

基于神经网络的多标签情绪分类研究

吴妍秀

(四川大学计算机学院,成都 610065)

摘要:

情绪分类是自然语言处理领域中非常有研究价值的任务。情绪分类任务的目的是将情绪标签自动分配给数据集中的每个文本样例,旨在对文本发布者的情绪进行识别。由于人类的情绪具有复杂性,所以情绪分类属于多标签分类任务。针对多标签情绪分类任务,简要介绍其研究现状,基于神经网络方法的前沿进展及相关数据集和评价指标。

关键词:

多标签分类;情绪分类;神经网络

0 引言

随着全球网络社交媒体的快速发展,越来越多用户在推特、微博等社交平台上发布内容,分享自己的个人观点、态度及情绪,挖掘海量文本中的此类信息变得尤为重要,使情绪分类成为热门研究方向。情绪分类具有很高的应用价值,它可以为舆情分析^[1],个性化推荐,股票预测^[2]等任务提供支持。

情绪具有多种分类方式:Ekman 等人^[3]将基础情绪划分为 joy、sadness、anger、fear、disgust、surprise 六类,在 SemEval-2007 竞赛 task 14: Affective Text^[4]任务中,1250 条新闻标题被以这六种情绪进行标注;Plutchik^[5]提出了八种基础情绪的分类,分别为 acceptance、anger、anticipation、disgust、joy、fear、sadness、surprise。如表 1 所示,一段文本中可能包含多种情绪,且情绪之间具有关联,识别句子中同时存在的多种情绪仍是一项具有挑战性的任务。

多标签分类任务的定义为:假设 $X = R^d$ 为 d 维的样本空间, $Y = \{y_1, y_2, \dots, y_q\}$ 为存在 q 个标签的标签空间,多标签分类任务从训练集中学习函数映射 $h: X \rightarrow 2^Y$ 。

表 1 SemEval2018 task 1^[6]数据集中的推特及对应情绪

推特	情绪标签
Whatever you decide to do make sure it makes you #happy.	joy, love, optimism
My roommate: it's okay that we can't spell because we have autocorrect. #terrible #firstworldprobs	anger, disgust

1 研究现状

1.1 多标签分类任务

多标签分类任务主要采用问题转化或算法适应的方法,问题转化方法将多标签分类问题转化为其他较为熟悉的分类问题,算法适应方法采用现有的机器学习技术加以改进来处理多标签数据。

二元关系法(BR)^[7]将多标签分类问题分解为 q 个独立的二分类问题,每个二分类任务对应标签空间中一个可能的标签;多标签 K 近邻(MLKNN)^[8]采用 K 近邻算法作为基础处理多标签数据,通过邻居节点的标签信息利用最大后验进行预测。上述两种方法为一阶方法,单独预测每个标签,忽略了标签的共现性,未考虑标签之间的关联,造成信息丢失。

基于排序的方法有排序-SVM^[9]及校准的标签排序算法(CLR)^[10],排序-SVM 算法采用最大间隔来处理多标签数据,对 q 个线性分类器进行优化,最小化排序损失,并利用核函数解决非线性情况;CLR 将多标签分类问题转化为标签排序问题,为每个标签对构建二分类器,共构建 $q(q-1)/2$ 个二分类器,通过人工插入的阈值标签得到最后的分类结果。上述两种方法为二阶方法,只考虑了两个标签之间的共现性,未考虑真实世界中的标签关联不止存在于二者之间。

分类器链(CC)^[11]将多标签分类问题转化为链式的多个二分类问题,链中前一个分类器的分类结果将输

入后续的分类器中;标签幂集(LP)^[12]将问题转化为单标签多分类问题,每个可能的标签组合被视为一个单独的标签,使标签数量级增长至 2^n 。上述两种方法为高阶方法,可以考虑标签之间的关联,但CC在训练过程中使用teacher forcing(即分类器的输入为上一个分类器的黄金标注,而不是其输出的实际预测结果)会导致预测时的暴露误差,LP转化问题的方式使标签数量呈幂级增长导致计算困难。

