和情感词应尽量分配到同一主题.评价单元提取的 P 和 R 如图 14 所示,其中,横坐标为主题个数 K,纵坐标为准确率 P 或召回率 R.

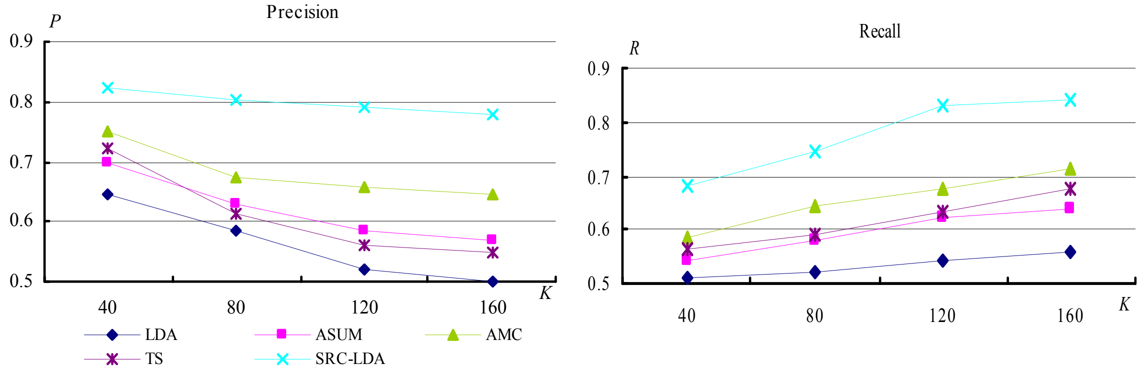


Fig.14 Precision and recall of appraisal expression extraction

图 14 评价单元提取的准确率和召回率

从图 14 可以看出:SRC-LDA 模型的准确率高于其他模型,且优势较明显,召回率的优势随着 K 的增加也更趋明显.关于准确率 P 的差异情况分析如下.

(1) 当 K=40 时,SRC-LDA 和 AMC 的 P 值相差最小,相差 7.2 个百分点;

(2) 当 K=160 时,SRC-LDA 和 AMC 的 P 值相差最大,相差 13.4 个百分点;

(3) SRC-LDA 比 AMC 的 P 值平均高出 11.7 个百分点.

原因分析:由于 LDA 倾向于发现全局特征词-情感词之间的关系,难以识别中、低频共现的局部特征词-情感词关系,所以准确率偏低;ASUM 和 TS 部分引入了主题词和情感词的关联,准确率均高于 LDA;AMC 没有从语义角度考查特征词-情感词的 must-link 约束,难以准确发现局部特征词-情感词之间的关联性;SRC-LDA 从语义角度获取特征词-情感词之间的 must-link 约束,增加了中、低频特征词-情感词之间关系发现的准确率.

关于召回率 R 的差异情况分析如下.

(1) 当 K=40 时,SRC-LDA 和 AMC 的 R 值相差最小,相差 9.4 个百分点;

(2) 当 K=120 时,SRC-LDA 和 AMC 的 R 值相差最大,相差 15.6 个百分点;

(3) SRC-LDA 比 AMC 的 R 值平均高出 12.1 个百分点.

原因分析:LDA 模型倾向于发现高频共现词语,这就导致了分配概率较高的词语在各主题下重复性较高,影响了中、低频特征词和情感词的识别;AUSM,TS 及 AMC 模型同样对应中、低频特征词和情感词具有不敏感性;对于 SRC-LDA 模型,一方面由于 cannot-link 约束提高了特征词和情感词的主题辨识度,减少了同一主题下特征词和情感词之间的错误匹配关系,另一方面,must-link 约束提高了主题下低频共现的特征词和情感词关系的识别率,使得能发现更多的局部特征词-情感词评价单元.例如,对于其他模型难以发现的一些低频共现词语关系,如(色彩,逼真)、(价格,公正)、(外观,圆润)和(颜色,饱满)等,SRC-LDA 模型利用特征词-情感词 must-link 可以识别这些评价单元.

4.4 SRC-LDA模型性能分析

为了对 SRC-LDA 模型性能进行更全面的分析,从以下两个方面考察不同约束知识对模型性能的影响:(1) 仅使用 must-link 约束,记为 M-SRC-LDA;(2) 仅使用 cannot-link 约束,记为 C-SRC-LDA.

