

梁家振

1. COSMIC: COmmonSense knowledge for eMotion Identification in Conversations (ACL)

● Motivation

本文旨在利用常识性知识解决话语层面情感识别中在情境传播、情感转移检测和相关情感类之间的区分等方面经常遇到困难。

● Method

本文提出了一个新的框架 COSMIC，它包含了不同的常识元素，如精神状态、事件和因果关系，并以此为基础学习参与对话的对话者之间的互动。

● Experimental result

通过学习不同的常识表示，COSMIC 解决了这些挑战，并在四个不同的基准会话数据集 (IE 上实现了最新的情感识别结果。

2. DialogXL: All-in-One XLNet for Multi-Party Conversation Emotion Recognition(AAAI)

● Motivation

在往常的 ERC 工作中，来自不同方面的对话话语交替出现，往往被组织成层次结构，这样的结构不利于预训练模型 (pre-train model) 的应用。

为了解决这个问题，本文提出了一个一体化的 XLNet 模型，即 DialogXL，它具有增强的记忆能力来存储更长的历史上下文，同时它还拥有对话感知的自我注意 (dialog-aware self-attention) 来处理多方结构。

● Method

本文首先将 XLNet 的递归机制从段级修改为语句级，以便更好地对会话数据建模。其次，本文引入了对话感知的自我注意，以取代 XLNet 中的普通自我注意，以捕获有用的说话者内部和说话者之间的依赖关系。

● Experimental result

本文通过对四个 ERC 基准进行了大量的实验，并提出了主流模型进行比较。实验结果表明，该模型在所有数据集上的性能都优于 baseline。还进行了消融实验和误差分析等其他实验，结果证实了 DialogXL 关键模块的作用。

3. MMGCN: Multimodal Fusion via Deep Graph Convolution Network for Emotion Recognition in Conversation(ACL)

● Motivation

为了探索更有效地利用多模态和远距离上下文信息的方法，本文提出了一种基于多模态融合图卷积网络的新模型 MMGCN。

● Method

为了获取跨多种模式的话语级上下文依赖性，我们提出了一个多模态融合图卷积网络 (MMGCN)。本文首先构建了一个谱域图卷积网络来编码多模态上下文信息，然后堆叠更多的层来构造一个深层的 GCN。另外，为了更全面的提升模型性能，本文还加入了说话人级别的上下文信息来嵌入。

● Experimental result

本文在两个公共基准数据集 (IEMOCAP 和 MELD) 上对我们提出的模型进行了评估，结果证明了 MMGCN 的有效性，在多模态会话设置下，它的性能显著优于其他 SOTA 方法。

李巧媚

1. Attention Biasing and Context Augmentation for Zero-Shot Control of Encoder-Decoder Transformers for Natural Language Generation(AAAI, CCF-A)

- Motivation:

基在基于神经网络自然语言生成（NLG）模型以实现理想输出，是机器翻译、对话系统等功能实现的重要控制方法，因此本文提出用 zero-shot 这种方式实现，该方式可省去额外注释数据和额外训练的需求。

- Method:

这篇文章是在基于 transformer 的自然语言生成框架上加上三个小部件，使之能够实现 zero-shot 的文本生成，上图是在推理时候使用控制旋钮在 encoder-decoder 的 transformer 框架上实现 zero-shot 受控文本生成，控制部件影响生成的过程，使生成的输出具有所需要的属性。

- Experimental Result:

利用两个控制旋钮去控制 EDT-NLG 模型，发现旋钮对 transformer 层是稳健的，CTX.AUG 与 ATTN.BIAS 共同使用时，p 值达到 0.619，说明旋钮对模型的控制能力极强。
(Written by Qiaomei Li)

2. Improving Multimodal Fusion with Hierarchical Mutual Information Maximization for Multimodal Sentiment Analysis(EMNLP, CCF-B)

- Motivation:

现有的特征融合方法如反向传播、操纵特征空间的几何属性等方法缺乏控制从原始输入到融合嵌入的信息流，可能会丢失数据的实际信息，并在其他模态的引入过程中产生噪声，故提出了一种用于多模态情感分析的分级互信息最大化框架（MMIM），互信息最大化发生在输入级别和融合级别，以减少有价值的任务相关信息的丢失。

- Method:

由特征融合和最大化信息两个部分组成，在融合部分，堆叠的线性激活层的融合网络 F 将单模态表示转换为融合结果 Z，然后通过回归多层感知器（MLP）进行最终预测。在最大化信息 MI 部分，估计并提升了两个级别的 MI 下限——输入级别和融合级别。两部分同时工作以产生反向传播的任务和 MI 相关的 losses，通过该模型学习将任务相关的信息注入融合结果，并提高主要任务中预测的准确性。

