Categorize the Opinions of Individuals using the Classifier Combination based on the Feature Selection

Leila Delikhoon

Department of Computer Engineering, Pooyesh highereducation institute,Ghom, Iran.

Abstract
Today, data mining and information analysis attract attention of many researchers. Also , cultural, social, social, and commercial motivations increase the importance of the data mining. Since the user', behavior is very decisive, analysis the user', opinion provides very useful information to managers. In this study, a new method is introduced to categorize the views and opinions of the users by using the classifier combination and conduct voting . This method takes more time for learning, but increases the classification precision . Since, the number of features of the data set is very high, in order to compensate the speed drop, a method is selected for feature selection . This method once is performed for the trial data , and then classification is performed just by the extracted features that compensates the speed reduction of the classifier combination. To test the suggested method, three data sets of the users' view from the Amazon website called Book, DVD, and Kitchen, have been used. The results of the classification were evaluated by the total accuracy, precision , and recall criteria. These criteria show that, in addition to reduce the feature, algorithm of feature selection, increases the classification efficiency that it is due to remove the noise and inappropriate features for classification. Also , results show that classifier combination is more effective than the classification by one base algorithm.

Keywords:
data mining, classification, classifier combination, feature selection

1. Introduction

Opinion mining (OM) is a modern and the newfound research field that deals with information recovery and knowledge extraction from the text by using the data mining (DM) and natural language processing (NLP) . The goal of OM is to enable the computer to recognize and state the emotions. OM is known as emotion analysis. The goal of data mining is to determine the emotional level of a document , namely that document is categorized as negative or positive. Suggested approach for data mining in the data set includes pre-processing of data, extracting the features, reducing the data (feature selection or selecting the superior pattern) and learning the classifications [1].
Wang et al (2014) presented a method to help the combined learning to classify the emotions. In this method , all learners are combined by Bagging, Boosting, and subspace methods and the results are evaluated for 10 databases by using the 10-fold cross-validator [2]. Ganjemi et al (2014) focus on the importance of preserving the opinions' emotion as one of the main challenges of the data mining [3]. Xia et al (2013) studied the issue of the evidence-based polarity change by using a set of relative combined vocal features with a sample selector [4]. Gindl et al (2010) separated the emotion terms and obscure emotions from those emotions that having constant polarity and positioning [5]. To analyze the language and extract the feature vector from the text, some tools are presented that the most famous of them have been used in the following. Tsai et al (2013) presented another method that enriches the language resources [6]. Pouria et al (2013) combined the SenticNet and WordNetAffect tools to obtain the emotional labels [7].
In this study, a method is presented to reduce the features and a method to combine the classifiers in order to improve the classification precision. First , the type of suggested feature selection algorithm is introduced and then the algorithm is described to combine the classifiers. In addition to the feature reduction in order to increase the classification speed, the suggested method to reduce the feature, increases the classification quality because it eliminates the useless features that reduced the classification precision. Although ,the suggested method for classifier combination increases the precision , but takes more time for learning . This speed drop is compensated by the feature reduction.
To select the feature , we used the Relief method. There are various algorithms for classification. The framework of all algorithms is same. Each classifier has one learning stage and one test stage. In this study, we used the three common base classifiers : k- nearest neighbor (KNN), support vector machine(SVM), and Naive Bayes (NB). There are many methods for classifier combination. The goal of all is to improve the classification performance. In this study, a method based on the random selection of the subset has been used. To evaluate the performance of the presented method, we used the 10-fold cross-validation algorithm. Values that were used to evaluate the performance and calculate the precision, are: Total accuracy: the ratio of true classified opinions to all opinions and calculated by the following relation:

Manuscript received August 5, 2026
Manuscript revised August 20, 2026
https://doi.org/10.22937/IJCSNS.2026.26.8.15

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN} = \frac{N_T}{N} \quad (1)$$

TP: the number of positive opinions that were recognized true.
TN: the number of negative opinions that were recognized true.
FP: the number of positive opinions that were recognized as negative, falsely.
FN: the number of negative opinions that were recognized as true, falsely.