上述两种情况和 SRC-LDA 的特征词、情感词和词语关联组合的准确率和召回率比较如图 15~图 17 所示,其中,横坐标为主题个数 K,纵坐标为准确率 P 或召回率 R.

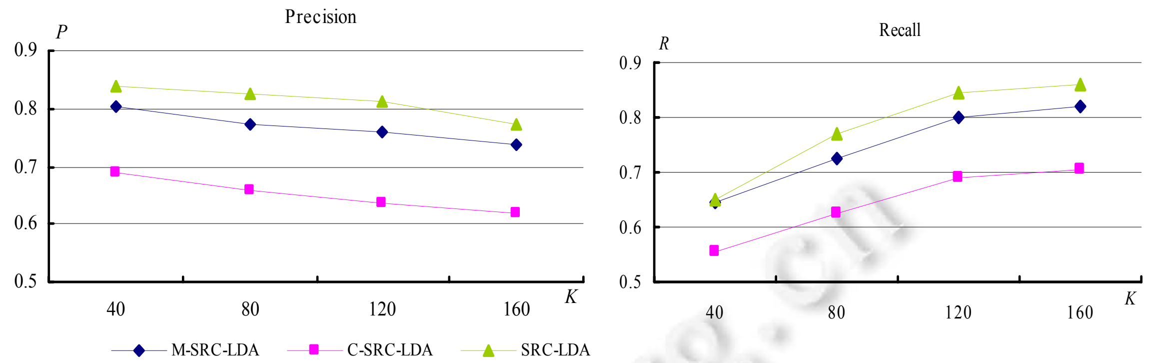


Fig.15 Precision and recall of aspect extraction

图 15 特征词提取的准确率和召回率

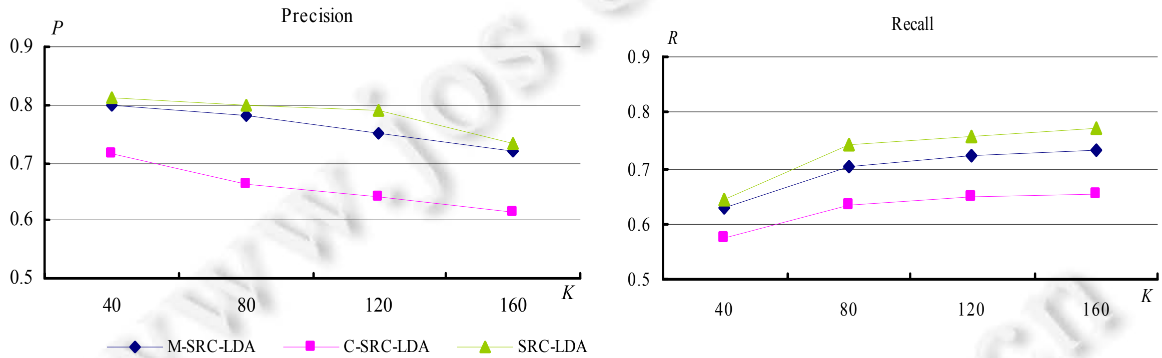


Fig.16 Precision and recall of opinion extraction

图 16 情感词提取的准确率和召回率

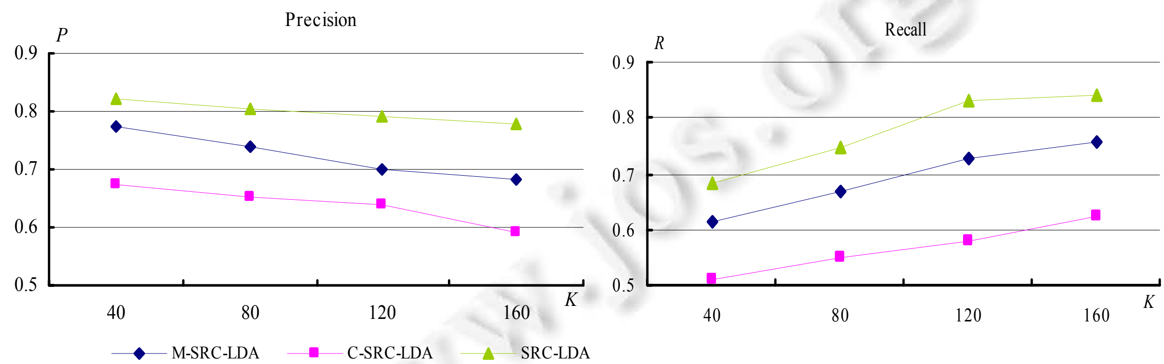


Fig.17 Precision and recall of appraisal expression extraction

图 17 评价单元提取的准确率和召回率

从图 15 可以看出:C-SRC-LDA 的准确率和召回率明显低于 M-SRC-LDA 和 SRC-LDA,M-SRC-LDA 和 SRC-LDA 之间相差不大.说明在特征词的提取中,仅使用 cannot-link 约束对于特征词的识别提升作用有限,而 must-link 约束对 LDA 模型产生了较大影响.原因在于 cannot-link 可以将不同的特征词尽量分配到不同主题,但一些低频特征词还是难以发现,而 must-link 利用同义关系可以更多地识别一些低频特征词,同时也将标准 LDA 中散布在各个主题中的全局特征词集中分配到少量主题,减少了这些高频特征词对中、低频局部特征词的分配干扰.

从图 16 可以看出:M-SRC-LDA 和 SRC-LDA 对情感词提取的准确率和召回率比较接近,因为利用特征词-情感词的 must-link 可以提高中、低频情感词的发现率,而使用 cannot-link 对情感词进行主题区分,可以增加主题对不同类情感词的识别度,但仅使用 cannot-link 不能有效提高中、低频情感词的识别率.

