Enhancing Sentiment Classification Performance Using Bi-Tagged Phrases

SENTIRE 2013

What is Sentiment Analysis?

- Sentiment analysis is to extract the opinion of the user from of the text documents.

- Identifying the orientation of opinions from the text.
  - This movie was awesome . [Sentiment] ☺
  - This movie was boring. [Sentiment] ☹

Applications !

- Helpful for Business in improving quality of the product based on users opinion.
- Help people in decision making.
- For government, know the opinion of people for a certain policy.

- For example:

Which model of a camera is liked by most of the users and which music is liked the most by people ? which laptop is best to purchase? etc.

Sentiment classification using machine learning

- The proposed approach consists of two phases.
- In the first phase, various features are extracted and feature selection methods are used to generate relevant sentiment-rich features.
- In the second phase, the relevant feature vector is passed to a machine-learning algorithm for sentiment classification.

Feature Engineering

- Two types of features are extracted:
  - unigrams and bi-tagged phrase.

- Bi-tagged phrases are extracted using POS-based fixed patterns and represent better indicators of sentiment information.

- Bi-tagged phrases conforming to certain pattern (as shown in Table 1) are extracted.

Feature Engineering

- Table 1. Rules for extraction of Bi-tagged phrases

S.No	First Word	Second Word
1	JJ/JJR/JJS	NN/NNS
2	RB/RBR/RBS	JJ/JJR/JJS
3	JJ/JJR/JJS	JJ/JJR/JJS
4	NN/NNS	JJ/JJR/JJS
5	RB/RBR/RBS	VB/VBD/VBN/VBG
6	VB/VBD/VBN/ VBG	NN/NNS
7	VB/VBD/VBN/ VBG	JJ/JJR/JJS
8	NN/NNS	RB/RBR/RBS
9	JJ/JJR/JJS	VB/VBD/VBN/VBG

Feature Engineering

- Initially, unigram and bi-grams are extracted from text.
- Next, bi-tagged phrase features are extracted.
- Further, prominent features are extracted using the IG feature selection method.
- Prominent features extracted from unigrams, bigrams and bi-tagged phrase are named as ***PromUni*** (prominent unigrams), ***PromBi*** (prominent bi-grams) and ***PromBiTa*** (prominent bi-tagged) features respectively

Feature Engineering

- The performance of unigram features increases when combined with bi-grams.
- Composite features are created using unigram with bi-grams and unigrams with bi-tagged features namely ***ComUniBi*** and ***ComUniBiTa***, respectively.
- Finally, ***PromUniBiTa*** feature set is created by combining prominent unigrams (*PromUni*) and prominent bitagged features (*PromBiTa*).

Feature Engineering

- Table 2. Description of the feature sets

Feature set	Feature extraction method
Unigram	Unigrams
Bi-gram	Bigrams
Bi-tagged Phrases	Bi-tagged features as discussed
ComUniBi	Composite of Unigrams and Bigrams
ComUniBiTa	Composite of Unigrams and Bi-tagged features
PromUni	Prominent unigram features using IG as Feature Selection
PromBiTa	Prominent Bi-tagged features using IG as Feature Selection
PromUniBiTa	Composite of Prominent unigram and prominent bi-tagged features

Datasets

- Movie Review Dataset (Pang B., and Lee L., 2004).
  - http://www.cs.cornell.edu/People/pabo/movie-review-data/
  - Dataset is consisting of 2000 reviews that contain 1000 positive and 1000 negative labeled reviews.

Experimental Setup

- Documents are initially pre-processed as follows: "NOT_" is concatenated to every word between negation words (no, not, never, isn't, didn't etc.) and punctuation marks following the negation word.
- Binary weighting scheme is used for representing text.
- Support Vector Machine (SVM) and Naïve Bayes (NB) classifiers are used for sentiment classification. The Weka tool is used for implementing the two classifiers.
- Evaluation is performed using 10-fold cross validation.
- Performance of all the feature vectors are reported using F-measure.

Performance measures

- *F-measure is used for performance evaluation*

- $F - Measure = \frac{2 * precision * recall}{(precision + recall)}$

- Precision for a class C is the fraction of total number of documents that are correctly classified to the total number of documents that classified to the class C.

- Recall is the fraction of total number of correctly classified documents to the total number of documents that belongs to class C.

Results

- F-Measure (%) for various feature sets

Features	SVM	NB	Feature size
Unigram	84.2	79.4	9045
Bi-gram	78.8	73.5	6050
Bi-tagged Phrases	75.3	71.8	4841
ComUniBi	86.7	81.1	15095
ComUniBiTa	87.6	82.3	13886
PromUni	85.8	85.4	1130
PromBiTa	86.5	73.7	1114
PromUniBiTa	89.4	86.2	2244

Comparison with existing methods

- Proposed method depends on the basic unigrams, simple bi-tagged phrases and IG which are easy to extract and computationally efficient as compared to other methods proposed in the literature for sentiment classification on movie review dataset.
- Proposed approach produces comparable results with very much less computational cost.

