

李巧媚

2021:

1. Missing Modality Imagination Network for Emotion Recognition with Uncertain Missing Modalities(ACL,CCF-A)

- Motivation:

在先前的工作中，多模态融合已被证明可以提高情感识别性能。然而，在实际应用中，我们经常遇到缺少模态的问题，并且哪些模态将被丢失是不确定的。在这种情况下，它使固定的多模态融合失败。在本文中提出了一个统一的模型，即缺失模态想象网络（MMIN），以处理不确定的缺失模态问题。

- Method:

MMIN 模型包含三个主要模型，用于提取特定模态嵌入的模态编码器网络；基于级联残差自编码器和循环一直性学习的想象模块，在给定相应的可用模态的表示的情况下，想象缺失模态的表示，之后收集自编码器中的潜在向量，形成联合多模态表示。最后通过基于联合多模态表示的情绪分类器用于预测情绪类别。

- Experimental Result:

在两个基准数据集 IEMOCAP 和 MSPIMPROV 上的实验表明，在不确定的缺失模态测试条件和全模态测试条件下，提出的 MMIN 模型都能显著提高多模态情感识别的性能。

[\(Written by Qiaomei Li\)](#)

2021:

2. Fusing Label Embedding into BERT: An Efficient Improvement for Text Classification(ACL,CCF-A)

- Motivation:

随着 BERT 等预先训练模型获得越来越多的关注，从数据增强实验到改进模型数学原理，大量的研究已经进一步提高了它们的性能。在本文中，作者提出了一种简洁有效的方法，在保持几乎相同的计算成本的情况下，利用标签嵌入技术来提高 BERT 的文本分类性能。

- Method:

首先，使用标签嵌入技术，嵌入了标签的内容之后，可以在训练模型的同时处理标签信息和输入特征。模型将标签和文本以 $w/[SEP]$ 标记作为输入，并对标签和文本使用不同的段嵌入，然后再通过注意力机制得到融合了标签信息和文本信息的表示向量。

- Experimental Result:

作者提出了一种简单而有效的方法，将标签嵌入融合到 BERT 中，同时利用其固有的输入结构和自我注意机制，这种方法对相对较小中型的数据集或者类别个数较少的数据集有显著的改进。将平均 tf-idf 得分最高的子词作为补充标签文本，实验结果表明这种方法可以进一步提升分类性能。[\(Written by Qiaomei Li\)](#)

2021:

3. Learning Modality-Specific Representations with Self-Supervised Multi-Task Learning for Multimodal Sentiment Analysis(AAAI,CCF-A)

- Motivation:

多模态数据集只是给三个模态定一个统一的标签，要对每一个单独模态标注非常费时费力，但是每一个单独模态的有不同的信息，一致性和差异性是很难同时获取。

- **Method:**

本文基于自监督学习 (self-supervised learning) 提出了标签生成模块用于获得每一个单独模态的标签, 并且联合多模态和单模态学习分别学习一致性和不同性。此外, 还设计了权重调节策略(weight-adjustment strategy)引导子任务聚焦在模态之间更加不同的采样上。

- **Experimental Result:**

作者在 MOSI、MOSEI、SIMS 三个多模态情绪分析数据集上进行基准测试, 并且在具有人工标注的 SIMS 数据上验证了相比于人工标注的性能有更好的效果。([Written by Qiaomei Li](#))

钟清山

2021

1. Is Encoder-Decoder Redundant for Neural Machine Translation? ([ACL](#))

- **Motivation:**

与目标语言中的自回归语言建模任务相比, 机器翻译只是有一个额外的源句作为上下文。鉴于现在的神经语言模型已经可以处理目标语言中相当长的上下文, 我们很自然地会问, 简单地连接源句和目标句并训练一个语言模型来进行翻译是否有效?

