

度-可控性 3 个模型的详细性能.以极性-强度模型为例,表中维度 1 表示前一维度——极性维度回归性能,维度 2 表示后一维度——强度维度回归性能.表 4 中 AdvNN 的每个维度的性能取自表 5 中各组合中皮尔森相关系数 r 最高的一项,例如读者情绪的极性维度在极性-可控性模型中的回归结果的 r 值较极性-强度模型中的 r 值更高,因此,表 4 中记录该最优 r 值以及相应的 MSE 值.

Table 4 The performances of various approach on the News domain
表 4 各种方法在新闻领域上的情绪回归性能

基准系统		极性		强度		可控性	
		r	MSE	r	MSE	r	MSE
读者情绪	SVR(BOW)	0.309	0.206	0.130	0.098	0.116	0.086
	SVR(TF-IDF)	0.307	0.212	0.139	0.094	0.155	0.094
	SVR(W2V)	0.272	0.198	0.144	0.094	0.139	0.090
	ANN	0.288	0.219	0.150	0.088	0.136	0.091
	LSTM	0.324	0.205	0.127	0.098	0.147	0.090
	C-LSTM	0.337	0.190	0.143	0.094	0.168	0.087
	AdvNN(r 最优)	0.353	0.182	0.168	0.091	0.179	0.080
作者情绪	SVR(BOW)	0.251	0.114	0.092	0.093	0.129	0.059
	SVR(TF-IDF)	0.263	0.122	0.065	0.092	0.112	0.058
	SVR(W2V)	0.242	0.121	0.081	0.095	0.143	0.055
	ANN	0.217	0.111	0.060	0.084	0.127	0.057
	LSTM	0.225	0.111	0.074	0.087	0.121	0.061
	C-LSTM	0.242	0.129	0.091	0.091	0.148	0.068
	AdvNN(r 最优)	0.290	0.120	0.093	0.085	0.160	0.052

从表 4 可以看出:

1) 总体而言,在新闻领域上,所有方法取得的 r 值都较低.最高的 r 值为 AdvNN 在读者极性情绪上取得的 0.353.作为参考,EMOBANK 全领域上标注者两两间标注结果的平均 r 值约为 0.6^[4].由此可见,EMOBANK 上的多维度情绪回归任务是一个较难的任务,不容易取得非常高的性能.注意到该领域上的作者强度情绪最高 r 值未超过 0.1.造成这一结果的可能原因在于新闻领域文本的作者撰写相关文本时倾向于不代入较强的情绪.例如,例句 3 的作者强度情绪为 3.25,为中等强度,而读者强度情绪为 4.00,为较高强度.例句 3 作为一条典型的报道负面消息的新闻文本,能够激发起读者较强的负面情绪.而作者在撰写这条文本时为了保证新闻的客观性,相比读者不会代入强烈的情绪(尽管例句 3 的作者情绪极性也为负面).同样,例句 4 作为一条报道正面消息的新闻文本,其读者强度情绪为较高的 3.80,而作者强度情绪为中等的 3.40.例句 5 说明在读者强度情绪较弱情况(2.00)下,作者情绪依然趋于中等(2.80).可见,在新闻领域上,作者情绪倾向于为一个适中的强度.

2) 与传统方法(SVR)相比,一般的基于神经网络的回归模型表现并无明显优势.其中,ANN 在所有基准系统中的总体性能最差,各情绪维度的平均 r 值为 0.163,低于其他基准系统.相比之下,SVR 方法的平均 r 值在 0.171 上下,优于 ANN,相比 LSTM 取得的平均 r 值 0.174 亦无太大劣势.而 C-LSTM 取得了基准系统中最高的平均 r 值 0.184,说明了文本分类任务中的先进系统在回归任务上亦能取得较好的效果.

3) 在新闻领域上,无论是读者情绪还是作者情绪,AdvNN 的最高 r 值均超过了其他所有基准系统的 r 值.其中,在作者极性维度上的提升最多,较该维度上最强的基准系统 SVR(TF-IDF)提升了 0.027. AdvNN 在作者强度维度回归上的提升最小,较该维度上最强的基准系统 SVR(BOW)仅有 0.001 的提升.此外,AdvNN 在新闻领域的平均 r 值达到 0.207,亦超过了平均 r 值最高的基准系统 C-LSTM.从这一结果来看,同样是在 LSTM 的基础上实现的神经网络模型,使用了对抗性学习的 AdvNN 相比融合了卷积神经网络的 C-LSTM 在本文的任务上能够进一步学习文本中的情绪维度特征,从而提高系统性能.

