. Fumanal-Idocin、O. Cord´on、G. P. Dimuro、A.-F. R. L´opez-de Hierro 和 H. Bustince，“社交网络分析中外部信息的量化：比较神话学应用”，《IEEE控制论汇刊》，2023年。[11] Y. Wei、H. Hu、Z. Xie、Z. Liu、Z. Zhang、Y. Cao、J. Bao、D. Chen 和 B. Guo，“提升CLIP微调性能”，载于*2023*年*IEEE/CVF*国际计算机视觉会议（*ICCV*），2023年，第5416–5426页。[12] M. V. Conde 和 K. Turgutlu，“CLIP-ART：面向细粒度艺术品分类的对比式预训练”，载于*IEEE/CVF*计算机视觉与模式识别会议论文集（*CVPR*研讨会），2021年6月，第3956–3960页。

[13]W. Lin、Z. Zhao、X. Zhang、C. Wu、Y. Zhang、Y. Wang 和 W. Xie，“Pmc-clip：利用生物医学文档进行对比式语言-图像预训练”，载于医学图像计算与计算机辅助干预国际会议，2023年。 [在线]。网址：https://api.semanticscholar.org/CorpusID:257496659

[14] J. Pareek 和 J. Jacob，“利用PCA和T-SNE进行数据压缩与可视化”，载于《信息通信技术与计算进展：AICTC 2019论文集》。Springer，2021年，第327–337页。

[15] G. Joshi、R. Walambe 和 K. Kotecha，“多模态深度神经网络可解释性综述”，IEEE Access，第9卷，59 821，2021年 59 800页– 。[16] J. M. Mendel 和 P. P. Bonissone，“关于基于规则的模糊系统中可解释人工智能（XAI）的批判性思考”，IEEE模糊系统汇刊，第29卷，第12期，第3579–3593页，2021年。[17] K. Chowdhary 和 K. Chowdhary，“自然语言处理”，《人工智能基础》，第603–649页，2020年。[18] S. Bird，“NLTK：自然语言工具包”，载于COLING/ACL 2006交互式演示环节论文集，2006年，第69–72页。[19] C. Hutto 和 E. Gilbert，“VADER：一种用于社交媒体文本情感分析的简约基于规则的模型”，载于AAAI国际网络与社交媒体会议论文集，第8卷，第1期，2014年，第216–225页。[20] I. Lauriola、A. Lavelli 和 F. Aiolli，“自然语言处理中的深度学习导论：模型、技术与工具”，《神经计算》，第470卷，第443–456页，2022年。[21] K. P. Kalyanathaya、D. Akila 和 P. Rajesh，“自然语言处理进展——当前研究趋势、开发工具和行业应用综述”，《近期技术与工程国际期刊》，第7卷，第5C期，第199–202页，2019年。[22] M. Barraco、M. Cornia、S. Cascianelli、L. Baraldi 和 R. Cucchiara，“CLIP特征在图像描述中的惊人效果：实验分析”，载于IEEE/CVF计算机视觉与模式识别会议论文集，2022年，第4662–4670页。[23] H. Xu、G. Ghosh、P.-Y. Huang、D. Okhonko、A. Aghajanyan、F. Metze、L. Zettlemoyer 和 C. Feichtenhofer，“VideoCLIP：用于零样本视频-文本理解的对比式预训练”，arXiv预印本 arXiv:2109.14084，2021年。[24] N. Garcia、B. Renoust 和 Y. Nakashima，“用于自动艺术分析的上下文感知嵌入”，载于2019年多媒体检索国际会议论文集，2019年，第25–33页。[25] J. Fumanal-Idocin、O. Cord、H. Bustince 等，“ArtXAI：可解释人工智能利用模糊技术为艺术图像筛选深度表征学习”，IEEE模糊系统汇刊，2023年。[26] P. Gao、S. Geng、R. Zhang、T. Ma、R. Fang、Y. Zhang、H. Li 和 Y. J. Qiao，“CLIP-Adapter：通过特征适配器改进视觉-语言模型”，arXiv，第abs/2110.04544卷，2021年。 [在线]。网址：https://api.semanticscholar.org/CorpusID:238583492