Precision
Precision is the ratio of the true recognized positive sample to all available positive samples.

$$Precision = \frac{TP}{TP + FP} \quad (2)$$

Recall :
Recall is the ratio of the true recognized positive sample to all samples that are recognized positive. It is noticed that some samples that were recognized positive, are false and placed in the FN category.

$$\frac{TP}{TP + FN} \quad (3)$$

The general goal of this study is to present a combined method to increase the precision of the individual opinion combination, so that the calculation time and program complexity don't increase , as much as possible. This method is useful for classifying the information in all fields. Finally, the specialized goal of this study is to classify the individual opinions, so that makes the simplicity and high precision. We can obtain useful information about user's favor in a cultural, politic, social issue or a commercial product, from the classified opinions.

2. Methodology

The suggested method was tested on three database of kitchen ،book ،DVD from the users' opinions, which was extracted from the Amazon website [8]. Each of these databases have 2000 opinions of the users about a product. Each opinion includes many features (nearly 4000 features) that each feature takes 0 or 1 value. These features of vector are according to the user's text that has been obtained by using the VLP methods and vocabulary concept. The label 1 or 2 is allocated to each label that indicate the negativity or positivity of the user's opinion about the product. Summery , we collected 2000 samples. Each sample has 4000 features and labels. Tests were performed in three stages for three base algorithms, namely SVM, KNN, NB, and the results were compared. In the first stage, samples were classified merely by base algorithms , without any algorithm of reducing the feature or combination. In the second stage, the feature reduction method of Relief was performed before the classification and samples with the selected features are classified by the base algorithm of NB, KNN, SVM. In the third stage, both feature reduction and classifier combination are used. Finally, the results are compared in terms of total accuracy, precision, and recall.

3. Approaches

Results were divided into three parts of accuracy, precision, and recall. In each section , the results for three data sets are presented. Classification results with base classifiers. In this section, each of datasets is performed by NB, KNN, and SVM classifiers without feature reduction and without ensemble to obtain a general view about the base algorithms. Classification results of this section are observed in the table 1 and figures 1-3.

Table 1- results of base classification

Data set DVD			**Data set Book**			**Kitchen data set**			Classification
Recall	Precision	Accuracy	Recall	Precision	Accuracy	Recall	Precision	Accuracy	type
79.10	69.23	71.25	71.10	69.83	70.10	76.9	77.35	77.15	**SVM**
65.10	57.11	58.10	66.80	55.74	56.50	76.4	56.41	58.45	**KNN**
86.40	73.46	77.45	70.50	79.18	75.85	85.2	79.94	81.9	**NB**