从图 17 可以看出:SRC-LDA 比 M-SRC-LDA,C-SRC-LDA 具有较明显的优势,说明同时使用 must-link 和 cannot-link约束可以有效提升LDA对于特征词-情感词评价单元的识别性能;C-SRC-LDA在召回率上明显低于其他模型,说明仅使用 cannot-link 约束虽然可以使不同类特征词和情感词尽量分配到不同主题,一定程度提高了特征词和情感词分配在同一主题的概率,但其关联性没有得到充分改善,还是难以发现一些低频的特征词-情感词关系;M-SRC-LDA 使用 must-link 约束可以增强局部特征词和情感词的关联性,同时可以发现更多的中、低频情感词,但缺少 cannot-link 约束容易导致不同类特征词和情感词在主题分配中的相互影响,尤其是来自全局特征词和全局情感词的干扰.

5 总结与展望

一方面,商品评论中存在很多低频的特征词和情感词,LDA 在主题分配中难以识别这类词语,使得低频特征和情感词的提取率不高;另一方面,商品评论文档中会同时出现多个不同特征的评价,LDA 很难辨别并将这些不相关特征分配到不同主题,造成特征词及其匹配的情感词没有实现较好的主题区分,难以有效提取特征和情感词.在中文商品评论中,特征词-特征词、特征词-情感词以及情感词-情感词之间蕴含着丰富的词语语义关系,可以利用这些语义知识来提高LDA对主题词语的识别度和区分度,改善标准LDA模型对一些低频特征词和情感词以及它们之间关系的提取率.

本文提出的 SRC-LDA 模型就是加入了 must-link 和 cannot-link 语义约束之后的 LDA 模型.一方面,通过 must-link 约束可以更多地发现低频的特征词和情感词,并将关联性强的特征词和情感词尽量分配到相同主题,提升中低频局部特征词和情感词的识别度,同时增加了词语间的主题聚合度;另一方面,通过 cannot-link 约束可以更多地发现评无相关的特征词和情感词,并将无关联的特征词和情感词尽量分配到不同主题,提升特征词和情感词的区分度.

实验结果表明:SRC-LDA 模型改善了特征词和情感词的主题内聚度,改进了特征词和情感词的主题区分度,从而提高了特征词、情感词和评价单元的提取率.对于特征词提取,SRC-LDA 模型比 AMC[27]的准确率平均高出 11.7 个百分点、召回率平均高出 11 个百分点;对于情感词提取,SRC-LDA 模型比 AMC 的准确率平均高出 10.1 个百分点、召回率平均高出 8.8 个百分点;对于评价单元提取,SRC-LDA 模型比 AMC 的准确率平均高出 11.7 个百分点、召回率平均高出 12.1 个百分点.