1.2 情绪分类任务

早期使用基于词典与规则的方法解决该问题:Balahur等人^[13]构建了一个能触发情绪的常识规则库EmotiNet,该规则库记录了各种触发情绪的事件,进而对没有明确提到情绪词汇的文本进行识别。但是人工构建词典及规则费时费力,覆盖率仍不够理想,且无法探测新词。

Almeida等人^[14]将BR、CC、LP、MLKNN等多标签分类任务的方法直接用于情绪多标签分类,传统的多标签分类方法存在未考虑标签关联或计算复杂的问题。

Li等人^[15]统计了Ren-CECps数据集^[16]中的标签共现频率及上下文出现相同标签的频率,提出了依赖因子图(DFG)模型,可以建模情绪标签及上下文之间的相关性。Zhu等人^[17]提出了语料库融合的方法利用两个情绪语料库,首先对两个语料库进行有监督的分类任务,然后根据语料库内约束和语料库外约束对预测结果进行优化。

2 神经网络方法

近年来,基于神经网络的方法快速发展,从传统机器学习人工提取特征的方式变为自动学习特征,由于自然语言序列输入且不定长的特性,可以使用循环神经网络RNN这类神经元节点按时间序列连接的神经网络,而长短时记忆网络LSTM及门控循环单元GRU改进了RNN梯度消失的问题,使用LSTM或GRU编码输入句子成为了主流。

Abdul-Mageed等人^[18]利用hashtag信息对语料进行远监督标注,使用GRU对文本进行细粒度的情绪多分类任务;Matsumoto等人^[19]利用emoji表情符号监督标注语料,分别实验BiLSTM、BiGRU、CNN三种神经网络的分类效果,最终效果BiLSTM及BiGRU优于

CNN。上述研究虽然使用远监督学习解决了人工标注语料费时费力,且标注数据不足的问题,但是将情绪分类看作普通的多分类问题,未考虑真实世界中的情绪分类应为多标签分类问题。

Yu等人^[20]通过迁移学习,利用情感分类任务提升样本较少的多标签情绪分类任务的效果,编码端使用BiLSTM将句子编码至两个特征空间,分别为共享空间及情绪任务专有空间,通过双向注意力正交优化,捕获情感词及情绪专有词的信息,最后将两个隐藏向量拼接传入多层感知机进行分类,但需要人工设定阈值得到最后的多个标签。Jabreel等人^[21]拼接句子和标签,形成句子标签对集合,使用标签嵌入作为查询向量计算每个词的注意力,输入BiGRU进行联合编码,将多标签分类问题转化为由同一个分类模型解决的二分类问题。上述研究未考虑标签与标签之间的共现相关性。

在LSTM的基础上,可以使用Seq2Seq^[22-23]进行多标签分类任务,将多标签分类问题转化为序列标注问题,该模型为编码-解码结构,编码端使用BiLSTM提取输入句子的特征,注意力机制通过聚焦文本序列的不同词,对词的隐藏状态求和生成上下文向量;解码端每个时刻接收上一个时刻的隐藏状态,及上一时刻的标签嵌入和该时刻由编码端得到的上下文向量的拼接,将生成过的标签进行mask操作从而避免重复生成,最终按序列生成情绪标签。该研究将解码端接收的上一时刻的标签嵌入改为了概率值加权的标签嵌入,减弱模型teacher forcing造成的暴露误差。

Fei等人^[24]提出了潜在情绪记忆网络得到先验的情绪分布,其中的潜在情绪模块使用变分自编码器重构输入的词袋来学习情绪分布;记忆模块使用潜在情绪模块得到的参数矩阵作为查询向量,对输入句子计算注意力,捕捉对应情绪相关的上下文特征。模型使用了多跳BiGRU得到输出,每个时刻输出一个情绪标签的二分类结果,潜在情绪模块得到的情绪分布矩阵及情绪特征矩阵为整个模型共享,记忆模块仅针对BiGRU的每个时刻。

