Document-level Sentiment Classification Based on Behavior-Knowledge Space Method

Zhifei Zhang Duoqian Miao Zhihua Wei

Tongji University
Department of Computer Science and Technology

8th Int'l Conf. on Advanced Data Mining and Applications, 2012

Outline

Outline

Background

ctrip 携程

总评: 5.0 卫生：5 服务：5 设施：5 位置：5

这个酒店真的非常好，房间又大又干净，周边环境也相当好，但酒店外的一条河实在太不卫生了。

360buy.com 京东商城

佳能A4000

优点：外表很漂亮做工非常不错。

不足：液晶屏清晰度不够好。

使用心得：开机速度非常快，一开就能使用，防抖做的很好。

豆瓣 douban

Prototype 2012-09-06

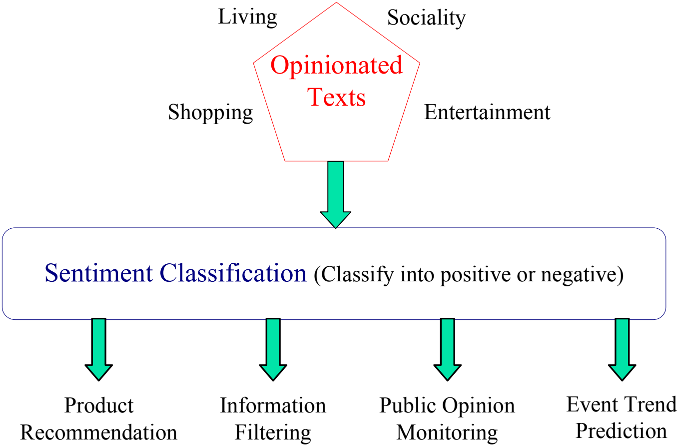
很不错，周迅的演技是个亮点，梁朝伟依旧很帅

新浪微博

李意如leah：钓鱼岛是中国滴，小日本想买？没门，没打你就是客气了。#钓鱼岛是中国的#

User-generated texts from the Internet contain subjective information.

Background (Cont.)



Sentiment Classification

Divided into three levels by the granularity of text:

- Word-level sentiment classification
  e.g., "good" (Positive) "bad" (Negative)
- Sentence-level sentiment classification
  e.g., "The price is high" (Negative)
- Document-level sentiment classification
  e.g., "The phone is expensive, but I like it." (Positive)

In this paper, we focus on document-level sentiment classification.

Sentiment Classification

Divided into three levels by the granularity of text:

- Word-level sentiment classification
  e.g., "good" (Positive) "bad" (Negative)
- Sentence-level sentiment classification
  e.g., "The price is high" (Negative)
- Document-level sentiment classification
  e.g., "The phone is expensive, but I like it." (Positive)

In this paper, we focus on document-level sentiment classification.

Sentiment Classification

Divided into three levels by the granularity of text:

- Word-level sentiment classification
  e.g., "good" (Positive) "bad" (Negative)
- Sentence-level sentiment classification
  e.g., "The price is high" (Negative)
- Document-level sentiment classification
  e.g., "The phone is expensive, but I like it." (Positive)

In this paper, we focus on document-level sentiment classification.

Sentiment Classification

Divided into three levels by the granularity of text:

- Word-level sentiment classification
  e.g., "good" (Positive) "bad" (Negative)
- Sentence-level sentiment classification
  e.g., "The price is high" (Negative)
- Document-level sentiment classification
  e.g., "The phone is expensive, but I like it." (Positive)

In this paper, we focus on document-level sentiment classification.

Sentiment Classification

Divided into three levels by the granularity of text:

- Word-level sentiment classification
  e.g., “good” (Positive) “bad” (Negative)
- Sentence-level sentiment classification
  e.g., “The price is high” (Negative)
- Document-level sentiment classification
  e.g., “The phone is expensive, but I like it.” (Positive)

In this paper, we focus on document-level sentiment classification.

Sentiment Classification

Divided into three levels by the granularity of text:

- Word-level sentiment classification
  e.g., “good” (Positive) “bad” (Negative)
- Sentence-level sentiment classification
  e.g., “The price is high” (Negative)
- Document-level sentiment classification
  e.g., “The phone is expensive, but I like it.” (Positive)

In this paper, we focus on document-level sentiment classification.

Document-level Sentiment Classification

To give the overall polarity of a document (Positive, Negative)

- Unsupervised methods[1]
  = accumulate the polarity of each opinion word
- Supervised methods[2]
  = convert into multi-class text classification task

[1] Turney, P.: Thumbs Up or Thumbs Down? Semantic Orientation Applied to Unsupervised Classification of Reviews. ACL 2002, pp. 417–424.