下一步工作希望继续挖掘中文商品评论中的语义知识来影响主题模型对于主题词的提取,更多发现符合语义要求的特征词和情感词.

References:

[1] Liu B. Sentiment Analysis and Opinion Mining. California: Morgan & Claypool Publishers, 2012. [doi: 10.2200/S00416ED1V01Y201204HLT016]

[2] Hu MQ, Liu B. Mining opinion features in customer reviews. In: Proc. of the 19th National Conf. on Artificial Intelligence (AAAI 2004). AAAI Press, 2004. 755−760.

[3] Popescu AM, Etzioni O. Extracting product features and opinions from reviews. In: Proc. of the Human Language Technology Conf. and the Conf. on Empirical Methods in Natural Language Processing (HLT/EMNLP 2005). Vancouver: ACL Press, 2005. 339−346. [doi: 10.1007/978-1-84628-754-1_2]

[4] Jakob N, Gurevych I. Extracting opinion targets in a single and cross-domain setting with conditional random fields. In: Proc. of the Conf. on Empirical Methods in Natural Language Processing (EMNLP 2010). Cambridge: ACL Press, 2010. 1035−1045.

[5] Jin W, Ho HH, Srihari RK. OpinionMiner: A novel machine learning system for Web opinion mining and extraction. In: Proc. of the 15th ACM Int’l Conf. on Knowledge Discovery and Data Mining (SIGKDD 2009). Paris: ACM Press, 2009. 1195−1204. [doi: 10.1007/978-1-84628-754-1_2]

[6] Su Q, Xu XY, Guo HL, Guo ZL, WuX, Zhang XX, Swen B, Su Z. Hidden sentiment association in Chinese web opinion mining. In: Proc. of the 17th Int'l Conf. on World Wide Web (WWW 2008). Beijing: ACM Press, 2008. 959−968. [doi: 10.1145/1367497.1367627]

[7] Wang RY, Ju JP, Li SS, Zhou GD. Feature engineering for CRFs based opinion target extraction. Journal of Chinese Information Processing, 2012,26(2):56−61 (in Chinese with English abstract).

[8] Liu HY, Zhao YY, Qin B, Liu T. Comment target extraction and sentiment classification. Journal of Chinese Information Processing, 2010,24(1):84−88 (in Chinese with English abstract).

[9] Wu YB, Zhang Q, Huang XJ, Wu LD. Phrase dependency parsing for opinion mining. In: Proc. of the 47th Annual Meeting of the Association for Computational Linguistics (ACL 2009). Singapore: ACL Press, 2009. 1553−1541.

[10] Zhao YY, Qin B, Che WX, Liu T. Appraisal expression recognition based on syntactic path. Ruan Jian Xue Bao/Journal of Software, 2011,22(5):887−898 (in Chinese with English abstract). http://www.jos.org.cn/1000-9825/3767.htm [doi: 10.3724.SP.J.1001.2011.03767]

[11] Qiu G, Liu B, Bu JJ, Chen C. Opinion word expansion and target extraction through double propagation. Computational Linguistics, 2011,37(1):9−27. [doi: 10.1162/coli_a_00034]

[12] Yao TF, Nie QY, Li JC, Li LL, Lou DC, Chen K, Fu Y. An opinion mining system for Chinese automobile reviews. In: Proc. of the Frontiers of Chinese Information Processing. Beijing: Tsinghua University Press, 2006. 260−281.

[13] Poria S, Cambria E, Ku LW, Gui C, Gelbukh A. A rule-based approach to aspect extraction from product reviews. In: Proc. of the 2nd Workshop on Natural Language Processing for Social Media (SocialNLP 2014). Dublin: ACM Press, 2014. 28−37. [doi: 10.3115/v1/W14-5905]

[14] Blei DM, Ng AY, Jordan MI. Latent dirichlet allocation. Journal of Machine Learning Research, 2003,3(3):993−1022.

