



# Machine learning based aspect level sentiment analysis for Amazon products

Neha Nandal<sup>1</sup> · Rohit Tanwar<sup>2</sup> · Jyoti Pruthi<sup>1</sup>

Received: 23 October 2019 / Revised: 8 February 2020 / Accepted: 13 February 2020 / Published online: 26 February 2020  
© Korean Spatial Information Society 2020

**Abstract** The field of sentiment analysis is widely utilized for analyzing the text data and then extracting the sentiment component out of that. The online commercial websites generate a huge amount of textual data via customer's reviews, comments, feedbacks and tweets every day. Aspect level analysis of this data provides a great help to retailers in better understanding of customer's expectations and then shaping their policies accordingly. However, a number of algorithms are existing these days to do aspect level sentiment detection on specified domains, but a few consider bipolar words (words which change polarity according to context) while doing analyses. In this paper, a novel approach has been presented that utilizes aspect level sentiment detection, which focuses on the features of the item. The work has been implemented and tested on Amazon customer reviews (crawled data) where aspect terms are identified first for each review. The system performs pre-processing operations like stemming, tokenization, casing, stop-word removal on the dataset to extract meaningful information and finally gives a rank for its classification in negativity or positivity.

**Keywords** Aspects · Machine learning · Support vector machines · Sentiment analysis · API crawler · Bipolar words

## 1 Introduction

Sentiment analysis is one of the techniques in natural language processing which helps in identification of sentiments that can allow entrepreneurs to get information about their customers' views through different online mediums like social media, surveys, e-commerce site reviews etc. This information can make one understand the reasons of product deterioration and the aspects which are affecting the same. The era of early 2000s was the time when sentiment analysis has aggrandized. Researchers have shown high interest in the area of sentiment analysis. Aspect level sentiment analysis came into existence as a part of sentiment analysis in which the main focus remains on particular aspects of the product/data.

Two terms named as 'Polarity' and 'Subjectivity' can be explored as parts of sentiment analysis. Subjectivity refers to the individual's beliefs, views or personal sentiments while polarity simply refers to the sentiments expressed in terms of positive, negative or neutral. Sentiment analysis covers the scope of working on sentence level, document level and sub-sentence level.

Different types of sentiment analysis can be performed on different domains i.e. one can do Fine-grained sentiment analysis by working on polarities in range from very negative to very positive, another analysis can be intent based or emotion detection and also aspect level sentiment analysis can be performed on data [1]. To perform sentiment analysis, traditional approach utilizing Lexicon based approach can be performed and another one is machine learning based approach. Both approaches have their own pros and cons. Aspect level sentiment analysis analyzes the data by concentrating on features or aspects of the data [2]. For example,

---

✉ Neha Nandal  
neha28nandal@gmail.com

<sup>1</sup> Department of Computer Science and Technology, Manav Rachna University, Faridabad, India

<sup>2</sup> Department of Systemics, University of Petroleum & Energy Studies, Dehradun, India

I like iphone for security but also there are some limits in utilizing the applications with it

In this example, aspects are *Security* and *ease to use apps* and polarity of the whole sentence cannot be predicted [3] as positive sentiments toward security of iphone and negative sentiment towards application utilization in iphone make the prediction complex about the iphone. Aspect level sentiment analysis provides a way to explore the data in detailed manner and can be highly utilizable to providers and users both to know important aspects of items which can affect their sale [4].

In 1950s, Researcher's commenced the work on semantic orientation, POS tagging which led to the origin of sentiment analysis. Early 2000s was the time of growth for sentiment analysis and during 2004, maximum number of publications came up in the area. Work on part-of-speech tagging has been presented in which an update made to the stochastic taggers [5] using a new rule. The presented novel approach showed improved performance. Along with the same, the utilization of hidden Markov models [6] also being done to increase the performance of sentiment analysis. Hidden Markov model played a very important role in the time series signal processes.

However, the first step of the research work is to collect data and moreover it should be organized in a proper format. An automated approach to collect the data can be highly useful for the work. A system named OPINE [7], an unsupervised approach has been prescribed for extraction of reviews and their respective polarities. Although the extracted information/data should be in an understandable format [8]. Context information in the customers reviews have been identified in the work. After extraction of data, preprocessing of the data plays a vital role in the task. Aspect level sentiment analysis brought a new sight to the sentiment analysis. One of the great contributions were the SemEval task [9] which caught attention of many researchers. The important and key points of the Aspect Level sentiment analysis field have been discussed. Aspect level sentiment analysis comprises of two parts mainly i.e. aspect classification and extraction [10].