4) 除 r 值之外,AdvNN 在 MSE 评估指标下的表现也较为优秀.除读者强度、作者极性和作者强度之外,AdvNN 均取得了最低的 MSE 值.尽管在作者强度维度上的 r 值提升较低,AdvNN 与 SVR(BOW)相比 MSE 下降了 0.008,但总体而言,在该维度上的性能却存在一定的提升. AdvNN 取得了平均 0.102 的 MSE 值,亦低于所有基准系统的平均 MSE 值.

例句 3: *Scam lures victims with free puppy offer.* (骗局使用提供免费的小狗的方式来引诱受害者.)

例句 4: Kidnapped AP photographer freed in Gaza. (被绑架的美联社摄影师在加扎被释放)

例句 5: I feel bad for Jorge he's a good friend of mine, Bengie said. (本吉说到: 我为霍赫感到难过, 他是我的好朋友)

Table 5 Detailed performances of AdvNNs for three combinations of emotion dimensions on the News domain

表 5 新闻领域上 3 种情绪维度组合的对抗式神经网络详细性能

	维度组合相应模型	维度 1		维度 2	
		r	MSE	r	MSE
读者情绪	极性-强度模型	0.339	0.190	0.168	0.091
	极性-可控性模型	0.353	0.182	0.174	0.083
	强度-可控性模型	0.162	0.097	0.179	0.080
作者情绪	极性-强度模型	0.290	0.120	0.093	0.085
	极性-可控性模型	0.287	0.118	0.160	0.052
	强度-可控性模型	0.090	0.085	0.149	0.061

进一步分析表 5, 可以发现:

1) 除了强度-可控性模型中作者强度情绪的 r 值低于两个基准系统之外, 各维度组合整体结果相比基准系统均有一定的提升. 其中, AdvNN 在读者强度情绪和作者极性情绪上的非最优 r 值亦超过最强基准系统至少 0.019.

2) 除作者强度情绪外, 其余情绪维度在不同组合下的平均 r 值亦超过相应维度的最强基准系统结果. 尽管在大多数情绪维度上 AdvNN 的所有维度组合性能超过了最强基准系统, 然而由于某些维度组合的 r 值提升幅度较小, 如极性-强度模型的读者情绪相比该维度最强基准系统 C-LSTM 仅有 0.002 的 r 值提升. 因此, 尽管 AdvNN 在新闻领域上的各组合的最优性能均超过了最强基准系统, 但在个别情况下的提升并不稳定.

3) AdvNN 在作者强度情绪这一维度上的较低性能也进一步证明, 由于新闻文本本身的特性, 新闻作者在撰写相关文本时为了维持新闻的相对客观性, 并不会产生明显的情绪强度上的波动, 因而难以在该维度上取得较好的回归结果.

表 6 给出了 AdvNN 与其他基准系统在小说领域上的情绪回归性能, 具体结果亦为 5 倍交叉实验的平均值. 表 7 进一步给出了 3 种情绪维度两两组合的详细性能. 从表 6 中可以看出:

1) 与新闻领域相比, 在小说领域上, 所有方法取得的 r 值亦较低. 而与新闻领域不同的是, 小说领域的结果之中无很低 (小于 0.1) 的 r 值. 造成这一结果的原因可能在于小说作者在撰写文本时会代入一定程度的个人情绪. 例如, 例句 6 标注了负面的读者情绪和作者情绪, 而读者强度情绪和作者强度情绪分别为 4.20 与 4.20, 均为较强的情绪. 而例句 7 的读者强度情绪和作者强度情绪分别为 2.67 与 2.20, 均为较弱的情绪. 由此可见, 小说领域上的作者情绪与新闻领域上的不同, 是和读者情绪一样会产生较为明显的强度波动. 因此, 实验中所有系统相对来说在小说领域强度维度上的 r 值均高出新闻领域很多. 此外, 整体来看, 所有系统取得的 r 值与新闻领域上一样并不高, 亦证明了 EMOBANK 语料上的多维度情绪回归任务较难.