[2] Pang, B., Lee, L., Vaithyanathan, S.: Thumbs Up? Sentiment Classification Using Machine Learning Techniques. EMNLP 2002, pp. 79–86.

Document-level Sentiment Classification

To give the overall polarity of a document (Positive, Negative)

- Unsupervised methods[1]
  = accumulate the polarity of each opinion word
- Supervised methods[2]
  = convert into multi-class text classification task

[1] Turney, P.: Thumbs Up or Thumbs Down? Semantic Orientation Applied to Unsupervised Classification of Reviews. ACL 2002, pp. 417–424.

[2] Pang, B., Lee, L., Vaithyanathan, S.: Thumbs Up? Sentiment Classification Using Machine Learning Techniques. EMNLP 2002, pp. 79–86.

Document-level Sentiment Classification

To give the overall polarity of a document (Positive, Negative)

- Unsupervised methods[1]
  = accumulate the polarity of each opinion word
- Supervised methods[2]
  = convert into multi-class text classification task

[1] Turney, P.: Thumbs Up or Thumbs Down? Semantic Orientation Applied to Unsupervised Classification of Reviews. ACL 2002, pp. 417–424.

[2] Pang, B., Lee, L., Vaithyanathan, S.: Thumbs Up? Sentiment Classification Using Machine Learning Techniques. EMNLP 2002, pp. 79–86.

Outline

Ensemble Learning

Ensemble learning is to combine outputs of multiple classifiers. Two conditions for validity[3]:

- the accurate rate of each basic classifier is not below 0.5
- the component classifiers disagree with one another (diversity)

Diversity can be measured by Q-statistic[4].

	H_j True	H_j False
H_i True	A	B
H_i False	C	D

$$Q_{ij} = \left| \frac{AD - BC}{AD + BC} \right| \tag{1}$$

[3] Dietterich, T.G.: Machine Learning Research: Four Current Directions. AI Magazine, 1997, 18(4): 97–136.

[4] Kuncheva, L.I., Whitaker, C.J.: Measures of Diversity in Classifier Ensembles. Machine Learning 2003, 51: 181–207.

Ensemble Learning

Ensemble learning is to combine outputs of multiple classifiers. Two conditions for validity[3]:

- the accurate rate of each basic classifier is not below 0.5
- the component classifiers disagree with one another (diversity)

Diversity can be measured by Q-statistic[4].

	H_j True	H_j False
H_i True	A	B
H_i False	C	D

$$Q_{ij} = \left| \frac{AD - BC}{AD + BC} \right| \quad (1)$$

[3] Dietterich, T.G.: Machine Learning Research: Four Current Directions. AI Magazine, 1997, 18(4): 97–136.

[4] Kuncheva, L.I., Whitaker, C.J.: Measures of Diversity in Classifier Ensembles. Machine Learning, 2003, 51: 181–207.

Ensemble Learning

Ensemble learning is to combine outputs of multiple classifiers. Two conditions for validity[3]:

- the accurate rate of each basic classifier is not below 0.5
- the component classifiers disagree with one another (diversity)

Diversity can be measured by Q-statistic[4].

	H_j True	H_j False
H_i True	A	B
H_i False	C	D

$$Q_{ij} = \left| \frac{AD - BC}{AD + BC} \right| \quad (1)$$

[3] Dietterich, T.G.: Machine Learning Research: Four Current Directions. AI Magazine, 1997, 18(4): 97–136.

[4] Kuncheva, L.I., Whitaker, C.J.: Measures of Diversity in Classifier Ensembles. Machine Learning 2003, 51: 181–207.

Ensemble Learning

Ensemble learning is to combine outputs of multiple classifiers. Two conditions for validity[3]:

- the accurate rate of each basic classifier is not below 0.5
- the component classifiers disagree with one another (diversity)

Diversity can be measured by Q-statistic[4].

	H_j True	H_j False
H_i True	A	B
H_i False	C	D

$$Q_{ij} = \left| \frac{AD - BC}{AD + BC} \right| \quad (1)$$

[3] Dietterich, T.G.: Machine Learning Research: Four Current Directions. AI Magazine, 1997, 18(4): 97–136.

[4] Kuncheva, L.I., Whitaker, C.J.: Measures of Diversity in Classifier Ensembles. Machine Learning 2003, 51: 181–207.

Ensemble Learning

Ensemble learning is to combine outputs of multiple classifiers. Two conditions for validity[3]:

- the accurate rate of each basic classifier is not below 0.5
- the component classifiers disagree with one another (diversity)

Diversity can be measured by Q-statistic[4].

	H_j True	H_j False
H_i True	A	B
H_i False	C	D

$$Q_{ij} = \left| \frac{AD - BC}{AD + BC} \right| \tag{1}$$

[3] Dietterich, T.G.: Machine Learning Research: Four Current Directions. AI Magazine, 1997, 18(4): 97–136.