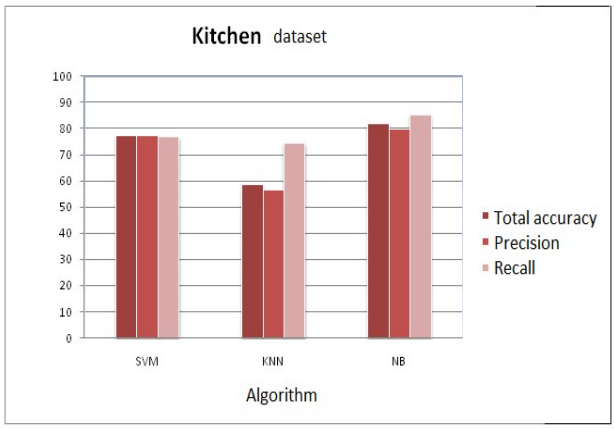

Figure 1- results of base classifiers for Kitchen dataset

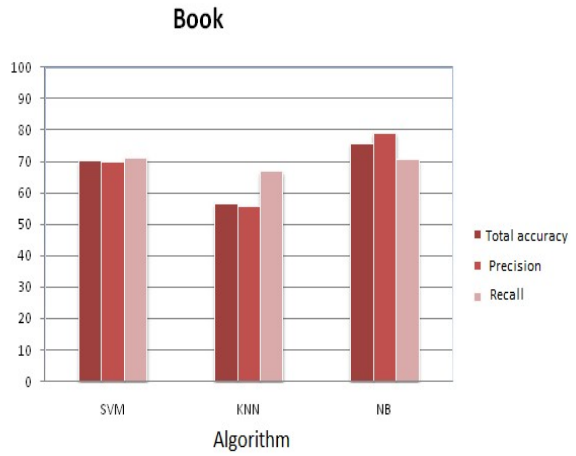

Figure 2- results of base classifiers for Book dataset

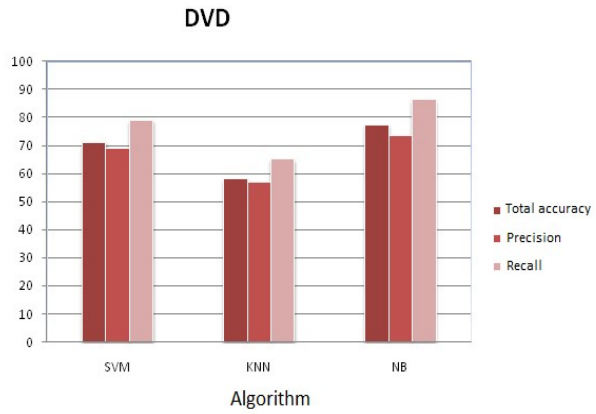

Figure 3- results of base classifiers for DVD dataset

In the figures 1=3, we observe that for three datasets, NB classifier has more precision than other two algorithms and KNN algorithm is weak in all three datasets. The reason is the large size of the samples, because the nearest neighbor in KNN is determined with regard to distance of samples. Since the values of dataset features are 0 or 1, the possibility of Similarity between the distances of samples is very high and there may be many neighbors with equal distances. Then , KNN can't present the true precision. In fact , KNN algorithm is better for samples that possess the fuzzy features and real numbers.

Results of feature reduction with Relief method

In this section , Relief algorithm with threshold value 5 is performed for all of the datasets and the classification is performed by the selected features. The results are presented in table 2.

Table 2- results of feature reduction

dataset DVD			dataset Book			Kitchen dataset			Classification algorithm
Recall	Precision	Accuracy	Recall	Precision	Accuracy	Recall	Precision	Accuracy	
73.7	71.61	72.15	72.60	71.02	71.35	78.00	78.24	78.15	SVM
61.00	58.56	58.9	67.2	55.69	56.60	73.20	56.96	58.90	KNN
86.30	72.55	76.65	71.90	79.78	76.65	85.8	79.99	82.15	NB