2) 不同于新闻领域, ANN 在小说领域上的性能较好, 其平均 r 值达到了 0.184, 在基准系统中仅次于 C-LSTM 的 0.190. LSTM 在小说领域上的表现明显差于其他基准系统, 平均 r 值仅为 0.160, 尤其是在读者极性维度和读者强度维度上与其他基准系统相比有较大差距. 而 SVR 在小说领域的性能依然稳定, 3 种特征输入得到的结果平均 r 值均在 0.170 以上, 最高为 SVR(TF-IDF) 取得的 0.182, 进一步证明了传统的 SVR 在回归任务中是不亚于一般的神经网络模型的先进系统.

3) AdvNN 在小说领域上亦取得了所有情绪维度下最高的 r 值, 并且在读者情绪与作者情绪的所有维度下较 r 值最高的基准系统有至少 0.019 的明显提升. 此外, AdvNN 在小说领域的平均 r 值达到了 0.219, 亦超过了平均 r 值最高的基准系统 C-LSTM. 在小说领域上的结果进一步证明, AdvNN 相比 C-LSTM 在本文的任务上能够进一步提升基于 LSTM 的神经网络回归模型的性能.

4) 小说领域上, AdvNN 在 MSE 评估指标下的表现也较为良好, 并在作者极性维度和作者强度维度中取得了最低的 MSE 值. 除了在读者可控性维度中的 MSE 值高出 MSE 值最低的基准系统 C-LSTM 较多之外, 在其他

未取得最低 MSE 值的情绪维度上相比该维度的最低值差距亦较小.然而与新闻领域不同的是,AdvNN 在所有情绪维度下取得了 0.124 的平均 MSE 值,与 SVR(TF-IDF)相同,略高于 C-LSTM 取得的 0.122,总体而言并未取得最佳的 MSE 值.

Table 6 The performances of various approach on the Fictions domain

表 6 各种方法在小说领域上的情绪回归性能

基准系统		极性		强度		可控性	
		r	MSE	r	MSE	r	MSE
读者情绪	SVR(BOW)	0.220	0.200	0.191	0.167	0.144	0.112
	SVR(TF-IDF)	0.200	0.176	0.196	0.122	0.144	0.105
	SVR(W2V)	0.217	0.187	0.162	0.124	0.166	0.110
	ANN	0.228	0.181	0.201	0.120	0.157	0.102
	LSTM	0.185	0.188	0.132	0.143	0.152	0.102
	C-LSTM	0.229	0.179	0.202	0.120	0.162	0.098
	AdvNN(r 最优)	0.253	0.181	0.225	0.122	0.193	0.110
作者情绪	SVR(BOW)	0.203	0.136	0.154	0.119	0.157	0.092
	SVR(TF-IDF)	0.209	0.130	0.183	0.119	0.161	0.091
	SVR(W2V)	0.199	0.135	0.132	0.120	0.155	0.087
	ANN	0.187	0.146	0.170	0.115	0.164	0.085
	LSTM	0.182	0.141	0.157	0.117	0.147	0.092
	C-LSTM	0.200	0.128	0.172	0.120	0.173	0.087
	AdvNN(r 最优)	0.243	0.125	0.202	0.114	0.199	0.089

例句 6: *She screamed I haven't socialized with Terra's elite for most of my life.* (她尖叫到:我人生中大部分时间都没能与地球的精英社交.)

例句 7: *The only other illumination came from a lurid moonlight filtered through thin branches and clouds casting its bone-pale glow onto the pine floorboards.* (唯一的其他照明来自一缕透过稀薄的树枝和云彩的月光,将其苍白的的光芒投射到松木地板上.)

进一步分析表 7,可以发现:

1) 在小说领域上,AdvNN 在所有维度上的非最优 r 值依然高于最强基准系统,并未出现新闻领域上 AdvNN 在个别维度的非最优 r 值低于基准系统的情况.其中,读者极性情绪、读者可控性情绪以及作者极性情绪上 AdvNN 的非最优 r 值相比最强基准系统都有至少 0.018 的提升.

2) 各情绪维度在不同模型中的平均 r 值亦均超过相应维度的最强基准系统结果,其中,AdvNN 在作者极性情绪上取得的平均 r 值达到 0.239,相比该维度上的最强基准系统 SVR(TF-IDF)的提升达到非常显著的 0.030.

3) 通过上述结果可以看出,小说领域的文本由于领域的特性,其情绪分析的难度相对新闻领域较低.两个领域上的结果对比说明了读者在阅读一段包含情绪的文本时并不受领域差异性的影响,对于情绪的表达会有较明显的波动,使得情绪分析系统能够较好地从中挖掘相关的特征来完成情绪回归的任务.而作者情绪则会较明显地受领域差异性的影响,尤其是作者强度情绪.小说领域的作者情绪强度不同于新闻领域,作者不需要保持情绪上的客观性,因此小说领域的作者强度情绪和读者强度情绪一样,能够比较容易被情绪回归系统识别.