[4] Kuncheva, L.I., Whitaker, C.J.: Measures of Diversity in Classifier Ensembles. Machine Learning, 2003, 51: 181–207.

Existing Work

- Xia et al. focus on ensemble of feature sets and classifiers, in which all classifiers belong to supervised learning.[5]
- Wan combines unsupervised learning methods using bilingual knowledge.[6]
- The common heterogeneous ensemble strategies: Voting, Bayes' Rule and Behavior-Knowledge Space.

In this paper, we focus on ensemble of unsupervised learning and supervised learning based on Behavior-Knowledge Space method.

[5] Xia, R., Zong C.Q., Li S.S.: Ensemble of Feature Sets and Classification Algorithms for Sentiment Classification. Information Sciences, 2011, 181: 1138–1152.

[6] Wan, X.J.: Using Bilingual Knowledge and Ensemble Techniques for Unsupervised Chinese Sentiment Analysis. EMNLP 2008, pp. 553–561.

Existing Work

- Xia et al. focus on ensemble of feature sets and classifiers, in which all classifiers belong to supervised learning.[5]
- Wan combines unsupervised learning methods using bilingual knowledge.[6]
- The common heterogeneous ensemble strategies: Voting, Bayes' Rule and Behavior-Knowledge Space.

In this paper, we focus on ensemble of unsupervised learning and supervised learning based on Behavior-Knowledge Space method.

[5] Xia, R., Zong C.Q., Li S.S.: Ensemble of Feature Sets and Classification Algorithms for Sentiment Classification. Information Sciences, 2011, 181: 1138–1152.

[6] Wan, X.J.: Using Bilingual Knowledge and Ensemble Techniques for Unsupervised Chinese Sentiment Analysis. EMNLP 2008, pp. 553–561.

Existing Work

- Xia et al. focus on ensemble of feature sets and classifiers, in which all classifiers belong to supervised learning.[5]
- Wan combines unsupervised learning methods using bilingual knowledge.[6]
- The common heterogeneous ensemble strategies: Voting, Bayes' Rule and Behavior-Knowledge Space.

In this paper, we focus on ensemble of unsupervised learning and supervised learning based on Behavior-Knowledge Space method.

[5] Xia, R., Zong C.Q., Li S.S.: Ensemble of Feature Sets and Classification Algorithms for Sentiment Classification. Information Sciences, 2011, 181: 1138–1152.

[6] Wan, X.J.: Using Bilingual Knowledge and Ensemble Techniques for Unsupervised Chinese Sentiment Analysis. EMNLP 2008, pp. 553–561.

Existing Work

- Xia et al. focus on ensemble of feature sets and classifiers, in which all classifiers belong to supervised learning.[5]
- Wan combines unsupervised learning methods using bilingual knowledge.[6]
- The common heterogeneous ensemble strategies: Voting, Bayes' Rule and Behavior-Knowledge Space.

In this paper, we focus on ensemble of unsupervised learning and supervised learning based on Behavior-Knowledge Space method.

[5] Xia, R., Zong C.Q., Li S.S.: Ensemble of Feature Sets and Classification Algorithms for Sentiment Classification. Information Sciences, 2011, 181: 1138–1152.

[6] Wan, X.J.: Using Bilingual Knowledge and Ensemble Techniques for Unsupervised Chinese Sentiment Analysis. EMNLP 2008, pp. 553–561.

Existing Work

- Xia et al. focus on ensemble of feature sets and classifiers, in which all classifiers belong to supervised learning.[5]
- Wan combines unsupervised learning methods using bilingual knowledge.[6]
- The common heterogeneous ensemble strategies: Voting, Bayes' Rule and Behavior-Knowledge Space.

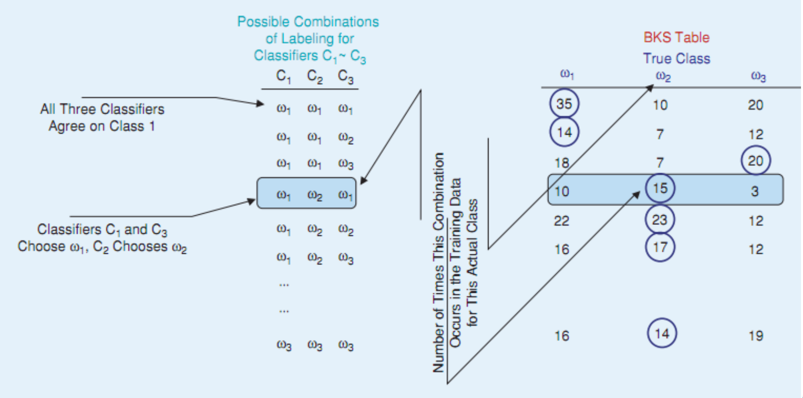
In this paper, we focus on ensemble of unsupervised learning and supervised learning based on Behavior-Knowledge Space method.