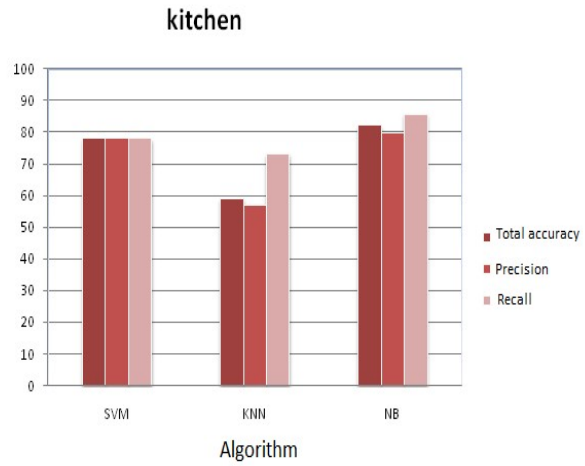

Figure 4- results of feature reduction in Kitchen dataset

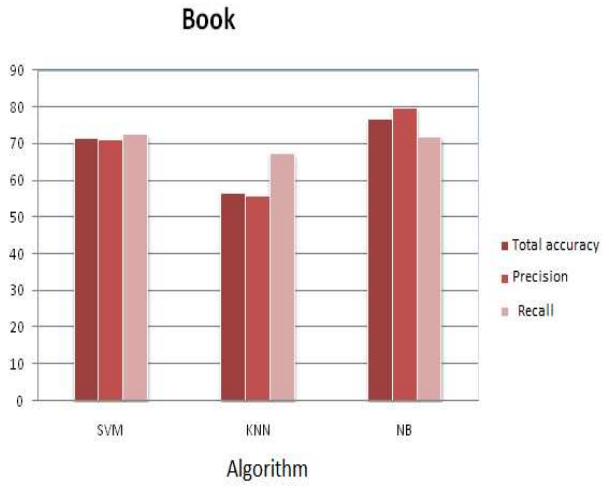

Figure 5- results of feature reduction in Book dataset

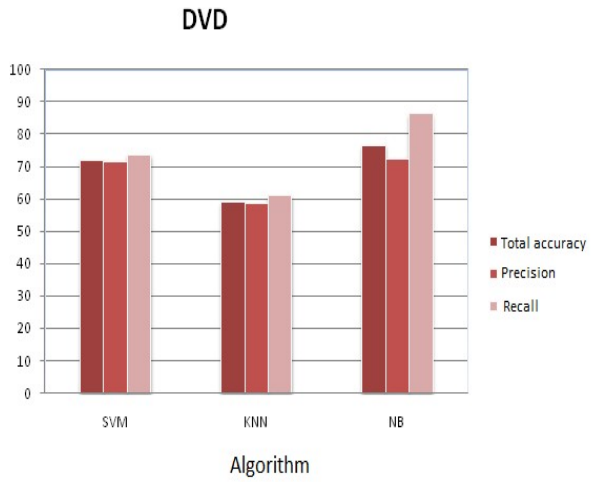

Figure 6- results of feature reduction in DVD dataset

Classifier combination with the feature reduction

In this section , after the feature reduction in the previous section, the ensemble method is performed by using the majority voting among the selected subsets. Results are summarized in table 3.

Table 3- results of feature reduction and ensemble

dataset DVD			dataset Book			Kitchen dataset			Classification algorithm
Recall	Precision	Accuracy	Recall	Precision	Accuracy	Recall	Precision	Accuracy	
99	94.28	96.51	99	87.61	92.53	96	89.71	92.53	SVM
99	79.20	86.56	97	88.18	92.03	99	87.61	92.53	KNN
90	86.53	88.05	89	88.11	88.55	93	85.32	88.55	NB

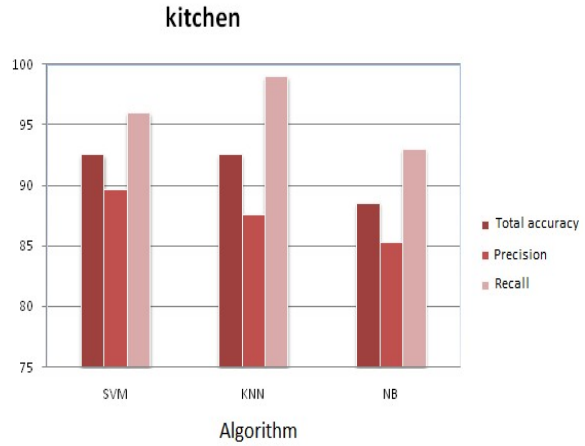

Figure 7- results of feature reduction and ensemble in Kitchen dataset

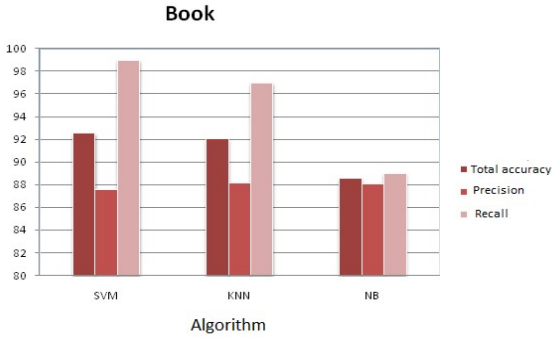

Figure 8- results of feature reduction and ensemble in Book dataset

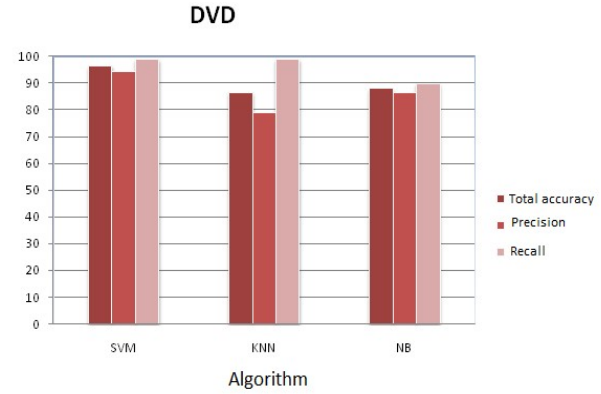

Figure 9- results of feature reduction and ensemble in DVD dataset

Above figures show that classifier combination can compensate the defects of a base classifier algorithm such as KNN, so that sometimes KNN , which was the weakest classifier in the main mode, shows a more quality with ensemble than NB and SVM. We can conclude that NB presents a high ability in the main mode and achieve the maximum precision, so that the Ensemble or feature reduction method does not affect its boosting, so much. In contrast, ensemble method has a significant positive effect on the base classifiers.

4. Result Comparison

In this section, we compare three stages, namely classification by base classifier, classification by feature reduction, and classification by using the feature reduction and ensemble , in terms of total accuracy, precision, and recall for all three databases. Following figure shows the total accuracy for all cases.