[27] Y. Ouali、A. Bulat、B. Matinez 和 G. Tzimiropoulos，“视觉-语言模型的黑盒小样本适配”，载于IEEE/CVF国际计算机视觉会议（ICCV）论文集，2023年10月，第15 534–15 546页。[28] B. Kosko，“模糊认知图”，《国际人机研究期刊》，第24卷，第1期，第65–75页，1986年。[29] J. Wieczynski、G. Lucca、E. Borges、L. Emmendorfer、M. Ferrero-Jaurrieta、G. Dimuro 和 H. Bustince，“在分类问题中应用d-XChoquet积分”，载于2022年IEEE国际模糊系统会议（FUZZ-IEEE）。电气电子工程师学会，2022年，第1–7页。[30] K. Indhuja 和 R. P. Reghu，“基于模糊逻辑的产品评论文档情感分析”，载于2014年第一届计算系统与通信国际会议（ICCSC）。电气电子工程师学会，2014年，第18–22页。[31] H. Liu 和 M. Cocea，“用于可解释情感分析的基于模糊规则的系统”，载于2017年第九届先进计算智能国际会议（ICACI）。电气电子工程师学会，2017年，第129–136页。[32] S. Vashishtha 和 S. Susan，“使用模糊规则，根据文本与语音线索的监督分类推断情感”，《计算机科学进展》，第167卷，第1370–1379页，2020年。[33] N. H. Phuong 和 V. Kreinovich，“模糊逻辑及其在医学中的应用”，《国际医学信息学期刊》，第62卷，第2–3期，第165–173页，2001年。[34] A. K. Shafizadeh Mohsen 和 B. Tom，“中风老年人家庭锻炼环境的主题分析”，2022年。 [在线]。网址：https://reshare.ukdataservice.ac.uk/856695/

[35] A. L. Maas、R. E. Daly、P. T. Pham、D. Huang、A. Y. Ng 和 C. Potts，“用于情感分析的词向量学习”，载于第*49*届计算语言学协会年会：人类语言技术论文集。美国俄勒冈州波特兰：计算语言学协会，2011年6月，第142–150页。 [在线]。可访问：http://www.aclweb.org/anthology/P11 1015 - M. Kiani、J. Andreu 和 H. Hagras，“用于时间依赖型可解释人工智能的时态二型模糊系统”，《*IEEE*人工智能汇刊》，2022年，即将出版。[37] J. Andreu-Perez、L. L. Emberson、M. Kiani、M. L. Filippetti、H. Hagras 和 S. Rigato，“基于可解释人工智能的分析：解读发展认知神经科学中的婴儿功能性近红外光谱数据”，《*Communications Biology*》，第4卷，第1期，第1077页，2021年。[38] S. Bird、E. Klein 和 E. Loper，《使用*Python*进行自然语言处理：利用自然语言工具包分析文本》。O’Reilly Media, Inc.，2009年。[39] A. Natekin 和 A. Knoll，“梯度提升机：教程”，《*Frontiers in Neurorobotics*》，第7卷，第21页，2013年。[40] S. M. Lundberg、G. Erion、H. Chen、A. DeGrave、J. M. Prutkin、B. Nair、R. Katz、J. Himmelfarb、N. Bansal 和 S.-I. Lee，“借助面向树模型的可解释人工智能，从局部解释到全局理解”，《*Nature Machine Intelligence*》，第2卷，第1期，第56–67页，2020年。[41] A. Moral、C. Castiello、L. Magdalena 和 C. Mencar，可解释模糊系统。Springer，2021年。[42] J. M. Mendel，“不确定规则型模糊系统”，导论与新方向，第684卷，2017年。