Table 7 Detailed performances of AdvNNs for three combinations of emotion dimensions on the Fictions domain

表 7 小说领域上 3 种情绪维度组合的对抗式神经网络详细性能

	维度组合相应模型	维度 1		维度 2	
		r	MSE	r	MSE
读者情绪	极性-强度模型	0.253	0.181	0.211	0.128
	极性-可控性模型	0.249	0.180	0.193	0.110
	强度-可控性模型	0.225	0.122	0.184	0.119
作者情绪	极性-强度模型	0.235	0.136	0.202	0.114
	极性-可控性模型	0.243	0.125	0.199	0.089
	强度-可控性模型	0.189	0.119	0.182	0.094

3.5 显著性测试

除了上述的情绪回归实验结果之外,我们进一步给出了本文所提方法的输出结果与 r 值最优的基准系统输出结果之间的配对双样本 t 检验结果.表 8 与表 9 分别给出了新闻与小说领域中各情绪维度上双样本 t 检验的 p 值.与实验结果一样,表 8 与表 9 中的 p 值亦为 5 倍交叉验证每一组测试集上输出结果之间显著性测试的平均值.

Table 8 Results of dependent t -test for paired samples between AdvNN and strongest baselines for each emotion dimension in the News domain
表 8 新闻领域上 AdvNN 与各情绪维度下最强基准系统之间的配对双样本 t 检验结果

		p
读者情绪	极性(vs. C-LSTM)	2.65e-11
	强度(vs. ANN)	6.03e-11
	可控性(vs. C-LSTM)	3.15e-7
作者情绪	极性(vs. SVR(TF-IDF))	1.58e-23
	强度(vs. SVR(BOW))	0.082
	可控性(vs. C-LSTM)	9.10e-9

Table 9 Results of dependent t -test for paired samples between AdvNN and strongest baselines for each emotion dimension in the Fictions domain
表 9 小说领域上 AdvNN 与各情绪维度下最强基准系统之间的配对双样本 t 检验结果

		p
读者情绪	极性(vs. C-LSTM)	1.12e-28
	强度(vs. C-LSTM)	2.20e-30
	可控性(vs. SVR(W2V))	1.15e-28
作者情绪	极性(vs. SVR(TF-IDF))	6.33e-36
	强度(vs. SVR(TF-IDF))	6.19e-14
	可控性(vs. C-LSTM)	4.11e-21

从表 8 所示结果可以看出,在新闻领域中,除了作者强度维度之外,本文提出的 AdvNN 相比各情绪维度上最强的基准系统有显著的性能提升($p < 0.05$).仅在作者强度维度上,AdvNN 相比最强基准系统 SVR(BOW)的提升不明显($p > 0.05$).而表 9 的结果则表明,在小说领域中的所有情绪维度上,AdvNN 相比各维度上的最强基准系统均有显著的性能提升.显著性测试的结果表明,本文提出的 AdvNN 相比其他基准系统不仅在 r 值评估指标下有明显的数值上的提升,并且这一提升具有统计显著性,从而进一步证明了 AdvNN 在 EMOBANK 语料中多维度情绪回归任务上的优越性.

4 结论与未来工作

本文针对多维度情绪回归这一任务提出了一种基于对抗式神经网络的多维度情绪回归方法.该方法包括了对抗式神经网络和对抗式学习算法.对抗式神经网络包括 3 个部分:特征抽取器、回归器以及判别器.以在一对情绪维度上进行对抗式学习的情绪回归模型为例,首先,本文实现了两个基于 LSTM 神经网络模型的特征抽取器,以分别从输入的文本词向量矩阵中抽取与特定情绪维度相关的特征向量.其次,本文使用两个基于全连接神经网络的回归器,以分别根据特征抽取器输出的特征对特定情绪维度进行打分.最后,本文实现了一个基于全连接神经网络的判别器,以判别特征抽取器抽取的特征属于哪个情绪维度.该方法借助判别器,在两个情绪维度的特征抽取器之间通过对抗式学习算法进行对抗式训练,从而使两个特征抽取器能够从文本输入中抽取泛化性更好的针对不同情绪维度的特征向量,继而获得更好的回归性能.在 EMOBANK 语料的新闻和小说两个领域上的实验结果表明,无论是作者情绪回归还是读者情绪回归,本文的方法在 3 个情绪维度上的 r 值均超过了所有的基准系统,其中包括了文本回归领域中的先进系统 ANN.除了文本回归先进系统之外,本文还将文本分类领域的先进系统 C-LSTM 进行修改后用于本文针对的多维度情绪回归任务进行比较,而本文提出的方法依然在所有情绪维度上取得了更好的 r 值.除实验结果之外,AdvNN 与各情绪维度上最强的基准系统间的显著性测试也表明,AdvNN 取得的性能提升具有统计上的显著性,进一步证明了 AdvNN 在 EMOBANK 新闻与小说领域上的多维度情绪回归任务上的优越性.