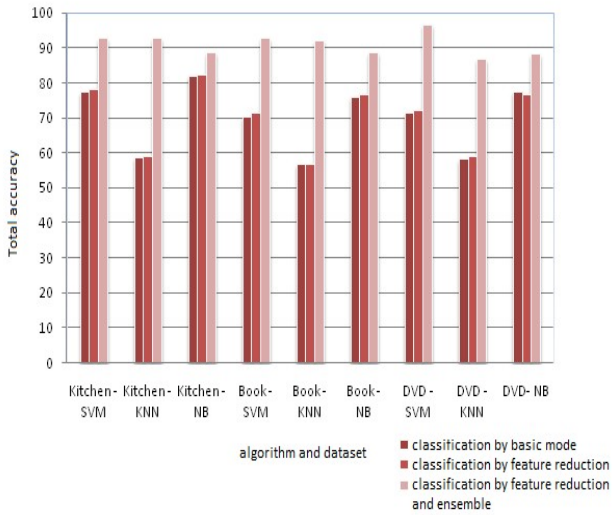

Figure 10- total accuracy comparison in methods

As seen in the figure 10, using the ensemble increases the total accuracy of classification in all algorithms and data sets, significantly. The reason is the increase in the learning ability of the combination. Except in the DVD dataset and NB algorithm, (namely, DVD-NB)in which the total accuracy is reduced after feature reduction, in all cases , the total accuracy increases after the feature reduction, but this enhancement is not significant. Since Relief algorithm does not perform any classification in its process, doesn't understand the classification and adds or removes the features merely based on the distance they make. Then , it cannot be expected that Relief algorithm has a significant effect on increasing the total accuracy. Even , it may be in some cases such as DVD-NB, the total accuracy is reduced. Then , Relief is merely an algorithm for feature reduction, so that the classification precision does not

reduce and it cannot be categorized in a feature reduction algorithm to increase accuracy. Another important point in the figure is that the high improvement of KNN performance after ensemble. Due to the binary type of the selected dataset of KNN , produces many similar neighbors that causes the accuracy drop. But , this method is performed by ensemble, the number of neighbors increases and the voting performs with high accuracy. In the base state , KNN performs classification by K neighbors, but in the ensemble method in which samples are divided into 10 subsets, KNN performs classification by 10K neighbors. Then , total accuracy increases, significantly. Therefore , we said that the suggested ensemble method is very effective in KNN algorithm than other algorithms.

For the final comparison, we can say that in the Kitchen dataset, ensemble method has the most total accuracy on KNN. In DVD dataset, ensemble method has the most total accuracy on SVM. In Book dataset, ensemble method has the most total accuracy on SVM. While in the base state , NB classifier has the most total accuracy than SVM and KNN, but in the ensemble state, its final total accuracy is less than SVM and KNN. We can conclude that although, ensemble method increases the total accuracy of NB, but can't obtain the best solution because NB use the maximum precision for classification in the base state. Following figures show the recall and precision for all trial states.