- **Method:**

在四个机器翻译数据集上进行了实验, 用国际口语翻译会议数据集进行实验。此外, 对于多语言实验, 创建了一个自定义的多语言。对于单语数据, 我们从英语新闻爬行单语语料库 4 中抽取 500 万个句子。

- **Result:**

结果表明, 去掉编码-解码器结构, 简单地将 MT 任务作为上下文语言建模, 就足以获得最先进的翻译结果。

(written by qingshan zhong) ([BibTex](#))

2021

2. Performance-Efficiency Trade-Offs in Adapting Language Models to Text Classification Tasks ([ACL](#))

- **Motivation:**

预先训练好的语言模型 (LMs) 在适应文本分类任务时获得了最先进的性能。然而, 当在实际应用程序中使用这些模型时, 效率方面的考虑是至关重要的。

- **Method:**

分别进行了微调 (F): 在 LM 体系结构中添加一个分类层, 并将该层的参数与其余的 LM 参数共同学习; 提示 (P): LM 被训练, 根据 LM 被训练的目标, 将下游任务重新措辞为单词预测。这种技术——也被称为基于提示的微调——被发现在使用小训练集时, 其性能优于标准的微调; 微调+ KD (F+KD): 使用微调 (F) 对一个 LM 进行训练, 然后执行 KD (知识蒸馏), 得到一个更小的模型; 提示+ KD (P+KD): 使用提示 (P) 训练一个 LM, 然后执行 KD 以获得一个更小的模型。我们证明了这个训练过程可以产生既紧凑又有样本效率的分类器。

- **Result:**

研究结果表明，尽管微调和提示可以很好地训练大型列车集上的大型逻辑模块，但有更有效的替代方案可以降低计算或数据成本。研究结果可以帮助 NLP 从业者确定基于计算预算和数据收集成本使预训练的 1m 适应文本分类任务的策略。

(written by qingshan zhong) [BibTex](#)

2021

3. Recent Advances in Pre-trained Language Models: Why Do They Work and How Do They Work (ACL)

- **Motivation:**

由于近年来预训练模型的发展，作者认为有必要和及时地通过一个组织良好的教程来告知 NLP 社区关于 plm 的最新进展。

- **Method:**

从经验和理论两个方面部分解释 plm 的有效性。经验：构建了一个特殊的训练前数据集来检查 PLM 的可转移性，并绘制出将 PLM 的可转移性与训练前数据集的特征联系起来；理论：通过严格的数学方法来了解 plm 的有效性；

- **Result:**

介绍了当前的预训练模型。

(written by qingshan zhong) [BibTex](#)

梁家振

2021 :

1、A Dual-Questioning Attention Network for Emotion-cause Pair Extraction with Context Awareness (IEEE)

- **Motivation**

摘要情感-原因对提取 (ECPE) 是情感分析中的一项新兴任务，旨在提取文档中的情感对及其对应的原因。这是一个比情绪原因提取 (ECE) 更具挑战性的问题，因为它不需要情绪信号，而情绪信号在 ECE 任务中起着重要作用。现有的工作遵循两个阶段的流程，第一步识别情绪和原因，第二步配对。然而，在没有上下文信息的情况下，跨步骤和对组合的错误传播限制了有效性。为了解决这个问题，本文提出了一个双问题注意力网络来缓解这些限制。

- **Method**

本文通过注意力网络独立地向上下文提问候选情绪和原因，以获得上下文和语义上的答案。

- **Experimental result**

实证结果表明，在多个评估指标方面，本文的方法优于 baseline。(Written by Jiazhen Liang)

2、An End-to-End Network for Emotion-Cause Pair Extraction (ACL)

- Motivation

情感原因对提取 (ECPE) 的任务旨在提取文档中所有潜在的情感条款及其相应的原因。与情绪原因提取 (ECE) 这项研究更为深入的任务不同, ECPE 不要求将情绪子句作为注释提供。以前关于 ECPE 的研究要么遵循多阶段方法, 其中情感提取、原因提取和配对是独立完成的, 要么使用复杂的架构来解决其局限性。为了提高架构效率, 本文提出了 ECPE 任务的端到端模型。