- Experimental Result:

在相同的参数设置下运行模型五次，发现 MMIM 产生了与许多基线方法更好或可比的结果，MMIM 实现了与 SOTA 非常接近的性能 (<0.5%)。这些结果初步证明了该方法在 MSA 任务中的有效性。(Written by Qiaomei Li)

Novelty Controlled Paraphrase Generation with Retrieval Augmented Conditional Prompt Tuning (AAAI, CCF-A)

- Motivation:

释义生成就是对指定的文本进行改写，要保持文本原来的含义，要将预训练语言模型应用于一些 NLP 的任务，最标准方法就是微调模型中的参数。但是随着这些预训练模型日益增大，参数逐渐增多，现在都是以亿为数量级的，微调参数并不是一个好的办法，因此该篇文章提出了新的方法在保留原文语义内容的同事，也发挥了词汇的新颖性。

- Method:

该篇文章提出了 RAPT，一种基于 KNN 的检索方式，可进一步提高释义生成的质量；

此外，该篇文章还提出了 NC-RAPT，对释义输出的新颖性进行可控输出。

- **Experimental Result:**

该篇文章在四个数据集上进行分析验证，PART 方式的增加在所有指标上都超过了 GPT2，且 NC-RAPT 的结果也符合预期，释义的输出具有很高的新奇度。(Written by Qiaomei Li)

黄骏

2020:

1. Building a Bridge: A Method for Image-Text Sarcasm Detection Without Pretraining on Image-Text Data (ACL, CCF-A)

- **Motivation:**

在以往的图像-文本讽刺检测的工作中，主要是对模态信息进行多步的预处理然后融合，这样会损失一些模态特征的细节。

- **Method:**

本文提出使用经过预训练的BERT和ResNet模型进行模态特征提取，能够减少数据预处理的时间；另外使用Bridge Layer将模态的特征空间相连，使用多层的多头自注意机制将模态信息进行融合，最后使用2D-Intra-Attention来探索文本与图像之间的对比度和差异。

- **Experimental Result**

在Twitter数据上，取得了更好的表现，性能优于SOTA。(Written by Jun Huang)

2021:

2.MMGCN: Multimodal Fusion via Deep Graph Convolution Network for Emotion Recognition in Conversation (ACL, CCF-A)

- **Motivation:**

现有的多模态ERC任务中，仅仅通过简单的特征串联来融合多模态信息，并没有充分利用模态间和模态内的信息，而且往往会忽略了提取Speaker的信息，另外远距离上下文信息通常也会包含着许多信息。

- **Method:**

本文中，作者提出了多模态融合图卷积（MMGCN），该方法能够有效地挖掘并融合多模态信息，同时捕获远距离上下文信息。MMGCN在每个模态中构建完全连接的图，并在不同模态中对应于相同话语的节点之间建立边缘，这样能够充分融合多模态信息。并且将Speaker的信息嵌入到GCN中。使用多层GCN可以捕获远距离上下文依赖，以进一步提高识别性能。

- **Experimental Result**

作者在IEMOCAP和MELD两个多模态对话数据集上进行实验，比对其他模型取得了较为不错的结果。这说明MMGCN在融合多模态信息以及捕获上下文依赖方面均有着优势。

(Written by Jun Huang)

3. DialogueTRM: Exploring the Intra- and Inter-Modal Emotional Behaviors in the Conversation (ACL, CCF-A)

- **Motivation:**

最近的ERC研究从自我依赖和人际依赖两个方面来探讨情绪的动态变化，但是忽略多模态会话中的时间依赖和空间依赖。而且一般模型不能满足模态的上下文偏好和多粒度融合问题。

- **Method:**

本文提出的DialogueTRM模型，可以从模态内和模态间的角度探讨不同的情感行为。使

用Transformer和BERT构建的Hierarchical Transformer (HT) 层可以通过改变注意Mask来在顺序结构和前馈结构之间切换,从而管理模态之间的时间依赖性和上下文偏好;另外设计了一个多粒度交互式融合层 (MGIF) 来处理特征空间的神经元粒度融合和向量粒度融合,这样可以实现空间依赖。

• Experimental Result

作者分别针对分类情绪和连续情绪两个不同的任务进行了实验,其中分类情绪任务使用了 IEMOCAP 和 MELD 两个数据集,连续情绪任务使用了 VEC 数据集。模型实现了 SOTA 性能,并且 MGIF 模块在 ERC 设置中优于 SOTA 融合方法。 ([Written by Jun Huang](#))