[5] Xia, R., Zong C.Q., Li S.S.: Ensemble of Feature Sets and Classification Algorithms for Sentiment Classification. Information Sciences, 2011, 181: 1138–1152.

[6] Wan, X.J.: Using Bilingual Knowledge and Ensemble Techniques for Unsupervised Chinese Sentiment Analysis. EMNLP 2008, pp. 553–561.

Behavior-Knowledge Space

Proposed by Huang and Suen[7]. The method using the history of the classifiers behavior allows getting free from the condition of independence.



[7] Huang, Y.S., Suen C.Y.: The Behavior-Knowledge Space Method for Combination of Multiple Classifiers. CVPR 1993, pp. 347–352.

Behavior-Knowledge Space (Cont.)

Definition

- The number of classifiers: K
- The number of classes: M
- The decision of a classifier: $e(i)$, $e(i) \in \{1, 2, ..., M, M+1\}$, $M+1$ is a rejection class
- The behavior knowledge space: BKS, the intersection of the decision of each classifier corresponds to a cell of BKS
- Each cell of the space: $BKS(e(1), e(2), ..., e(K))$

Behavior-Knowledge Space (Cont.)

Each cell of the BKS contains three features:

- The total number of samples with Class m: $n_{e(1),\ldots,e(K)}(m)$
- The total number of samples: $T_{e(1),\ldots,e(K)}$

$$T_{e(1),\ldots,e(K)} = \sum_{m=1}^{M} n_{e(1),\ldots,e(K)}(m) \tag{2}$$

- The best representative class: $R_{e(1),\ldots,e(K)}$

$$R_{e(1),\ldots,e(K)} = \arg\max_{m}\{n_{e(1),\ldots,e(K)}(m) | 1 \leq m \leq M\} \tag{3}$$

Behavior-Knowledge Space (Cont.)

The degree of belief that a sample belongs to Class m, $BEL(m)$

$$BEL(m) = \frac{n_{e(1),\ldots,e(K)}(m)}{T_{e(1),\ldots,e(K)}} \tag{4}$$

The ensemble of classifiers gives the input x the following class

$$D(x) = \begin{cases} R_{e(1),\ldots,e(K)} & \text{when } T_{e(1),\ldots,e(K)} > 0 \text{ and } BEL(R_{e(1),\ldots,e(K)}) \geq \alpha \\ M+1 & \text{otherwise} \end{cases} \tag{5}$$

where α is a rejection threshold, $0 \leq \alpha \leq 1$.

Behavior-Knowledge Space (Cont.)

The degree of belief that a sample belongs to Class m, $BEL(m)$

$$BEL(m) = \frac{n_{e(1),\dots,e(K)}(m)}{T_{e(1),\dots,e(K)}} \tag{4}$$

The ensemble of classifiers gives the input x the following class

$$D(x) = \begin{cases} R_{e(1),\dots,e(K)} & \text{when } T_{e(1),\dots,e(K)} > 0 \text{ and } BEL(R_{e(1),\dots,e(K)}) \geq \alpha \\ M+1 & \text{otherwise} \end{cases} \tag{5}$$

where α is a rejection threshold, $0 \leq \alpha \leq 1$.

Outline

Problem Statement

A problem with two classes and four classifiers.

Two classes: Positive, Negative

Four classifiers

- Unsupervised: SWS, WSC[8]
- Supervised: SVM, KNN[9]

So, $K = 4$ and $M = 2$ in BKS method.

[8] Chen, Y.F., Miao, D.Q., Zhang, Z.F.: Semantic Orientation Computing Based on Concepts. CAAI Trans. on Intelligent Systems, 2011, 6(6): 489–494.

[9] Tan, S.B., Zhang, J.: An Empirical Study of Sentiment Analysis for Chinese Documents. Expert Systems with Applications, 2008, 34: 2622–2629.

Problem Statement

A problem with two classes and four classifiers.

Two classes: Positive, Negative

Four classifiers

- Unsupervised: SWS, WSC[8]
- Supervised: SVM, KNN[9]

So, $K = 4$ and $M = 2$ in BKS method.

[8] Chen, Y.F., Miao, D.Q., Zhang, Z.F.: Semantic Orientation Computing Based on Concepts. CAAI Trans. on Intelligent Systems, 2011, 6(6): 489–494.

[9] Tan, S.B., Zhang, J.: An Empirical Study of Sentiment Analysis for Chinese Documents. Expert Systems with Applications, 2008, 34: 2622–2629.