[15] Titov I, McDonald RT. Modeling online reviews with multi-grain topic models. In: Proc. of the 17th Int'l Conf. on World Wide Web (WWW 2008). Beijing: ACM Press, 2008. 111−120. [doi: 10.1145/1367497.1367513]

[16] Andrzejewski D, Zhu X, Craven M. Incorporating domain knowledge into topic modeling via Dirichlet forest priors. In: Proc. of the 26th Annual Int'l Conf. on Machine Learning (ICML 2009). Montreal: ACM Press, 2009. 25−32. [doi: 10.1145/1553374.1553378]

[17] Zhai ZW, Liu B, Xu H, Jia PF. Constrained LDA for grouping product features in opinion mining. In: Proc. of the Pacific-Asia Conf. on Knowledge Discovery and Data Mining (PAKDD 2011). Shenzhen: Springer-Verlag Press, 2011. 448−459. [doi: 10.1007/978-3- 642-20841-6_37]

[18] Chen ZY, Mukherjee A, Liu B, Hsu M, Castellanos M, Ghosh R. Exploiting domain knowledge in aspect extraction. In: Proc. of the Conf. on Empirical Methods on Natural Language Processing (EMNLP 2013). Seattle: ACL Press, 2013. 1655−1667.

[19] Bagheri A, Saraee M, Jong FD. ADM-LDA: An aspect detection model based on topic modeling using the structure of review sentences. Journal of Information Science, 2014,40(5):621−636. [doi: 10.1177/0165551514538744]

[20] Ma BZ, Yan ZJ. Product features extraction of online reviews based on LDA model. Computer Integrated Manufacturing Systems, 2014,20(1):96−103 (in Chinese with English abstract).

[21] Chen ZY, Mukherjee A, Liu B. Aspect extraction with automated prior knowledge learning. In: Proc. of the 52nd Annual Meeting of the Association for Computational Linguistics (ACL 2014). Baltimore: ACL Press, 2014. 347−358. [doi: 10.3115/v1/P14-1033]

[22] Lin C, He Y. Joint sentiment/topic model for sentiment analysis. In: Proc. of the 18th ACM Conf. on Information and Knowledge Management (CIKM 2009). Hongkong: ACM Press, 2009. 375−384. [doi: 10.1145/1645953.1646003]

[23] Lu B, Ott M, Cardie C, Tsou BK. Multi-Aspect sentiment analysis with topic models. In: Proc. of the 11th IEEE Int'l Conf. on Data Mining (ICDM 2011). Vancouver: IEEE Press, 2011. 81−88. [doi: 10.1109/ICDMW.2011.125]

[24] Jo Y, Oh AH. Aspect and sentiment unification model for online review analysis. In: Proc. of the 4th ACM Int'l Conf. on Web Search and Data Mining (ICWS 2011). Washington: IEEE Press, 2011. 815−824. [doi: 10.1145/1935826.1935932]

[25] Moghaddam S, Ester M. ILDA: Interdependent LDA model for learning latent aspects and their ratings from online product reviews. In: Proc. of the 34th Int'l Conf. on Research and Development in Information Retrieval (SIGIR 2011). Beijing: ACM Press, 2011. 665−674. [doi: 10.1145/2009916.2010006]

[26] Sun Y, Zhou XG, Fu W. Unsupervised topic and sentiment unification model for sentiment analysis. Acta Scientiarum Naturalium Universitatis Pekinensis, 2013,49(1):102−108 (in Chinese with English abstract).

[27] Chen ZY, Liu B. Mining topics in documents: Standing on the shoulders of big data. In: Proc. of the 20th Int'l Conf on Knowledge Discovery and Data Mining (SIGKDD 2014). New York: ACM Press, 2014. 1116−1125. [doi: 10.1145/2623330.2623622]

[28] Dermouche M, Kouas L, Velcin J, Loudcher S. A joint model for topic-sentiment modeling from text. In: Proc. of the Symp. on Applied Computing (SAC 2015). Salamanca: ACM Press, 2015. 819−824. [doi: 10.1145/2623330.2623622]

[29] Ouyang JH, Liu YH, Li XM, Zhou XT. Multi-Grain sentiment/topic model based on LDA. Acta Electronica Sinica, 2015,43(9): 1875−1880 (in Chinese with English abstract).

[30] Li FT, Huang ML, Zhu XY. Sentiment analysis with global topics and local dependency. In: Proc. of the 24th Conf. on Artificial Intelligence (AAAI 2010). Atlanta: AAAI Press, 2010. 1371−1376.