3 数据集及评价指标

3.1 数据集

近年来,由于该任务受到广泛关注,有多个研究及竞赛提出并标注了数据集,如:SemEval-2007数据集、

NLPCC-2014 数据集、Ren-CECps 数据集、SemEval-2018 数据集等。

目前使用较多的情绪多标签分类数据集为英文微博文本的 SemEval-2018 task 1 的竞赛数据集及中文博客的 Ren-CECps 数据集。如表 2 所示, SemEval-18 数据集已划分训练集、验证集、测试集, 情绪标签为 disgust、anger、joy、sadness、optimism、fear、anticipation、pessimism、love、surprise、trust 十一类; CEC 数据集为未划分的结构化数据, 情绪标签为 joy、hate、love、sorrow、anxiety、surprise、anger、expect 八类。

统计得出的句子对应标签个数如表 3 所示, 两个数据集的每个句子样例均包含一种或多种情绪。

表 2 数据集情况

数据集	句子数	训练集	验证集	测试集	标签数
SemEval2018	10983	6838	886	3259	11
Ren-CECps	35096	-	-	-	8

表 3 情绪标签数量的分布情况

每个句子情绪标签个数	句子数 (Ren-CECps)	句子数 (SemEval2018)
1	18812	1563
2	11416	4442
3	1824	3419

参考文献:

- [1]LIU Z M, LIU L. Public emotion monitoring and surveillance in emergencies[J]. Systems Engineering, 2010, 7(15): 65-73.
- [2]DONG Y H, CHEN H, QIAN W N, et al. Micro-blog social moods and chinese stock market: the influence of emotional valence and arousal on shanghai composite index volume[J]. International Journal of Embedded Systems, 2015, 7(2): 148.
- [3]EKMAN P, FRIESEN W V. Universal and cultural differences in facial expression of emotion[C]//Nebraska Symposium of Motivation, 1972.
- [4]STRAPPARAVA C, MIHALCEA R. Semeval-2007 task 14: Affective text[C]//Proceedings of the 4th International Workshop on Semantic Evaluations, 2007.
- [5]PLUTCHIK R. A general psychoevolutionary theory of emotion[J]. Emotion Theory Research & Experience, 2000, 21(4-5): 529-553.
- [6]MOHAMMAD S M, BRAVO-MARQUEZ F, SALAMEH M, et al. SemEval-2018 Task 1: Affect in Tweets[C]//International Workshop on Semantic Evaluation (SemEval-2018), 2018.
- [7]BOUTELL M R, LUO J B, SHEN X P, et al. Learning multi-label scene classification[J]. Pattern Recognition, 2004, 37(9): 1757-1771.
- [8]ZHANG M L, ZHOU Z H. ML-KNN: a lazy learning approach to multi-label learning[J]. Pattern Recognition, 2007, 40(7): 2038-2048.
- [9]ELISSEEFF A E, WESTON J. A kernel method for multi-labelled classification[C]//Neural Information Processing Systems 14, 2002.
- [10]FÜRNKRANZ J, HÜLLERMEIER E, LOZAMENCÍA E, et al. Multilabel classification via calibrated label ranking[J]. Machine Learning, 2008, 73(2): 133-153.
- [11]READ J, PFAHRINGER B, HOLMES G, et al. Classifier chains for multi-label classification[C]//Proceedings of the European Conference on Machine Learning and Knowledge Discovery in Databases, 2009.
- [12]TSOUMAKAS G, KATAKIS I, VLAHAVAS I. Random k-labelsets for multilabel classification[J]. IEEE Transactions on Knowledge