Problem Statement

A problem with two classes and four classifiers.

Two classes: Positive, Negative

Four classifiers

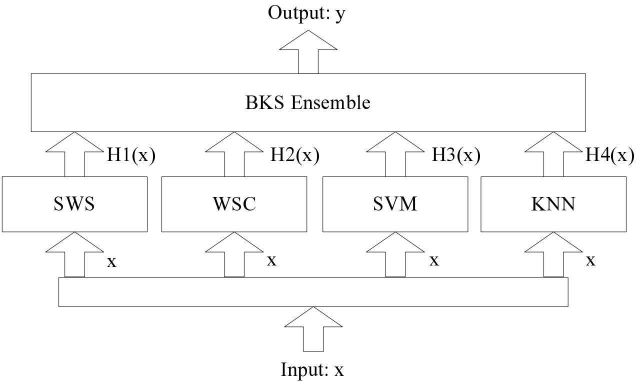
- Unsupervised: SWS, WSC[8]
- Supervised: SVM, KNN[9]

So, $K = 4$ and $M = 2$ in BKS method.

[8] Chen, Y.F., Miao, D.Q., Zhang, Z.F.: Semantic Orientation Computing Based on Concepts. CAAI Trans. on Intelligent Systems, 2011, 6(6): 489–494.

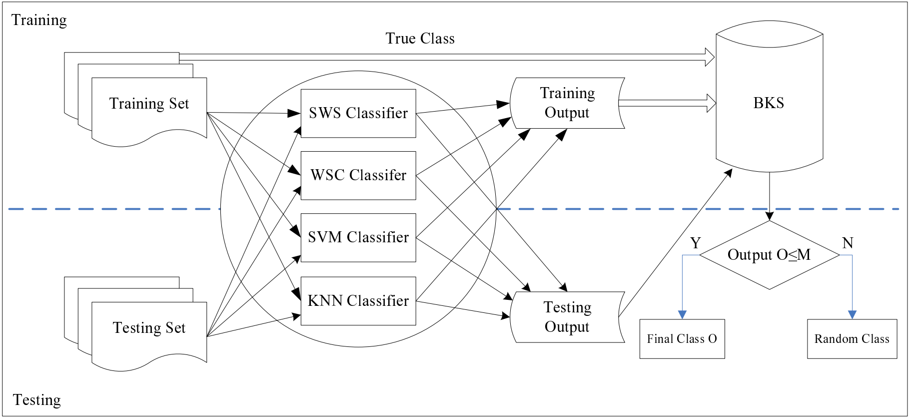
[9] Tan, S.B., Zhang, J.: An Empirical Study of Sentiment Analysis for Chinese Documents. Expert Systems with Applications, 2008, 34: 2622–2629.

Problem Statement (Cont.)



Here, $H1(x)$, $H2(x)$, $H3(x)$, $H4(x)$ and y are all from {Positive, Negative, Reject}.

Method Illustration



Method Illustration (Cont.)

Algorithm: Document-level SC Based on BKS

Input: Training set, TS; Test set, NS;
Output: Decision classes of test samples, D;

1. Build up two supervised classifiers, SVM and KNN, using the training set TS;
2. Obtain the decision classes by four basic classifiers for each sample in TS, denoted by $TID_{|TS|\times 4}$;
3. From $TID_{|TS|\times 4}$ and the true classes in TS, compute T and R for each cell of BKS by Eq.2 and Eq.3;
4. Obtain the decision classes by four basic classifiers for each sample in NS, denoted by $NID_{|NS|\times 4}$;
5. From $NID_{|NS|\times 4}$ and BKS, make decisions for all test samples by Eq.5, denoted by D';
6. For all the i-th test sample in NS, if $D'(i) < M+1$, then $D(i) = D'(i)$, otherwise $D(i)$ is randomly generated from $\{1, ..., M\}$; $(M = 2)$
7. Return D.

Method Illustration (Cont.)

Algorithm: Document-level SC Based on BKS

Input: Training set, TS; Test set, NS;
Output: Decision classes of test samples, D;

1. Build up two supervised classifiers, SVM and KNN, using the training set TS;
2. Obtain the decision classes by four basic classifiers for each sample in TS, denoted by $TID_{|TS|\times 4}$;
3. From $TID_{|TS|\times 4}$ and the true classes in TS, compute T and R for each cell of BKS by Eq.2 and Eq.3;
4. Obtain the decision classes by four basic classifiers for each sample in NS, denoted by $NID_{|NS|\times 4}$;
5. From $NID_{|NS|\times 4}$ and BKS, make decisions for all test samples by Eq.5, denoted by D';
6. For all the i-th test sample in NS, if $D'(i) < M + 1$, then $D(i) = D'(i)$, otherwise $D(i)$ is randomly generated from $\{1, ..., M\}$; $(M = 2)$
7. Return D.