然而,本文的方法目前仅为两个维度之间的对抗,并且在一些维度下取得的提升并不明显(例如新闻领域上的作者强度情绪维度).此外,一些情绪维度的对抗组合取得的性能提升也并不非常理想,甚至有个别组合在某个情绪维度下取得的 r 值相比部分基准系统有所下降.因此,我们未来的工作将首先集中于同时进行三维度之间的对抗.此外,我们还将改进对抗式学习框架与算法,例如引入共享参数的神经网络层来进一步学习抽取不同维度间的公共特征,以期能够在所有情绪维度对抗的组合上都取得稳定且明显的提升.

References:

- [1] Mishne G. Experiments with mood classification in blog posts. In: Shlomo A, Jussi K, James G. S, eds. Proc. of the ACM SIGIR 2005 Workshop on Stylistic Analysis of Text for Information Access. New York: Association for Computing Machinery, 2005. 321–327.
- [2] Yao Y, Wang S, Xu R, Liu B, Gui L, Lu Q, Wang X. The construction of an emotion annotated corpus on microblog text. Journal of Chinese Information Processing, 2014,28(5):83–91 (in Chinese with English abstract).
- [3] Abdul-Mageed M, Ungar L. Emonet: Fine-grained emotion detection with gated recurrent neural networks. In: Barzilay R, Kan M Y, eds. Proc. of the 55th Annual Meeting of the Association for Computational Linguistics, Vol. 1. Stroudsburg: Association for Computational Linguistics, 2017. 718–728. [doi: 10.18653/v1/P17-1067]
- [4] Buechel S, Hahn U. EMOBANK: Studying the impact of annotation perspective and representation format on dimensional emotion analysis. In: Lapata M, Blunsom P, Koller A, eds. Proc. of the 15th Conf. of the European Chapter of the Association for Computational Linguistics, Vol. 2. Stroudsburg: Association for Computational Linguistics, 2017. 578–585.
- [5] Yang C, Lin KHY, Chen HH. Emotion classification using Web blog corpora. In: Yao Y, ed. Proc. of the 2007 IEEE/WIC/ACM Int'l Conf. on Web Intelligence. Washington: IEEE Computer Society, 2007. 275–278. [doi: 10.1109/WI.2007.51]
- [6] Tripathi S, Acharya S, Sharma RD, Mittal S, Bhattacharya S. Using deep and convolutional neural networks for accurate emotion classification on deap dataset. In: Singh SP, Markovitch S, eds. Proc. of the 31st AAAI Conf. on Artificial Intelligence. Menlo Park: Association for the Advancement of Artificial Intelligence, 2017. 4746–4752.
- [7] Ekman P. Chapter 3: Basic emotions. In: Handbook of Cognition and Emotion. Hoboken: John Wiley & Sons, 1999. 45–60. [doi: 10.1002/0470013494.ch3]
- [8] Huang L, Li S, Zhou G. Emotion corpus construction on microblog text. In: Proc. of the 16th Workshop on Chinese Lexical Semantics. Heidelberg: Springer-Verlag, 2015. 204–212 (in Chinese with English abstract).
- [9] Zhu S, Li S, Chen Y, Zhou G. Corpus fusion for emotion classification. In: Matsumoto Y, Prasad R, eds. Proc. of the COLING 2016, the 26th Int'l Conf. on Computational Linguistics: Technical Papers. Osaka, 2016. 3287–3297.