Researchers [11] showed their study on techniques of sentiment analysis on tweets. They explained about different approaches of sentiment analysis like document level, text level approach etc. In [12], researchers presented a work on customer Ad sharing sentiments. They came up with a concept that sentiment analysis provides better understanding of intentions of customer for sharing Ads online.

Researchers [13] focused their research work on study of emotions and sentiment analysis. They introduced natural language processing, then model of emotions and other approaches. The papers they have taken in consideration

for survey are either from DH or from computational linguistic venue.

## 2 Methodology and work flow

A model has been designed for better sentiment analysis using ensemble approach to improve correctness and efficiency which is implemented on reviews collected for trending keywords. Overall flow of the work is shown in Fig. 1.

Firstly, the data stream has been collected using the API and information has been extracted. Next, the extraction of aspects has been done, which are properties of a product. Next step is to map the sentiments to ratings.

Each and every step of work methodology have its importance. The methodology flow of the work has been discussed in this section.



Fig. 1 Work flow and methodology

## 2.1 Data collection

For any sentiment analysis to work we need data first, we developed a Scrapy based web crawler [14] to fetch user reviews on given products from Amazon. Basic flow of Crawler is shown in Fig. 2.

For data collection, an application processing interface has been developed to store the data. Scrapy has been utilized for data extraction; it's processing and saving data in .csv format.

The data has been stored in tabular format with attributes as date of the review, URL of the review, rating on the review, user name and user review. The number of product reviews collected is shown in Fig. 3.

## 2.2 Aspect identification

Identification of aspects specifies identification of words or phrases which relates to the features of the review comments. For example, say product is earphones; the important aspects of an earphone are shown in Fig. 4.

Human identifier plays an important role in identifying aspect terms to manually store the aspects and their sentiments. Aspects can be identified by following aspect aggregation i.e. the terms that are synonyms of each other (for example, 'battery' and 'charging') which can be done by utilizing supervised approach.

## 2.3 Preprocessing of data

This phase is one of the most important phases in which cleaning of data and removal of stop words etc. happens to improve the effectiveness of results.

- *Vectorization* phase combinable provides a record of data that will be required for classification of reviews and a technique of vector space model is utilized for the same.
- *POS tagging* is part of speech tagging which permits to tag each word of data to the POS i.e. verb, adverb, noun, pronoun, adjective etc.
- *Stemming and lemmatization* helps to reduce spatiality in the words. For example, the words like 'bright', 'brighter', and 'brightening' are taken as one word 'bright'.
- *Stop word removal* works on removing those words from data which do not affect the final sentiment value of the data.

## 2.4 Evaluation and classification

Support vector machines has been used as a classifier which is a supervised learning having concept of hyper plane to deal with complex problems.

Table 1 presents some sample *bipolar* words which shows the change in polarities of words in presence of some context.

It is important to understand nature of words for identification of bipolar words. The words when used with context changes its original polarity. For example, "Dark Glossy", in which the word "dark" owns negative polarity but when it is utilized with "Glossy", the polarity of the whole phrase changed as Positive. The example is elaborated in Table 2.

After identification of bipolar words and its adjustment, the next part is the classification and evaluation. Support vector machine has been used as a classifier. Among three