Table.3 Performance of Various Methods on Movie Review Dataset

Paper	Approach	Best accuracy
Pang et al. (2004) [8]	Minimum cut algorithm, SVM	87.1
Prabowo et al. 2009. [9]	Hybrid SVM	87.3
O'keefee et al. 2009. [5]	SentiWordNet based features and feature selection with SVM,NB classifier	87.15
Ng et al.(2006) [13]	SVM with various features	90.5
Tu et al. (2012) [12]	Dependency forest based with MaxEnt	91.6
Abbasi et al., 2008 [1]	Genetic Algorithms (GA), Information Gain (IG), IG + GA	91.7
Proposed Approach	Unigrams, Bi- tagged phrases, IG	89.4

Conclusion

- Main objective of this paper is to investigate the performance of the bi-tagged features for supervised classification.
- In this paper, Bi-tagged features are used in addition to unigrams for enhancing the performance of the sentiment classification.
- Experimental results show that composite feature of prominent unigrams and prominent bi-tagged features perform better than other features for movie review sentiment classification.
- The main reason for this observation is that that bigrams contains very important sentiment information but with lots of noisy features which surpass the effect of context and sentiment information.
- However, Bi-tagged phrases are the sentiment-rich bi-grams which contain only subjective information that is very important for sentiment classification.
- Experimental results show the effectiveness of the proposed method.

References

[1] A. Abbasi, C. Chen, A. Salem, "Sentiment analysis in multiple languages: Feature selection for opinion classification in web forums", ACM Transactions on Information Systems 26(3), 2008

[2] B. Agarwal, N. Mittal, "Optimal Feature Selection Methods for Sentiment Analysis", Proceedings of the International Conference on Intelligent Text Processing and Computational Linguistics (CICLing), vol. 7817, pp. 13-24, 2013

[3] E. Cambria, T. Mazzocco, A. Hussain, "Application of multidimensional scaling and artificial neural networks for biologically inspired opinion mining", Biologically Inspired Cognitive Architectures 4, pp. 41-53, 2013

[4] B. Liu, "Sentiment Analysis and Opinion Mining", Synthesis Lectures on Human Language Technologies, Morgan & Claypool Publishers, 2012

[5] T. O'Keefe, I. Koprinska, "Feature Selection and Weighting Methods in Sentiment Analysis", Proceedings of the 14th Australasian Document Computing Symposium, 2009

[6] V. Hatzivassiloglou, K.R. McKeown, "Predicting the semantic orientation of adjectives", Proceedings of the 35th Annual Meeting of the ACL and the 8th Conference of the European Chapter of the ACL, pp. 174-181, 1997

[7] B. Pang,, L. Lee, S. Vaithyanathan, "Thumbs up? Sentiment classification using machine learning techniques", Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP), pp. 79–86, 2002

[8] B. Pang, L . Lee, "A sentimental education: sentiment analysis using subjectivity summarization based on minimum cuts", Proceedings of the Association for Computational Linguistics, pp. 271–278, 2004

[9] R. Prabowo, M. Thelwall, "Sentiment analysis: A combined approach", Journal of Informetrics 3(2), pp. 143-157, 2009

[10] S. Tan, J. Zhang, "An empirical study of sentiment analysis for chinese documents", Expert Systems with Applications 34, pp. 2622- 2629, 2008

References

[11] P. Turney, "Thumbs Up or Thumbs Down? Semantic Orientation Applied to Unsupervised Classification of Reviews", Proceedings of the Association for Computational Linguistics, 2002.

[12] Z. Tu, W. Jiang, Q. Liu, S. Lin, "Dependency Forest for Sentiment Analysis" Proceedings of the First CCF Conference, Natural Language Processing and Chinese Computing, pp 69-77, 2012.

[13] V. Ng, S. Dasgupta, S. M. Niaz Arifin, "Examining the Role of Linguistic Knowledge Sources in the Automatic Identification and Classification of Reviews", Proceedings of the COLING/ACL, pp. 611–618, 2006

[14] S. Wang, D. Li, S. Song, Y. Wei, H. Li, "A Feature Selection Method Based on Fisher's Discriminant Ratio for Text Sentiment Classification", Proceedings of the International Conference on Web Information Systems and Mining, pp. 88- 97, 2009

[15] WEKA. Open Source Machine Learning Software Weka. http://www.cs.waikato.ac.nz/ml/weka

[16] R. Xia, C. Zong, S. Li, "Ensemble of Feature Sets and Classification Algorithms for Sentiment Classification", Journal of Information Sciences 181(6), pp. 1138-1152, 2011

[17] R. Mukras,N. Wiratunga, R. Lothian, "Selecting Bi-Tags for Sentiment Analysis of Text", Proceedings of 27th SGAI International Conference on Innovative Techniques and Applications of Artificial Intelligence, pp. 181-194, 2008

[18] E. Cambria, D. Rajagopal, D. Olsher, D. Das, "Big social data analysis", R. Akerkar (ed.) Big Data Computing, ch. 13, pp. 401-414, Taylor & Francis, 2013

[19] R. Xia, C.Q. Zong, X.L. Hu, E. Cambria. "Feature ensemble plus sample selection: A comprehensive approach to domain adaptation for sentiment classification", IEEE Intelligent Systems 28(3), pp. 10-18, 2013

[20] E. Cambria, Y. Song, H. Wang, A. Hussain, "Isanette: A common and common sense knowledge base for opinion mining", Proceedings of IEEE ICDM, pp. 315-322, 2011

Thank You

☺