[31] Zhao WX, Jiang J, Yan HF, Li XM. Jointly modeling aspects and opinions with a MaxEnt-LDA hybrid. In: Proc. of the 2010 Conf. on Empirical Methods in Natural Language Processing (EMNLP 2010). Cambridge: ACL Press, 2010. 56−65.

[32] Mukherjee A, Liu B. Aspect extraction through semi-supervised modeling. In: Proc. of the 50th Annual Meeting of the Association for Computational Linguistics (ACL 2012). Jeju Island: ACL Press, 2012. 339−348.

[33] Ma C, Wang M, Chen X. Topic and sentiment unification maximum entropy model for online review analysis. In: Proc. of the 24th Int'l Conf. on World Wide Web Companion (WWW 2015). Florence: ACM Press, 2015. 649−654. [doi: 10.1145/2740908.2741704]

[34] Li FT, Wang S, Liu SH, Zhang M. SUIT: A supervised user-item based topic model for sentiment analysis. In: Proc. of the 28th Conf. on Artificial Intelligence (AAAI 2014). Quebec: AAAI Press, 2014. 1636−1642.

[35] Heyrani-Nobari G, Chua TS. User intent identification from online discussions using a joint aspect-action topic model. In: Proc. of the 28th Conf. on Artificial Intelligence (AAAI 2014). Quebec: AAAI Press, 2014. 1221−1227.

[36] Yang ZH, Kotov A, Mohan A, Lu SY. Parametric and non-parametric user-aware sentiment topic models. In: Proc. of the 38th Int'l Conf. on Research and Development in Information Retrieval (SIGIR 2015). Santiago: ACM Press, 2015. 413−422. [doi: 10.1145/2766462.2767758]

[37] Che WX, Li ZH, Liu T. LTP: A Chinese language technology platform. In: Proc. of the 23rd Int'l Conf. on Computational Linguistics: Demonstrations (COLING 2010). Beijing: ACM Press, 2010. 3−16.

附中文参考文献:

[7] 王荣洋,鞠久朋,李寿山,周国栋.基于 CRFs 的评价对象抽取特征研究.中文信息学报,2012,26(2):56−61.

[8] 刘鸿宇,赵妍妍,秦兵,刘挺.评价对象抽取及其倾向性分析.中文信息学报,2010,24(1):84−88.

[10] 赵妍妍,秦兵,车万翔,刘挺.基于句法路径的情感评价单元识别.软件学报,2011,22(5):887−898. http://www.jos.org.cn/1000-9825/3767.htm [doi: 10.3724.SP.J.1001.2011.03767]

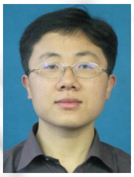
[12] 姚天昉,聂青阳,李建超,李林琳,娄德成,陈珂,付宇.一个用于汉语汽车评论的意见挖掘系统.见:中文信息处理前沿进展会议论文集.北京:清华大学出版社,2006.260−281.

[20] 马柏樟,颜志军.基于潜在狄利特雷分布模型的网络评论产品特征抽取方法.计算机集成制造系统,2014,20(1):96−103.

[26] 孙艳,周学广,付伟.基于主题情感混合模型的无监督文本情感分析.北京大学学报(自然科学版),2013,49(1):102−108.

[29] 欧阳继红,刘燕辉,李熙铭,周晓堂.基于 LDA 的多粒度主题情感混合模型.电子学报,2015,43(9):1875−1880.

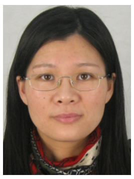
彭云(1972－),男,江西宜春人,博士生,副教授,CCF 专业会员,主要研究领域为情感分析,数据挖掘,自然语言处理.



刘德喜(1975－),男,博士,教授,博士生导师,CCF 高级会员,主要研究领域为 Web 数据管理,信息检索,自然语言处理.

万常选(1962－),男,博士,教授,博士生导师,CCF 高级会员,主要研究领域为 Web 数据管理,情感分析,信息检索,数据挖掘.

刘喜平(1981－),男,博士,副教授,CCF 专业会员,主要研究领域为信息检索,数据挖掘.



江腾蛟(1976－),女,博士,讲师,主要研究领域为情感分析,Web 数据管理.

廖国琼(1969－),男,博士,教授,博士生导师,CCF 高级会员,主要研究领域为数据库,数据挖掘,社会网络.