- Method

由于英语 ECPE 语料库不可用, 本文改编了 NTCIR-13 ECE 语料库, 并在此数据集上建立了 ECPE 任务的基线。

- Experimental result

本文在现有的数据集上, 建议的方法产生了显著的性能改进(F1 得分增加 6.5%), 并且实现了与现有的先进技术方法竞争力相当的性能。([Written by Jiazhen Liang](#))

3、Perspective-taking and Pragmatics for Generating Empathetic

Responses Focused on Emotion Causes (EMNLP)

- Motivation

移情是一种基于他人情感状态推理的复杂认知能力。

为了更好地理解他人并在对话中表达更强烈的同理心, 本文认为必须同时解决两个问题:

- (i) 从对方的话语中识别出哪个词是引起对方情绪的原因;
- (ii) 在反应生成中反映这些特定词。

然而, 以前的识别文本中的情感原因词的方法需要子话语级注释, 这可能会要求很高。为了解决这个问题, 本文提出了一种新的解决方案。

- Method

本文从社会认知中获得灵感, 利用生成性估计器从没有词级标签的话语中推断出情感导致的词。此外, 本文引入了一种基于语用学的新方法, 使对话模型在生成过程中关注输入中的目标词。

- Experimental result

本文的方法适用于任何没有额外训练的对话模型。本文展示了我们的方法改进了多个表现最佳的对话代理, 在自动和人工评估方面产生更集中的移情反应。([Written by Jiazhen Liang](#))

黄骏

2022:

1、Utilizing BERT InterMediate Layers for MultiModal Sentiment Analysis (IEEE ICME CCF-B)

- Motivation:

使用 BERT 提取文本特征可以获得良好的上下文表示, 但是将 BERT 应用于多模态数据集的探索仍然有限, BERT 中间层可以分别捕获短语级、语法级和语义级信息, 在多模态融合阶段利用这些级别的信息可以得到细粒度的融合结果。

•Method:

作者设计了一个 BERT 外部的多模态分层融合结构,将声学 and 视觉特征作为补充信息,与 BERT 的多个中间层信息进行分层融合。为了增强分层融合框架的记忆能力,引入 GRU 来记忆各个阶段的融合特征。此外使用双重恢复损失将丢失的信息分解为情感相关和情感无关的特征。这种损失确保了在将情感相关信息恢复到网络后,特征将变得更具有鉴别性,从而促进情感分析。

•Experimental Result:

作者在 MOSI 和 MOSEI 两个基准数据集上进行实验,Acc2 比 MAG-BERT 高 1.7%,比 MISA 高 2.5%。另外通过消融实验进一步验证了利用 BERT 中间层进行模态融合的有效性。

([Written by Jun Huang](#))

2、Label Anchored Contrastive Learning for Language Understanding (NAACL CCF-B)

•Motivation:

对比学习通过自监督学习在 CV、NLP 等领域取得了惊人的进展,但是在有监督环境下的 CL 方法还没有得到充分的探索。

•Method:

作者提出一种新的标签锚定监督对比学习方法 (LaCon),该方法结合了对比学习和标签嵌入技术的优点。应用一个多头机制来捕捉文本语义的不同方面,另外加入以标签为中心的对比损失,通过以标签为锚,鼓励标签表示与小批次中属于同一类的对应实例更相似。最后通过加入标签嵌入正则化,使标签间的相似度尽可能低。

•Experimental Result:

为了验证 LaCon 的有效性,作者对八种语言理解任务进行了广泛的实验。分析了 LaCon 每个成分的贡献,也可可视化了学习的实例和标签表示,展示了每个损失成分的必要性,以及 LaCon 在表示学习上的优势,而不是交叉熵微调的 BERT。([Written by Jun Huang](#))

3、Fusion Label Enhancement for Multi-Label Learning (IJCAI CCF-A)

•Motivation:

标签增强作为一种有效挖掘不同标签隐含相对重要性信息的方法,被广泛应用与各种 MLL 任务。由于以往基于 LE 的 MLL 方法中的标签增强过程与预测模型的训练过程是解耦的,导致 LE 的目标与训练过程不匹配,最终影响整个学习系统。

•Method:

在本文中,作者提出了一种名为融合标签增强多标签学习(FLEM)的新方法,以有效地整合 LE 过程和训练过程。具体而言,本文设计了一种匹配和交互机制,利用一种新的交互标签增强损失来避免恢复的标签分布与预测模型的需求不匹配。同时,提出了一个统一的标签分布损失,以建立恢复的标签分布与预测模型训练之间的对应关系。

•Experimental Result:

大量的实验结果表明,与现有的 MLL 方法相比, FLEM 可以带来显著的性能改进。

([Written by Jun Huang](#))

张泽浩

2022:

1、MISC: A Mixed Strategy-Aware Model integrating COMET for Emotional Support Conversation(ACL,CCF-A)

- **Motivation:**

论文关注情感支持对话，情感支持对话发生在寻求者和支持者之间，支持者的目的是随着对话的进行逐渐减少寻求者的焦虑。已有的关于情感聊天的研究通过对话级的情感标签来预测用户的情感，对话内容是粗粒度和静态的。然而，情感是复杂的，用户的情感强度会在对话的发展过程中发生变化。因此，有必要在每一次说话时都说出寻求者的细微心理状态。另外，大多数共情聊天机器人都经过训练，能够根据预测的粗粒度情绪类别做出情绪回复，而不考虑如何应对寻求者的情绪问题。

- **Method:**

首先像一般情景对话模型一样将上下文进行编码，使用 COMET 模型来生成用户对话过程中心理状态信息，并引入了注意力机制来进行信息增强，可同时获得了心理状态的编码和对话级的编码，进而获得对应的响应策略，可以简单使用多层感知机对其进行分类获得各种策略的概率分布。而模型利用这个概率分布进一步对响应策略进行混合建模。

- **Experimental Result:**

通过与 SOTA 模型 BlenderBotJoint 的比较，可以发现在预测更准确的响应策略方面 MISC 更有效。BlenderBotJoint 在第一个解码步骤预测一个单一策略，而 MISC 使用策略码本建模混合响应策略，允许解码器学习并更自然地表现出同理心。比较结果表明，将反应策略作为一项额外任务进行预测，并考虑情绪支持对话的策略复杂度是有益的。([Written by Zehao Zhang](#))

2022:

2、Vision-Language Pre-Training for Multimodal Aspect-Based Sentiment Analysis(ACL,CCF-A)

- **Motivation:**

对于基于多模态方面的情感分析 (MABSA)，现有方法要么分别使用预训练的视觉、文本模型，忽略了模态间的对齐关系。要么使用经通用预训练任务的视觉-语言模型，不足以识别细粒度的 aspect、opinion 和模态间的对齐关系。

- **Method:**

为了解决这些限制，作者提出了一种针对 MABSA 的任务特定视觉语言预训练框架 (VLPMABSA)，该框架是用于所有预训练和下游任务的统一多模态编码器-解码器架构。首先构建一个基于 BART 的多模态生成框架 为视觉-语言预训练和下游 MABSA 任务。然后作者提出三种类型的视觉-语言预训练任务。为了评估本文方法的有效性，作者使用 Twitter 数据集 MVSA-Multi 实现粗粒度的文本-图片情感分析。

- **Experimental Result:**

实验结果表明，作者的方法在三个 MABSA 子任务上通常优于最先进的方法。进一步的分析证明了每个预训练任务的有效性。([Written by Zehao Zhang](#))

2022:

3、Multi-Granularity Semantic Aware Graph Model for Reducing Position Bias in Emotion Cause Pair Extraction(ACL,CCF-A)