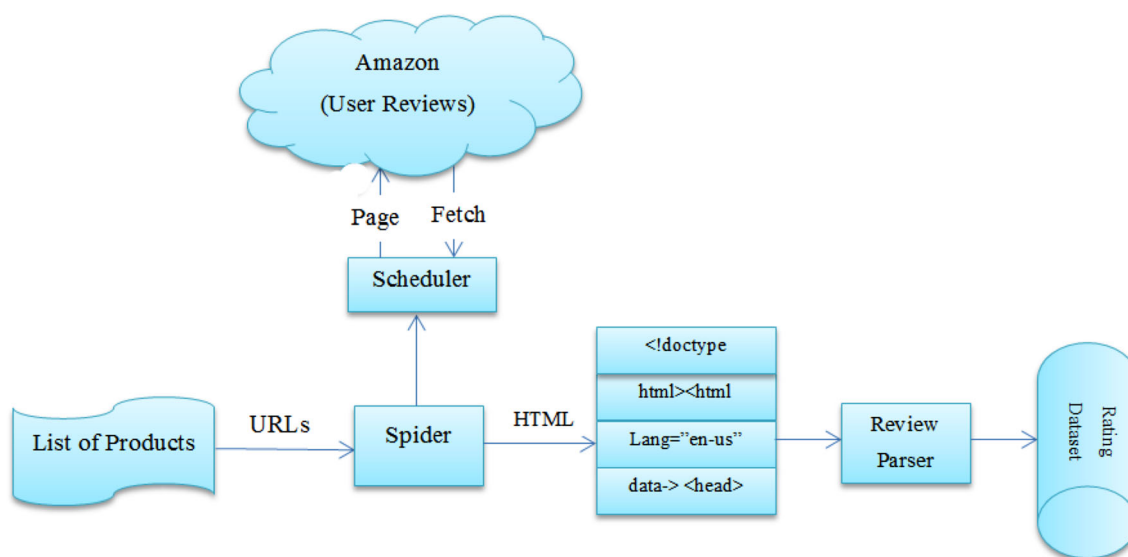
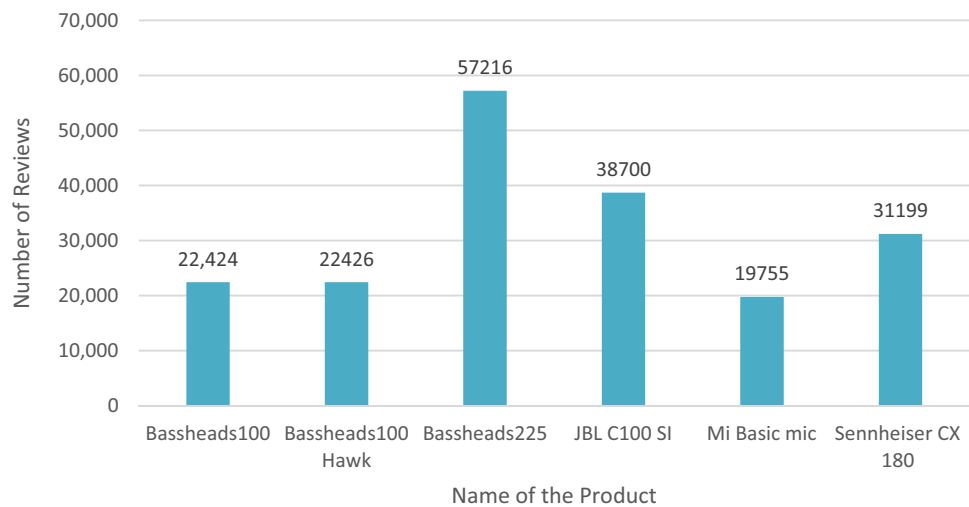
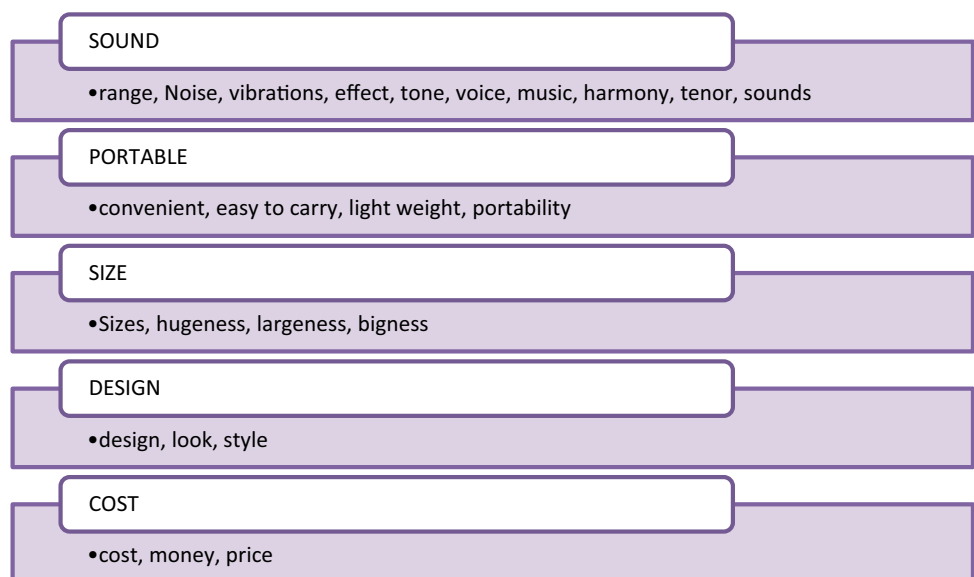


Fig. 2 Flow of crawler

**Fig. 3** Showing no. of product reviews of crawled Amazon products which were collected in the dataset



**Fig. 4** Sample aspects of a product



kernels of SVM, RBF kernel provided best results. The working of SVM classifier is shown in Fig. 5.