- [10] Barrett LF. Solving the emotion paradox: Categorization and the experience of emotion. Personality and Social Psychology Review, 2006,10(1):20–46.
- [11] Duong L, Cohn T, Bird S, Cook P. Low resource dependency parsing: Cross-lingual parameter sharing in a neural network parser. In: Zong C, Strube M, eds. Proc. of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th Int'l Joint Conf. on Natural Language Processing, Vol. 2. Stroudsburg: Association for Computational Linguistics, 2015. 845–850.
- [12] Liang C, Berant J, Le Q, Forbus KD, Lao N. Neural symbolic machines: Learning semantic parsers on freebase with weak supervision. In: Barzilay R, Kan MY, eds. Proc. of the 55th Annual Meeting of the Association for Computational Linguistics, Vol. 1. Stroudsburg: Association for Computational Linguistics, 2017. 23–33. [doi: 10.18653/v1/P17-1003]
- [13] Wang J, Yu LC, Lai KR, Zhang X. Dimensional sentiment analysis using a regional CNN-LSTM model. In: Erk K, Smith NA, eds. Proc. of the 54th Annual Meeting of the Association for Computational Linguistics, Vol. 2. Stroudsburg: Association for Computational Linguistics, 2016. 225–230.
- [14] Mishra A, Dey K, Bhattacharyya P. Learning cognitive features from gaze data for sentiment and sarcasm classification using convolutional neural network. In: Barzilay R, Kan MY, eds. Proc. of the 55th Annual Meeting of the Association for Computational Linguistics, Vol. 1. Stroudsburg: Association for Computational Linguistics, 2017. 377–387. [doi: 10.18653/v1/P17-1035]
- [15] Cai R, Zhang X, Wang H. Bidirectional recurrent convolutional neural network for relation classification. In: Erk K, Smith NA, eds. Proc. of the 54th Annual Meeting of the Association for Computational Linguistics, Vol. 1. Stroudsburg: Association for Computational Linguistics, 2016. 756–765.
- [16] Yu M, Yin W, Hasan KS, Santos C, Xiang B, Zhou B. Improved neural relation detection for knowledge base question answering. In: Barzilay R, Kan MY, eds. Proc. of the 55th Annual Meeting of the Association for Computational Linguistics, Vol. 1. Stroudsburg: Association for Computational Linguistics, 2017. 571–581. [doi: 10.18653/v1/P17-1053]
- [17] Meng F, Lu Z, Wang M, Li H, Jiang W, Liu Q. Encoding source language with convolutional neural network for machine translation. In: Zong C, Strube M, eds. Proc. of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th Int'l Joint Conf. on Natural Language Processing, Vol. 1. Stroudsburg: Association for Computational Linguistics, 2015. 20–30.