3.2 评价指标

针对多标签分类问题,除了传统的 F1 值度量以外,可以使用汉明损失^[25]、杰卡德系数进行度量。

汉明损失用于评价被误分类的样本标签对:

$$HL = \frac{1}{n} \sum_{i=1}^n \frac{1}{l} XOR(h(x_i), Y_i)$$

其中 n 代表样本个数, l 代表标签个数, Y_i 为真实标注, $h(x_i)$ 为预测结果, 汉明损失越小, 代表模型预测结果越准确。

杰卡德系数用于衡量两个标签集合之间的相似性:

$$Ja = \frac{1}{n} \sum_{i=1}^n \frac{|h(x_i) \cap Y_i|}{|h(x_i) \cup Y_i|}$$

当集合 Y_i 、 $h(x_i)$ 都为空时, Ja 定义为 1。杰卡德系数越大, 代表模型预测结果越准确。

4 结语

本文对基于神经网络的多标签情绪分类方法进行了研究。多标签情绪分类对于业界挖掘用户信息有很大帮助, 神经网络的发展使该任务的效果有了进一步提升。但由于神经网络的黑盒特性, 仍有很多特征提取及解码方式值得探索。

- and Data Engineering, 2011, 23(7): 1079–1089.
- [13] BALAHUR A, HERMIDA J M, MONTORO A. Detecting Implicit expressions of sentiment in text based on commonsense knowledge [C]//Proceedings of the 2nd Workshop on Computational Approaches to Subjectivity and Sentiment Analysis (ACL-HLT 2011), 2011.
- [14] ALMEIDA A M G, CERRI R, PARAISO E C, et al. Applying multi-label techniques in emotion identification of short texts[J]. Neuro-computing, 2018, 320(DEC.3): 35–46.
- [15] LI S S, HUANG L, WANG R, et al. Sentence-level emotion classification with label and context dependence[C]//Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing, 2015.
- [16] QUAN C Q, REN F J. Sentence emotion analysis and recognition based on emotion words using ren-CECps[J]. International Journal of Advanced Intelligence Paradigms, 2010, 2(1): 105–117.
- [17] ZHU S Y, LI S S, CHEN Y, et al. Corpus fusion for emotion classification[C]//Proceedings of the 26th International Conference on Computational Linguistics (COLING 2016), 2016.
- [18] ABDUL-MAGEED M, UNGAR L. EmoNet: fine-grained emotion detection with gated recurrent neural networks[C]//Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics, 2017.
- [19] MATSUMOTO K, YOSHIDA M, KITA K. Classification of emoji categories from tweet based on deep neural networks[C]//the 2nd International Conference, 2018.
- [20] YU J F, MARUJO L, JIANG J, et al. Improving multi-label emotion classification via sentiment classification with dual attention transfer network[C]//Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing, 2018.
- [21] JABREEL M, MORENO A. A deep learning-based approach for multi-label emotion classification in tweets[J]. Applied Ence, 2019, 9(6): 1123.
- [22] SUTSKEVER I, VINYALS O, LE Q V. Sequence to sequence learning with neural networks[J]. Advances in Neural Information Processing Systems, 2014.
- [23] YANG P C, SUN X, LI W, et al. SGM: sequence generation model for multi-label classification[C]//COLING 2018, 2018.
- [24] FEI H, ZHANG Y, REN Y F, et al. Latent emotion memory for multi-label emotion classification[C]//Proceedings of the 34th Association for the Advancement of Artificial Intelligence (AAAI), 2020.
- [25] ZHANG M L, ZHOU Z H. A review on multi-label learning algorithms[J]. IEEE Transactions on Knowledge & Data Engineering, 2014, 26(8): 1819–1837.

作者简介:

吴妍秀(1996–),女,重庆人,硕士研究生,研究方向为自然语言处理

收稿时间:2020-11-17 修稿时间:2020-12-30

Survey of Multi-Label Emotion Classification Based on Neural Network

WU Yan-xiu

(College of Computer Science, Sichuan University, Chengdu 610065)

Abstract:

Emotion classification is a valuable research task in the field of natural language processing. The purpose of emotion classification task is to automatically assign the emotion labels to each text in the dataset, aiming to identify the emotion of text publishers. Because of the complexity of human emotions, emotion classification is a multi-label classification task. For multi-label emotion classification task, this paper briefly introduces the research status, the state-of-the-art neural network based methods, relevant datasets and evaluation metrics.

Keywords:

Multi-Label Classification; Emotion Classification; Neural Network