### 3 Results

The work has been implemented on Platform of Python and Matlab has been utilized for analysis purpose. The Python was selected to do more of sentiment analysis using support vector machines. The support vector machines with its three kernels i.e. linear, polynomial and radial basis function (RBF) has been used for analysis. With Matlab, graphs have been plotted for analysis of confusion matrix, learning rate, area under the curve.

For the evaluation of the proposed work we have used various metrics such as learning rate, mean squared error (MSE), accuracy, precision, recall, confusion matrix and

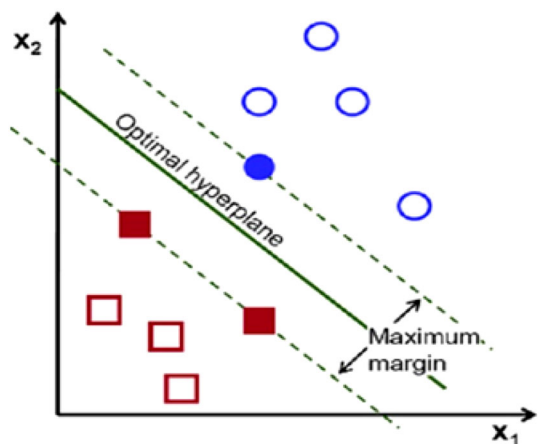
**Table 1** Bi-polar words with context changing polarity (sample)

Word	Context	Polarity shift	Adjustment
Belated	Birthday	+	0.81
Vice	Versa	+	0.79
Stress	Reliever	+	0.48
Deep	Breath	+	0.83
Long	Awaited	+	0.44
Dark	Glossy	+	0.87
Dark	Body	+	0.43
Well	Oh	-	0.45
Hot	Outside	-	0.86
Better	Feels	-	0.72
Enough	Money	-	0.37
More	Noise	-	0.85
Little	Old	-	0.70
Super	Duper	+	0.52

**Table 2** Bi-polar word adjustment with context changing polarity (example)

All these phones are superb, no doubt, but once you use the iPhone 7, you know that the difference between these phones and iPhone 6s has got much bigger because of its *dark Glossy* body and camera

Bi polar word	Context	Polarity shift	Original rating	Adjustment	Final rating
Dark	glossy	Negative to positive	4.0	+ 0.87	3.87



**Fig. 5** Support vector machine

roc curves. The learning rate in learning can best used to determines to what extent newly acquired information overrides old information i.e. during testing the how much of learned information is sustained in the testing phase.

The MSE which measures the quality the proposed ALSA algorithm is calculated using summing error

achieved from all the product reviews divided by the total reviews using following formula [2]:

$$MSE = \sum_{i=1}^n Actual\ rating - Mapped\ rating$$

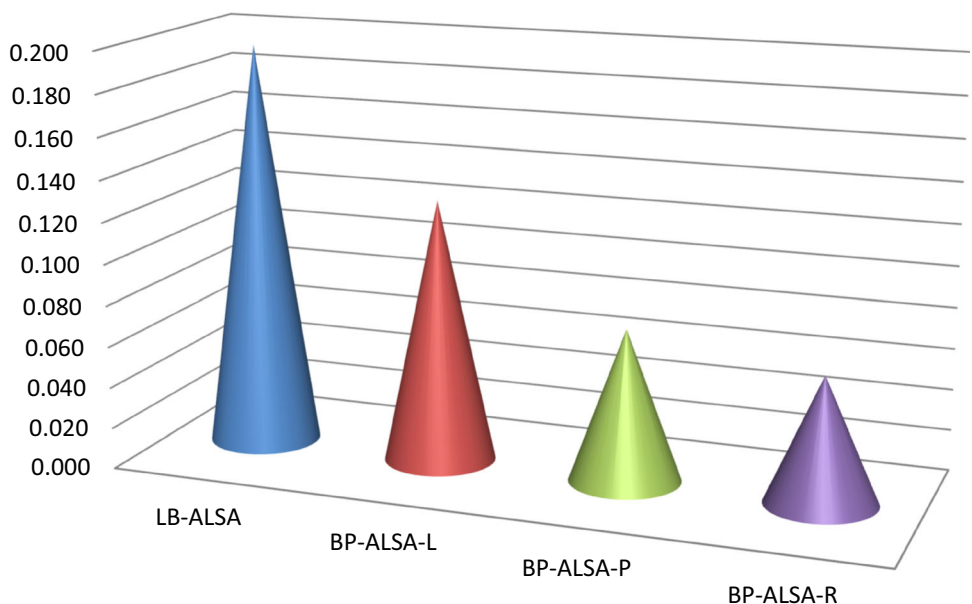
The perceived accuracy of the four classifiers LB-ALSA, BP-ALSA-L, BP-ALSA-P, BP-ALSA-R. BP-ALSA with RBF kernel being the best of all is shown in Fig. 6.

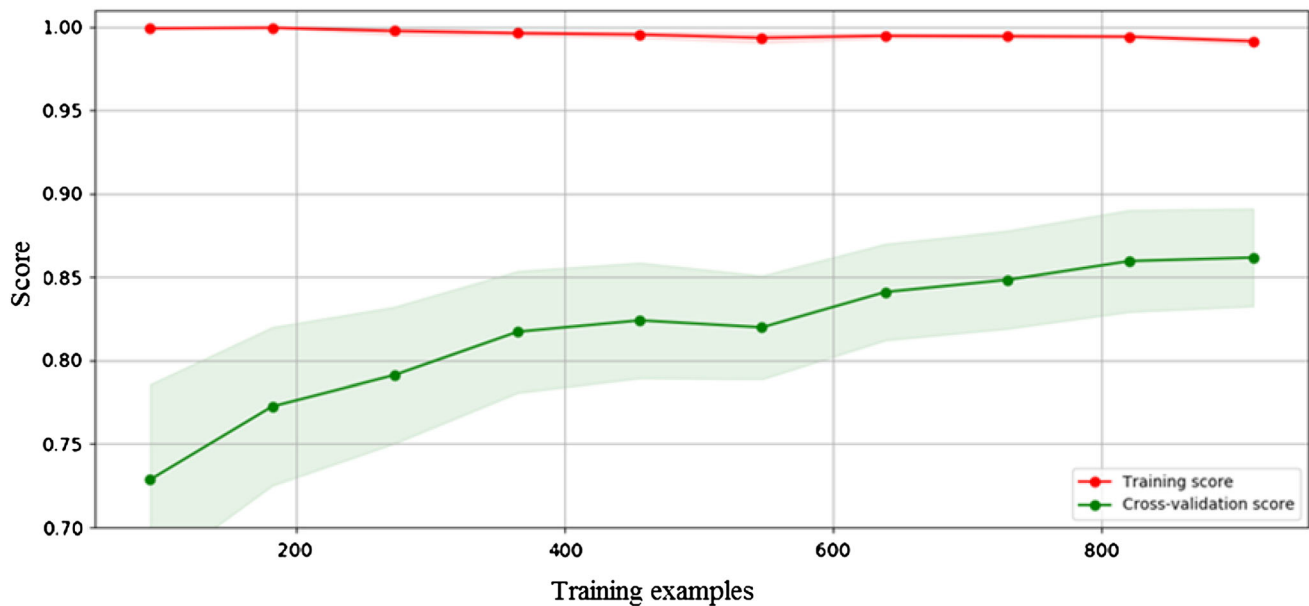
Also as just accuracy and MSE can be inefficient in analysis of a classifier [x] we must also consider other very important classification metrics such as sensitivity (true positive rate or recall) and specificity

$$Recall\ rate = \frac{True\ positive\ sentiments}{Total\ positive\ sentiments}$$

Where, true positive sentiments correspond to the number of the true positive reviews detected and total positive sentiments is the total number of positive samples including true positives and false negatives. Specificity or true negative rate (TNR) measures how well the SVM was able to recognize negative samples. It is defined as

**Fig. 6** MSE with exponential trend line showing step by step reduction in error





**Fig. 7** Learning curve of bipolar identified aspect level sentiment analysis

$$\text{Specificity} = \frac{\text{True negative sentiments}}{\text{Total negative sentiments}}$$

Where, true negative sentiments correspond to the number of the true negative sentiments detected and total negative sentiments the total number of samples that are negative in the dataset. Figure 7 shows the Learning rate of BP-ALSA algorithm during testing or cross validation phase achieving up to 97% score meaning that about only approximately 3% of information loss is found during the testing.