- [18] Eriguchi A, Hashimoto K, Tsuruoka Y. Tree-to-sequence attentional neural machine translation. In: Erk K, Smith NA, eds. Proc. of the 54th Annual Meeting of the Association for Computational Linguistics, Vol. 1. Stroudsburg: Association for Computational Linguistics, 2016. 823–833.
- [19] Goodfellow I, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, Courville A, Bengio Y. Generative adversarial nets. In: Ghahramani Z, Welling M, Cortes C, Lawrence ND, Weinberger KQ, eds. Advances in Neural Information Processing Systems 27 (NIPS 2014). New York, 2014. 2672–2680.
- [20] Mirza M, Osindero S. Conditional generative adversarial nets. CoRR, abs/1411.1784, 2014.
- [21] Radford A, Metz L, Chintala S. Unsupervised representation learning with deep convolutional generative adversarial networks. CoRR, abs/1511.06434, 2015.
- [22] Zhang Y, Gan Z, Carin L. Generating text via adversarial training. In: Lee DD, Sugiyama M, Luxburg UV, Guyon I, Garnett R, eds. Proc. of the NIPS 2016 Workshop on Adversarial Training. New York, 2016. 1–6.
- [23] Zhao JJ, Kim Y, Zhang K, M.Rush A, LeCun Y. Adversarially regularized autoencoders for generating discrete structures. CoRR, abs/1706.04223, 2017.
- [24] Wu Y, Bamman D, Russell S. Adversarial training for relation extraction. In: Palmer M, Hwa R, Riedel S, eds. Proc. of the 2017 Conf. on Empirical Methods in Natural Language Processing. Stroudsburg: Association for Computational Linguistics, 2017. 1778–1783. [doi: 10.18653/v1/D17-1187]
- [25] Liu P, Qiu X, Huang X. Adversarial multi-task learning for text classification. In: Barzilay R, Kan MY, eds. Proc. of the 55th Annual Meeting of the Association for Computational Linguistics, Vol. 1. Stroudsburg: Association for Computational Linguistics, 2017. 1–10. [doi: 10.18653/v1/P17-1001]
- [26] Pak A, Paroubek P. Twitter as a corpus for sentiment analysis and opinion mining. In: Calzolari N, Choukri K, Maegaard B, Mariani J, Odijk J, Piperidis S, Rosner M, Tapias D, eds. Proc. of the 7th Conf. on Int'l Language Resources and Evaluation (LREC 2010). Luxembourg: European Language Resources Association, 2010. 1320–1326.
- [27] Xu G, Meng X, Wang H. Build Chinese emotion lexicons using a graph-based algorithm and multiple resources. In: Huang CR, Jurafsky D, eds. Proc. of the 23rd Int'l Conf. on Computational Linguistics (Coling 2010). Beijing: Coling 2010 Organizing Committee, 2010. 1209–1217.
- [28] Yang M, Peng B, Chen Z, Zhu D, Chow, KP. A topic model for building fine-grained domain-specific emotion lexicon. In: Toutanova K, Wu H, eds. Proc. of the 52nd Annual Meeting of the Association for Computational Linguistics, Vol. 2. Stroudsburg: Association for Computational Linguistics, 2014. 421–426.
- [29] Lin KHY, Yang C, Chen HH. What emotions do news articles trigger in their readers. In: Kraaij W, Vries, AP, Clarke CLA, Fuhr N, Kando N, eds. Proc. of the 30th Annual Int'l ACM SIGIR Conf. on Research and Development in Information Retrieval. New York: Association for Computing Machinery, 2007. 733–734. [doi: 10.1145/1277741.1277882]
- [30] Aman S, Szpakowicz S. Identifying expressions of emotion in text. In: Matoušek V, Mautner P, eds. Proc. of the 10th Int'l Conf. on Text, Speech and Dialogue. Heidelberg: Springer-Verlag, 2007. 196–205. [doi: 10.1007/978-3-540-74628-7_27]
- [31] Mohammad SM, Turney PD. Emotions evoked by common words and phrases: Using mechanical turk to create an emotion lexicon. In: Inkpen D, Strapparava C, eds. Proc. of the NAACL HLT 2010 Workshop on Computational Approaches to Analysis and Generation of Emotion in Text. Stroudsburg: Association for Computational Linguistics, 2010. 26–34.
- [32] Das D, Bandyopadhyay S. Sentence level emotion tagging on blog and news corpora. *Journal of Intelligent Systems*, 2010,19(2): 145–162. [doi: 10.1515/JISYS.2010.19.2.145]
- [33] Bertero D, Siddique FB, Wu CS, Wan Y, Chan RHY, Fung P. Real-time speech emotion and sentiment recognition for interactive dialogue systems. In: Su J, Duh K, Carreras X, eds. Proc. of the 2016 Conf. on Empirical Methods in Natural Language Processing. Stroudsburg: Association for Computational Linguistics, 2016. 1042–1047.
- [34] Felbo B, Mislove A, Søgaard A, Rahwan I, Lehmann S. Using millions of EMOJI occurrences to learn any-domain representations for detecting sentiment, emotion and sarcasm. In: Palmer M, Hwa R, Riedel S, eds. Proc. of the 2017 Conf. on Empirical Methods in Natural Language Processing. Stroudsburg: Association for Computational Linguistics, 2017. 1615–1625. [doi: 10.18653/v1/D17-1169]