Method Illustration (Cont.)

Algorithm: Document-level SC Based on BKS

Input: Training set, TS; Test set, NS;
Output: Decision classes of test samples, D;

1. Build up two supervised classifiers, SVM and KNN, using the training set TS;
2. Obtain the decision classes by four basic classifiers for each sample in TS, denoted by $TID_{|TS|\times 4}$;
3. From $TID_{|TS|\times 4}$ and the true classes in TS, compute T and R for each cell of BKS by Eq.2 and Eq.3;
4. Obtain the decision classes by four basic classifiers for each sample in NS, denoted by $NID_{|NS|\times 4}$;
5. From $NID_{|NS|\times 4}$ and BKS, make decisions for all test samples by Eq.5, denoted by D';
6. For all the i-th test sample in NS, if $D'(i) < M + 1$, then $D(i) = D'(i)$, otherwise $D(i)$ is randomly generated from $\{1, ..., M\}$; $(M = 2)$
7. Return D.

Method Illustration (Cont.)

Algorithm: Document-level SC Based on BKS

Input: Training set, TS; Test set, NS;
Output: Decision classes of test samples, D;

1. Build up two supervised classifiers, SVM and KNN, using the training set TS;
2. Obtain the decision classes by four basic classifiers for each sample in TS, denoted by $TID_{|TS| \times 4}$;
3. From $TID_{|TS| \times 4}$ and the true classes in TS, compute T and R for each cell of BKS by Eq.2 and Eq.3;
4. Obtain the decision classes by four basic classifiers for each sample in NS, denoted by $NID_{|NS| \times 4}$;
5. From $NID_{|NS| \times 4}$ and BKS, make decisions for all test samples by Eq.5, denoted by D';
6. For all the i-th test sample in NS, if $D'(i) < M + 1$, then $D(i) = D'(i)$, otherwise $D(i)$ is randomly generated from $\{1, ..., M\}$; $(M = 2)$
7. Return D.

Method Illustration (Cont.)

Algorithm: Document-level SC Based on BKS

Input: Training set, TS; Test set, NS;
Output: Decision classes of test samples, D;

1. Build up two supervised classifiers, SVM and KNN, using the training set TS;
2. Obtain the decision classes by four basic classifiers for each sample in TS, denoted by $TID_{|TS|\times 4}$;
3. From $TID_{|TS|\times 4}$ and the true classes in TS, compute T and R for each cell of BKS by Eq.2 and Eq.3;
4. Obtain the decision classes by four basic classifiers for each sample in NS, denoted by $NID_{|NS|\times 4}$;
5. From $NID_{|NS|\times 4}$ and BKS, make decisions for all test samples by Eq.5, denoted by D';
6. For all the i-th test sample in NS, if $D'(i) < M + 1$, then $D(i) = D'(i)$, otherwise $D(i)$ is randomly generated from $\{1, ..., M\}$; $(M = 2)$
7. Return D.

Method Illustration (Cont.)

Algorithm: Document-level SC Based on BKS

Input: Training set, TS; Test set, NS;
Output: Decision classes of test samples, D;

1. Build up two supervised classifiers, SVM and KNN, using the training set TS;
2. Obtain the decision classes by four basic classifiers for each sample in TS, denoted by $TID_{|TS|\times 4}$;
3. From $TID_{|TS|\times 4}$ and the true classes in TS, compute T and R for each cell of BKS by Eq.2 and Eq.3;
4. Obtain the decision classes by four basic classifiers for each sample in NS, denoted by $NID_{|NS|\times 4}$;
5. From $NID_{|NS|\times 4}$ and BKS, make decisions for all test samples by Eq.5, denoted by D';
6. For all the i-th test sample in NS, if $D'(i) < M + 1$, then $D(i) = D'(i)$, otherwise $D(i)$ is randomly generated from $\{1, ..., M\}$; $(M = 2)$
7. Return D.

Method Illustration (Cont.)

Algorithm: Document-level SC Based on BKS

Input: Training set, TS; Test set, NS;
Output: Decision classes of test samples, D;

1. Build up two supervised classifiers, SVM and KNN, using the training set TS;
2. Obtain the decision classes by four basic classifiers for each sample in TS, denoted by $TID_{|TS|\times 4}$;
3. From $TID_{|TS|\times 4}$ and the true classes in TS, compute T and R for each cell of BKS by Eq.2 and Eq.3;
4. Obtain the decision classes by four basic classifiers for each sample in NS, denoted by $NID_{|NS|\times 4}$;
5. From $NID_{|NS|\times 4}$ and BKS, make decisions for all test samples by Eq.5, denoted by D';
6. For all the i-th test sample in NS, if $D'(i) < M + 1$, then $D(i) = D'(i)$, otherwise $D(i)$ is randomly generated from $\{1, ..., M\}$; $(M = 2)$
7. Return D.