## 4 Conclusion

Sentiment analysis is one of the latest challenging research area to work with. The work presented here handles one of the biggest challenges of bipolar words in sentiment analysis. It is needed for the firms to use latest tools and approaches to optimize aspect level sentiment analysis. The identification of the words changing polarity in presence of context and its effect on the overall rating of the product along with the particular aspect has been analyzed on the work and results collected were impressive. It is highly required now-a-days to have a tool or software which can help both customer and developers to understand the behavior of product in market. Future work will be done with a vision to solve more challenging areas like spam and fake, negations, sarcasm etc. with latest tools.

**Funding** No funding is received from any sources for doing this work.

## Compliance with ethical standards

**Conflict of interest** The authors declare that they have no conflict of interest.

## References

1. Haque, T. U., Saber, N. N., & Shah, F. M. (2018). Sentiment analysis on large scale Amazon product reviews. In *IEEE international conference on innovative research and development (ICIRD)*. 11–12 May, Bangkok, Thailand.
2. Bertero, D., Siddique, F. B., Wu, C. S., Wan, Y., Chan, R. H. Y., & Fung, P. (2016). Real-time speech emotion and sentiment recognition for interactive dialogue systems. In *Proceedings of the 2016 conference on empirical methods in natural language processing* (pp. 1042–1047). November 1–5, Austin, TX.
3. Garanayak, M., Mohanty, S. N., Jagadev, A. K., & Sahoo, S. (2019). Recommended system using item based collaborative filtering (CF) and K-means. *International Journal of Knowledge-Based and Intelligent Engineering Systems*, 23(2), 93–101.
4. Safeek, I., & Kalideen, M. R. (2017). Preprocessing on Facebook data for sentiment analysis. In *Proceedings of 7th international symposium, SEUSL* (pp. 69–78). 7th & 8th December.
5. Brill, E. (1994). Some advances in transformation-based part of speech tagging. In *Proceedings of the twelfth national conference on artificial intelligence* (pp. 722–727). Menlo Park, CA: AAAI Press.
6. Ghahramani, Z. (2001). An introduction to hidden Markov models and Bayesian networks. *International Journal of Pattern Recognition and Artificial Intelligence*, 15(1), 9–42.
7. Popescu, A.-M., & Etzioni, O. (2007). Extracting product features and opinions from reviews. *Natural language processing and text mining* (pp. 9–28). London: Springer.
8. Liu, Z., Yang, N., & Cao, S. (2016). Sentiment-analysis of review text for micro-video. In *2nd IEEE international conference on computer and communications (ICCC)*. 14–17 Oct, Chengdu, China.

9. Pontiki, M., Galanis, D., Pavlopoulos, J., Papageorgiou, H., Androutsopoulos, I., & Manandhar, S. (2014). SemEval-2014 task 4: Aspect based sentiment analysis. In *Proceedings of the 8th international workshop on semantic evaluation* (pp. 27–35). Dublin.
10. Liu, B., Hu, M., & Cheng, J. (2005). Opinion observer: Analyzing and comparing opinions on the web. In *Proceedings of international conference on world wide web*.
11. Alsaeedi, A., & Khan, M. Z. (2019). A study on sentiment analysis techniques of Twitter data. *International Journal of Advanced Computer Science and Applications*, 10(2), 361–374.
12. Kulkarni, K. K., Kalro, A. D., Sharma, D., & Sharma, P. (2020). A typology of viral ad sharers using sentiment analysis. *Journal of Retailing and Customer Services*. <https://doi.org/10.1016/j.jretconser.2019.01.008>.
13. Kim, E., & Klinger, R. (2018). *A survey on sentiment and emotion analysis for computational literary studies*. [arXiv:1808.03137](https://arxiv.org/abs/1808.03137), 9 Aug.
14. Mitchell, R. (2018). *Web scraping with python: collecting more data from the modern web*. Newton: O'Reilly Media Inc.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.