- [35] Yu LC, Wang J, Lai KR, Zhang X. Predicting valence-arousal ratings of words using a weighted graph method. In: Zong C, Strube M, eds. Proc. of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th Int'l Joint Conf. on Natural Language Processing, Vol. 2. Stroudsburg: Association for Computational Linguistics, 2015. 788–793.
- [36] Wang J, Yu LC, Lai KR, Zhang X. Dimensional sentiment analysis using a regional CNN-LSTM model. In: Erk K, Smith NA, eds. Proc. of the 54th Annual Meeting of the Association for Computational Linguistics, Vol. 2. Stroudsburg: Association for Computational Linguistics, 2016. 225–230.
- [37] Buechel S, Hahn U. Emotion Analysis as a regression problem-dimensional models and their implications on emotion representation and metrical evaluation. In: Kaminka GA, Fox M, Bouquet P, Hüllermeier E, Dignum V, Dignum F, Van Harmelen F, eds. Proc. of the 22nd European Conf. on Artificial Intelligence. Amsterdam: IOS Press, 2016. 1114–1122. [doi: 10.3233/978-1-61499-672-9-1114]
- [38] Preoțiuc-Pietro D, Schwartz HA, Park G, Eichstaedt JC, Kern M, Ungar L, Shulman EP. Modelling valence and arousal in facebook posts. In: Balahur A, Goot E, Vossen P, Montoyo A, eds. Proc. of the 7th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis. Stroudsburg: Association for Computational Linguistics, 2016. 9–15.
- [39] Yu LC, Lee LH, Hao S, Wang J, He Y, Hu J, Lai KR, Zhang X. Building Chinese affective resources in valence-arousal dimensions. In: Knight K, Nenkova A, Rambow O, eds. Proc. of the 2016 Conf. of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. Stroudsburg: Association for Computational Linguistics, 2016. 540–545.
- [40] Arjovsky M, Chintala S, Bottou L. Wasserstein gan. CoRR, abs/1701.07875, 2017.
- [41] Li Z, Zhang Y, Wei Y, Wu Y, Yang Q. End-to-end adversarial memory network for cross-domain sentiment classification. In: Sierra C, ed. Proc. of the Int'l Joint Conf. on Artificial Intelligence (IJCAI 2017). 2017. 2237–2243. [doi: 10.24963/ijcai.2017/311]
- [42] Peng M, Zhang Q, Jiang Y, Huang X. Cross-domain sentiment classification with target domain specific information. In: Gurevych I, Miyao Y, eds. Proc. of the 56th Annual Meeting of the Association for Computational Linguistics, Vol. 1. Stroudsburg: Association for Computational Linguistics, 2018. 2505–2513.
- [43] Mikolov T, Sutskever I, Chen K, Corrado G, Dean J. Distributed representations of words and phrases and their compositionality. In: Burges CJC, Bottou LL, Welling M, Ghahramani Z, Weinberger KQ, eds. Advances in Neural Information Processing Systems 26 (NIPS 2013). New York, 2013. 3111–3119.
- [44] Hochreiter S, Schmidhuber J. Long short-term memory. *Neural Computation*, 1997,9(8):1735–1780. [doi: 10.1162/neco.1997.9.8.1735]
- [45] Kingma DP, Ba J. ADAM: A method for stochastic optimization. CoRR, abs/1412.6980, 2014.
- [46] Graves A. Generating sequences with recurrent neural networks. CoRR, abs/1308.0850, 2013.
- [47] Glorot X, Bengio Y. Understanding the difficulty of training deep feedforward neural networks. In: Teh YW, Titterton DM, eds. Proc. of the 13th Int'l Conf. on Artificial Intelligence and Statistics. 2010. 249–256.
- [48] Smola AJ, Schölkopf B. A tutorial on support vector regression. *Statistics and Computing*, 2004,14(3):199–222. [doi: 10.1023/B:STCO.0000035301.49549.88]
- [49] Lamos V, Aletras N, Preoțiuc-Pietro D, Cohn T. Predicting and characterising user impact on Twitter. In: Wintner S, Goldwater S, Riezler S, eds. Proc. of the 14th Conf. of the European Chapter of the Association for Computational Linguistics. Stroudsburg: Association for Computational Linguistics, 2014. 405–413.
- [50] Bitvai Z, Cohn T. Non-linear text regression with a deep convolutional neural network. In: Zong C, Strube M, eds. Proc. of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th Int'l Joint Conf. on Natural Language Processing, Vol. 2. Stroudsburg: Association for Computational Linguistics, 2015. 180–185.
- [51] Zhou C, Sun C, Liu Z, Lau FCM. A C-LSTM neural network for text classification. CoRR, abs/1511.08630, 2015.

附中文参考文献:

- [2] 姚源林,王树伟,徐睿峰,等.面向微博文本的情绪标注语料库构建.中文信息学报,2014,28(5):83–91.
- [8] 黄磊,李寿山,周国栋.微博文本的情绪语料库构建.见:第 16 届汉语词汇语义学国际研讨会(CLSW 2015)会议论文集.2015. 204–212.



朱苏阳(1989-),男,江苏苏州人,博士生,主要研究领域为自然语言处理,情感分析,情绪分析.



周国栋(1967-),男,博士,教授,博士生导师,CCF 杰出会员,主要研究领域为自然语言理解,机器翻译,信息抽取,信息检索,机器学习.



李寿山(1980-),男,博士,教授,博士生导师,CCF 专业会员,主要研究领域为自然语言处理,情感计算.

www.jos.org.cn

www.jos.org.cn