Method Illustration (Cont.)

Algorithm: Document-level SC Based on BKS

Input: Training set, TS; Test set, NS;
Output: Decision classes of test samples, D;

1. Build up two supervised classifiers, SVM and KNN, using the training set TS;
2. Obtain the decision classes by four basic classifiers for each sample in TS, denoted by $TID_{|TS|\times 4}$;
3. From $TID_{|TS|\times 4}$ and the true classes in TS, compute T and R for each cell of BKS by Eq.2 and Eq.3;
4. Obtain the decision classes by four basic classifiers for each sample in NS, denoted by $NID_{|NS|\times 4}$;
5. From $NID_{|NS|\times 4}$ and BKS, make decisions for all test samples by Eq.5, denoted by D';
6. For all the i-th test sample in NS, if $D'(i) < M + 1$, then $D(i) = D'(i)$, otherwise $D(i)$ is randomly generated from $\{1, ..., M\}$; $(M = 2)$
7. Return D.

Outline

Experiment Setting

- Data set: ChnSentiCorp-Htl-ba-4000 from ChnSenti-Corp[10](2000 positive and 2000 negative texts about hotels)
- Rejection threshold: set to be 0.55 according to experiences
- Evaluation metrics: Precision, Recall, and F_1

$$F_1 = \frac{(PP + PN) \times (RP + RN)}{PP + PN + RP + RN} \quad (6)$$

PP (precision for positive), RP (recall for positive)

PN (precision for negative), RN (recall for negative)

[10] http://www.searchforum.org.cn/tansongbo/corpus-senti.htm

Experiment Setting

- Data set: ChnSentiCorp-Htl-ba-4000 from ChnSenti-Corp[10](2000 positive and 2000 negative texts about hotels)
- Rejection threshold: set to be 0.55 according to experiences
- Evaluation metrics: Precision, Recall, and F_1

$$F_1 = \frac{(PP + PN) \times (RP + RN)}{PP + PN + RP + RN} \tag{6}$$

PP (precision for positive), RP (recall for positive)

PN (precision for negative), RN (recall for negative)

[10] http://www.searchforum.org.cn/tansongbo/corpus-senti.htm

Experiment Setting

- Data set: ChnSentiCorp-Htl-ba-4000 from ChnSenti-Corp[10](2000 positive and 2000 negative texts about hotels)
- Rejection threshold: set to be 0.55 according to experiences
- Evaluation metrics: Precision, Recall, and F_1

$$F_1 = \frac{(PP + PN) \times (RP + RN)}{PP + PN + RP + RN} \quad (6)$$

PP (precision for positive), RP (recall for positive)

PN (precision for negative), RN (recall for negative)

[10] http://www.searchforum.org.cn/tansongbo/corpus-senti.htm

Effectiveness of BKS Ensemble

Results of four classifiers

Classifier	PP	RP	PN	RN	F_1
SWS	0.850	0.852	0.851	0.850	0.851
WSC	0.681	0.778	0.748	0.631	0.709
SVM	0.898	0.826	0.839	0.905	0.867
KNN	0.857	0.878	0.875	0.854	0.866

The accurate rate of each classifier is higher than 0.5 so that one condition of effective ensemble learning is satisfied.

Effectiveness of BKS Ensemble

Results of four classifiers

Classifier	PP	RP	PN	RN	F_1
SWS	0.850	0.852	0.851	0.850	0.851
WSC	0.681	0.778	0.748	0.631	0.709
SVM	0.898	0.826	0.839	0.905	0.867
KNN	0.857	0.878	0.875	0.854	0.866

The accurate rate of each classifier is higher than 0.5 so that one condition of effective ensemble learning is satisfied.

Effectiveness of BKS Ensemble (Cont.)

Diversity measure of four classifiers

	SWS	WSC	SVM	KNN
SWS	—	0.51	0.61	0.45
WSC	0.51	—	0.24	0.26
SVM	0.61	0.24	—	0.90
KNN	0.45	0.26	0.90	—

The diversity of these four classifiers is guaranteed so that our ensemble method is effective.