- **Motivation:**

在典型的 ECPE 数据集中，情绪和原因的相对距离分布极不平衡，情绪-原因配对提取 (ECPE) 任务旨在从文档中提取情绪和原因作为配对。现有方法忽略了距离子句之间的有效语义连接，导致对位置不敏感数据的泛化能力较差。为了缓解这一问题，作者提出了一种新的多粒度语义感知图模型 (MGSAG)，以结合细粒度和粗粒度语义特征，而不考虑距离限制。

- **Method:**

通过关键字引入的细粒度语义和子句之间的粗粒度语义来实现子句表示学习,基于子句关键字二分图获得细粒度语义感知子句表示,基于完全连接的子句图生成粗粒度语义感知子句表示。模型由文档编码,细粒度语义感知图(FGSAG),粗粒度语义感知图(CGSAG),成对分类四个部分组成。

● **Experimental Result:**

结果表明,MGSAG 超过了现有最先进的 ECPE 模型。特别是,在位置不敏感数据的情况下,MGSAG 显著优于其他模型。。([Written by Zehao Zhang](#))

李外

1. Generation of Compound Emotions Expressions with Emotion Generative Adversarial Networks (EmoGANs)(IEEE,ACF-A)

- motivate:

人类情绪的面部表情在深入了解人类认知方面起着至关重要的作用。它们对于设计人机交互模型至关重要。尽管人类的情绪状态不限于基本情绪,如快乐、悲伤、愤怒、恐惧、厌恶和惊讶,但目前的大多数研究都集中在这些基本情绪上。

- method:

在这项研究中,我们提出了一种新的方法来创建复合情绪的面部表情,这些表情是由基本情绪的表情组合而来的。在我们的实验中,我们以无监督和有监督的方式训练我们提出的模型,即情绪生成对抗网络(EmoGAN),以提高生成图像的质量。

- experimental result:

为了证明所提出的方法的有效性,我们使用扩展的 Cohn Kanade 数据集(CK+)和日本女性面部表情数据集(JAFFE)作为输入,并可视化从我们提出的表情生成的图像。在实验中,我们提出的方法可以操纵基本的面部表情来创建复合情绪的面部表情。

(Written by WaiLi, <https://ieeexplore.ieee.org/abstract/document/9240306>)

2. Speech Driven Talking Face Generation From a Single Image and an Emotion Condition(IEEE,ACF-A)

- motivate:

视觉情感表达在视听言语交际中起着重要作用。在这项工作中,我们提出了一种在语音驱动的谈话脸生成中呈现视觉情感表达的新方法。

- method:

提出了一种在语音驱动的谈话脸生成中呈现视觉情感表达的新方法。具体而言,我们设计了一个端到端的说话脸生成系统,该系统将语音话语、单脸图像和类别情感标签作为输入,以呈现与语音同步的说话脸视频,并表达条件化的情感。对图像质量、视听同步和视觉情感表达的客观评估表明,所提出的系统优于最先进的基线系统。对视觉情感表达和视频真实性的主观评价也证明了所提出系统的优越性。此外,我们使用生成的音频和视觉模式中情绪不匹配的视频进行了人类情绪识别试点研究。

- experimental result:

结果表明,在这项任务中,人类对视觉模态的反应比音频模态更显著。(Written by WaiLi,<https://ieeexplore.ieee.org/abstract/document/9496264>)

3. Spectral Norm Normalized Network for Facial Expression Generation(IEEE,ACF-A)

- motivate:

在目前的表情识别技术研究中,识别性能下降的原因是表情数据库样本量小、样本量不均衡等问题。

- method:

在本文中，我们研究了生成对抗模型中面部表情图像的生成，重点研究了满足 Lipschitz 约束的解决方案。我们建议基于谱范数归一化网络 (SNN-EG) 生成面部表情图像。该方法满足 K -Lipschitz 正则约束，可以扩展表达样本的数量，丰富样本的多样性，实现对小样本和样本不平衡问题的解决。

- experimental result:

实验表明，该方法对人脸表情的合成和识别具有较好的效果。(Written by WaiLi, <https://ieeexplore.ieee.org/abstract/document/9350747>)