钟清山

1. A Computational Approach to Understanding Empathy Expressed in Text-Based Mental Health Support (EMNLP 2020)

Motivation:

数以百万的人使用基于文本的平台进行心理健康支持,在这些情况下理解同理心是至关重要的。

methd:

设计了一个基于 robert 的多任务双编码器模型,用于识别对话中的共情,并提取其预测背后的基本原理

result:

有效的识别了同理心潜在的基本原理,对心理健康平台很重要,有助于通过基于模型的反馈改善对同辈人的心理支持。(written by qingshan zhong) [BibTex](#)

2. Improving Empathetic Response Generation by Recognizing Emotion Cause in Conversations (EMNLP 2021)

motivation

目前的移情反应生成方法主要是学习一个模型来预测一个情绪标签,并基于该标签产生一个反应,并取得了很好的效果。然而,情感的原因,作为移情反应的一个重要因素,却被忽略了。

methd:

设计了情绪推理器和回复生成器,情绪反应器的任务是对上下文级别情绪预测和词级情绪原因的检测 (定位产生情绪原因的词),回复生成器通过情绪推理器提供的情感信息生成共情回答 (分别设计了软硬门控机制对生成共情回答的生成进行了研究)。

result:

模型可以产生出更有意义和更有同理心的回复。(written by qingshan zhong) [BibTex](#)

3. Towards Facilitating Empathic Conversations in Online Mental Health Support: A Reinforcement Learning Approach (WWW 21)

motivation:

在线心理健康平台上,高度移情的对话很少见。

Methd:

使用一个基于 GPT-2 的转换语言网络,生成候选的共情句子且将其插入到合适的位置。

Result:

模型可以产生更多的共情、具体和多样化的反应,并且在相关任务如风格潜移和共情对

话生成中优于 NLP 方法。(written by qingshan zhong) [BibTex](#)

李外

1. Context — Non-Acted Text and Keystrokes Database and Learning Methods to Recognize Emotions (AAAI,CCF-A)

● Motivation:

本文提出了一个多模态文本-按键数据集和相关的学习方法,用于识别隐藏在小文本中的人类情感

● Method:

在多个场景下收集了 69 个参与者的键击的文本数据,在一个受控制的环境中的视频诱导刺激。在刺激诱导之后,参与者以一种无指导的方式写下他们对给定场景的评论。之后,键击和文本中的特征从数据集种被提取。这些方法以及一些学习方法一起用来识别隐藏在短文种的情感。

● Experimental Result:

结果表明,SVM 和判别分析的性能较好,基于树的特征选择的特征提取效果都较好。对 20 种不同的训练和验证数据组合进行了训练和验证,同时对未看到的数据进行了测试,实验结果表明,混合模型和基于文本的模型获得了较高的精度。

2. Context — Contrast and Generation Make BART a Good Dialogue Emotion Recognizer(AAAI,CCF-A)

● Motivation:

区分不同的情绪类别也不是很简单的,因为它们通常具有语义上相似的情绪,为此,我们采用监督对比学习,使不同的情绪互斥,从而更好地识别相似的情绪。

● Method:

为了实现实验目标,我们使用预先训练好的编码器-解码器模型 BART 作为我们的主干模型,将反应生成作为 BART 的辅助任务,即 CoGBART,用于对话中的情绪识别(ERC)。

● Experimental Result:

在四个数据集上的实验表明,我们提出的模型在对话情绪识别方面比现有的模型获得了更有利的结果,消融研究进一步证明了监督对比损失和生殖损失的有效性。

3. Context — DRER: Deep Learning-Based Driver's Real Emotion Recognizer (AAAI,CCF-A)

● Motivation:

在智能车辆中,监控驾驶员的身体状况是非常重要的,以往的研究大多集中在面部表情识别来监测驾驶员的情绪状态,然而,在驾驶时,许多因素阻止了司机暴露出他们脸上的情绪,为了解决这一问题,我们提出了一种基于深度学习的驾驶员真实情绪识别器(DRER),这是一种基于深度学习的算法来识别驾驶员的真实情绪最近根据他们的面部表情进行了识别。

● Method:

在 CNN 的基础上,提出了 FER 模型,通过拍摄驾驶员的面部图像来识别驾驶员的面部表情状态。然后,在 DNN 的基础上,提出了 SFER 模型,通过融合识别的面部表情状态来识别驾驶员的真实情绪状态。

● Experimental Result:

该方法在驾驶情况下对驾驶员诱发情绪的识别准确率达到 86.8%。