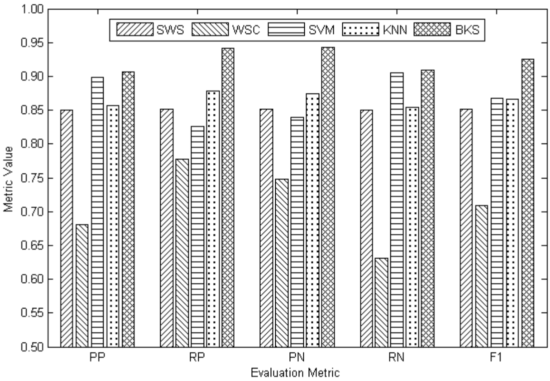
Effectiveness of BKS Ensemble (Cont.)

Diversity measure of four classifiers

	SWS	WSC	SVM	KNN
SWS	—	0.51	0.61	0.45
WSC	0.51	—	0.24	0.26
SVM	0.61	0.24	—	0.90
KNN	0.45	0.26	0.90	—

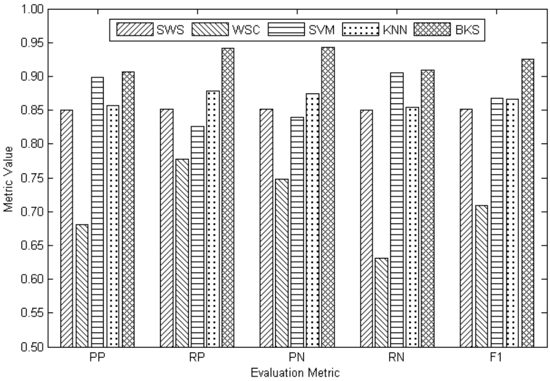
The diversity of these four classifiers is guaranteed so that our ensemble method is effective.

Performance of BKS Ensemble



- BKS ensemble method outperforms four basic classifiers, F_1 is up to 92.5%.
- The difference of performances in positive and negative samples is not significant.

Performance of BKS Ensemble



- BKS ensemble method outperforms four basic classifiers, F_1 is up to 92.5%.
- The difference of performances in positive and negative samples is not significant.

Performance of BKS Ensemble (Cont.)

Comparison results of three ensemble methods

Method	PP	RP	PN	RN	F_1	Random Times
Voting	0.884	0.883	0.883	0.884	0.884	162
Bayes' Rule	0.880	0.909	0.911	0.884	0.896	0
BKS	0.906	0.941	0.943	0.909	0.925	14

- BKS is better than the other two ensemble methods.
- BKS needs few times to make random decisions, which is acceptable.

Performance of BKS Ensemble (Cont.)

Comparison results of three ensemble methods

Method	PP	RP	PN	RN	F_1	Random Times
Voting	0.884	0.883	0.883	0.884	0.884	162
Bayes' Rule	0.880	0.909	0.911	0.884	0.896	0
BKS	0.906	0.941	0.943	0.909	0.925	14

- BKS is better than the other two ensemble methods.
- BKS needs few times to make random decisions, which is acceptable.

Outline

Our work

- An ensemble of two unsupervised and two supervised methods is proposed for document-level sentiment classification.
- The ensemble classifier based on Behavior-Knowledge Space method is effective and outperforms each basic classifier.
- Our method is better than the other two ensemble methods, Voting and Bayes' Rule.

Future work:

- Experiments with more corpora and more classification algorithms.
- A solution to random decisions in BKS ensemble.

Our work

- An ensemble of two unsupervised and two supervised methods is proposed for document-level sentiment classification.
- The ensemble classifier based on Behavior-Knowledge Space method is effective and outperforms each basic classifier.
- Our method is better than the other two ensemble methods, Voting and Bayes' Rule.

Future work:

- Experiments with more corpora and more classification algorithms.
- A solution to random decisions in BKS ensemble.

Our work

- An ensemble of two unsupervised and two supervised methods is proposed for document-level sentiment classification.
- The ensemble classifier based on Behavior-Knowledge Space method is effective and outperforms each basic classifier.
- Our method is better than the other two ensemble methods, Voting and Bayes' Rule.

Future work:

- Experiments with more corpora and more classification algorithms.
- A solution to random decisions in BKS ensemble.

Our work

- An ensemble of two unsupervised and two supervised methods is proposed for document-level sentiment classification.
- The ensemble classifier based on Behavior-Knowledge Space method is effective and outperforms each basic classifier.
- Our method is better than the other two ensemble methods, Voting and Bayes' Rule.

Future work:

- Experiments with more corpora and more classification algorithms.
- A solution to random decisions in BKS ensemble.

Our work

- An ensemble of two unsupervised and two supervised methods is proposed for document-level sentiment classification.
- The ensemble classifier based on Behavior-Knowledge Space method is effective and outperforms each basic classifier.
- Our method is better than the other two ensemble methods, Voting and Bayes' Rule.

Future work:

- Experiments with more corpora and more classification algorithms.
- A solution to random decisions in BKS ensemble.

Thank you !

Q & A?

Email:tjzhifei@163.com