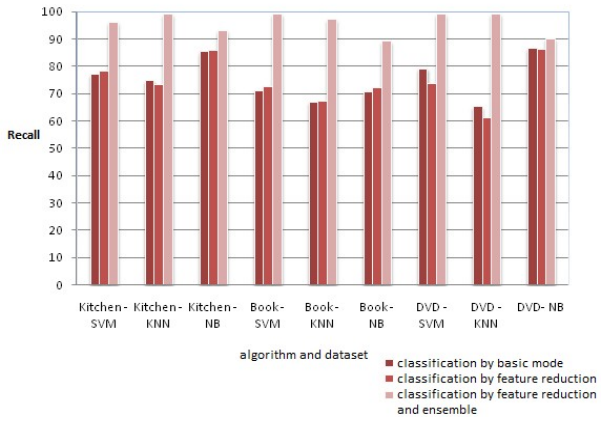

Figure 11- comparison of method recall

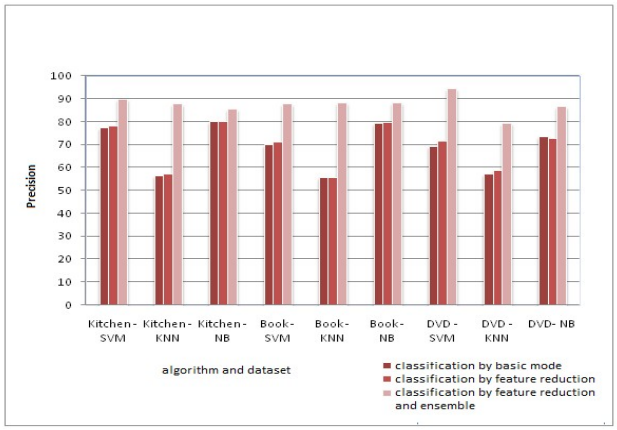

Figure 12- comparison of method precision

In both figures, we can observe the improvement after ensemble and feature reduction , clearly. These figures confirm all results about the total accuracy. Up now, the suggested methods were tested on the three datasets called Kitchen, Book, DVD, which extracted from the opinions of users in Amazon web site. In the opinion classification, it observed that many of the variables and features of the opinions can be ignored because after Relief performing and sample feature reduction, the classification quality was not reduced and even increased in all evaluations of total accuracy, precision, and recall. Only, in some cases in the second section of simulation, namely classifier combination as performing the selection among classifiers, the classification quality increased , significantly. These results show that the main goal of the study, namely increasing the classification precision, was obtained and also the combined speed reduction was compensated by decreasing the sample features by Relief method. Moreover , we observed that combined learning method or ensemble method is very effective on KNN, so that despite this algorithm was weaker than other algorithms in the base state, but presented a high precision in the combined state.

5. Conclusion

In the present study, a method for feature reduction and a method for classifier combination were introduced and the efficiency of both methods was presented. In addition to perform the feature reduction properly in the datasets of users' opinion and reduced 4000 features to 500 features, Relief method also increased the classification quality. With regard to the results, Relief algorithm has the following advantages:

Despite the wrapper method, the Relief algorithm does not require to study all possible subsets. In the Wrapper method, there are 2^N subsets for N features that all of them should be investigated.

In the tested samples, about 12% of features remained after performing the Relief method and other features were discarded . Then, this method performs the feature reduction properly. This feature reduction increases the processing speed in the next stages.

Since features are weighted based on the nearest defeat and failure and finally the low – weighted features are discarded, it is not necessary to use any classification algorithm in the feature reduction process. This characteristic shows the generality of relief method because is not limited to an especial classifier. Moreover , performing the classification for each selected subset is time- consuming. Therefore , Relief method works much faster than methods based on Wrapper.

In addition to the feature reduction and the processing speed increase, By simultaneous performing the classification quality will be increased and the efficiency will be stable. The obtained results confirm these claims.

Although , Relief method has a less calculation than other algorithms, but has a longer running time. The calculations can be reduced by using the post-innovative methods or using the artificial intelligence algorithms. In the present study, we use the three SVM, KNN, and NB classifiers for voting. For the future study, other classifiers can be used . Another idea is to use the combined method such as Bagging and Boosting in weighted voting.this study focuses more on the opinion classification, but the suggested method is useful in all classification fields. For example, the suggested method can use for classifying the medical databases in order to classify the patients, predict the disease intensity, or diagnose the disease.

References

[1] Subrahmanian, Venkatramana S., and Diego Reforgiato. "AVA: Adjective-verb-adverb

[2] Liu, Cheng-Lin. "Classifier combination based on confidence transformation." *Pattern Recognition* 38.1 (2005): 11-28.

[3] Gangemi, Aldo, Valentina Presutti, and Diego Reforgiato Recupero. "Frame-based detection of opinion holders and topics: a model and a tool." *IEEE Computational Intelligence Magazine* 9.1 (2014): 20-30.

[4] Xia, Rui, et al. "Feature ensemble plus sample selection: domain adaptation for sentiment classification." *IEEE Intelligent Systems* 28.3 (2013): 10-18.

[5] Gindl, Stefan, Albert Weichselbraun, and Arno Scharl. "Cross-domain contextualisation of sentiment lexicons." (2010).

[6] Tsai, Angela Charng-Rurng, et al. "Building a concept-level sentiment dictionary based on commonsense knowledge." *IEEE Intelligent Systems* 28.2 (2013): 22-30.

[7] Poria, Soujanya, et al. "Enhanced SenticNet with affective labels for concept-based opinion mining." *IEEE Intelligent Systems* 28.2 (2013): 31-38.

[8] McAuley, Julian, Rahul Pandey, and Jure Leskovec. "Inferring networks of substitutable and complementary products." *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM, 2015.