

# 意見分析コーパスに関する現状調査

関 洋平<sup>1,a)</sup>

概要：インターネット，Web を利用した電子メディアの発展とともに，ユーザ個人が発信する意見の影響は大きくなり，その利用方法は変遷し続けている．電子的に蓄積された意見情報を分析することにより，消費活動や行動判断に有用な情報へのアクセスを支援する研究として，主に自然言語処理，情報検索，データ工学などの分野で，意見分析あるいは意見マイニングの研究が行われている．意見を分析するための表層的な特徴は，文書ジャンルあるいはメディアごとに異なることが知られており，ニュース，レビュー，ブログ，マイクロブログなど，様々なメディアごとにコーパスが重要な役割を果たす．本稿では，意見分析についての最近の研究におけるコーパスの使用状況を明らかにするとともに，意見分析コーパスの代表的な事例について紹介する．

## 1. はじめに

ここ数年，Web などの大量の電子化テキストに現れる他者が発信した意見情報を抽出し，集約や可視化を行うことで，世論調査や評判分析といった応用を実現する研究が進んでいる [1], [2], [3], [4]．これらの研究を総称して，意見分析 (*Sentiment Analysis*) あるいは意見マイニング (*Opinion Mining*) と呼ぶ [1]．対象となる文書ジャンルは，報道機関が配信するニュース，Web 上のレビューサイト，個人が自身の体験や意見を記述するブログやマイクロブログなどであり，政策や選挙のための情報分析，世論調査，商品や映画やレストラン・ホテルなどのサービスの評判分析，トレンド分析，などについて実用化が進められている．

意見を分析するためには，意見とは何かを定式化する必要がある．ただし，一口に意見情報といってもその特徴はさまざまであり，文書ジャンル (例：新聞，ブログ，SNS・QA サイトなどのコミュニティサイト，雑誌，会議録等) やドメイン (政治，映画，商品，恋愛相談等) に応じて，文書中で使用される概念や語彙の傾向は異なり，その話題についての意見情報の傾向も異なり，さまざまな文書ジャンル，ドメインを対象とした意見分析コーパスの需要は高い．

本論文の構成は以下の通りである．2 節では，意見分析についての最近の研究におけるコーパスの使用状況を，その研究内容とともに明らかにする．3 節では，意見分析コーパスの代表的な事例とその特徴を紹介する．最後に，4 節で，まとめと今後の展望について述べる．

## 2. 意見分析についての最近の研究におけるコーパスの使用状況

意見分析の研究は，2002 年ごろから活発に行われており [5], [6], [7]，初期の研究は，レビュー文書を用いた教師あり学習 [5] や教師なし学習 [6] に基づく文書単位の肯定・否定の分類が活発に行われていた．一方で，ニュース文書を対象としたアノテーションコーパスの作成も同時期に始まっている [7]．その後，数多くの国際ワークショップ (*AAAI 2004 Spring Symposium: Exploring Attitude and Affect in Text: Theories and Applications*[8]; *COLING-ACL 2006 Workshop: Sentiment and Subjectivity in Text; Workshop on Computational Approaches to Subjectivity and Sentiment Analysis (WASSA) 2010-2012; Sentiment Analysis Symposium*<sup>\*1</sup>等) や評価型ワークショップ (*TREC Blog Track 2006-2010; NTCIR Multilingual Opinion Analysis Tasks; SemEval-2007: Affective Text; TAC 2008 Opinion Question Answering and Summarization Tasks; TREC Microblog Track 2011*-等) を経て，発展を続けている．また，日本でも，2005 年の第 12 回言語処理学会年次大会併設ワークショップ「感情・評価・態度と言語」，2007 年の電子情報通信学会 言語理解とコミュニケーション研究会主催の「主観表現処理の最前線」シンポジウム，2008 年以降の楽天株式会社主催の「楽天研究開発シンポジウム」など，着実に研究の場を広げている．これらの研究の変遷は，[1], [2], [3] などにまとめられている．

これらを踏まえ，意見分析についての最近の研究にお

<sup>1</sup> 筑波大学  
University of Tsukuba

<sup>a)</sup> yohei@slis.tsukuba.ac.jp

<sup>\*1</sup> <http://sentimentsymposium.com/>

るコーパスの使用状況を明らかにするために、自然言語処理、情報検索の国際会議で発表された 31 件の意見分析の研究について、コーパスの使用状況と研究内容を整理した結果を以下に示す\*2。

- (1) ニュースを主な対象としたもの (6 件)
  - (a) 意見性、極性 (肯定, 否定, 中立), 意見の強弱の分類
    - (i) 行列空間モデルに基づく定式化 [9]
    - (ii) 談話関係の考慮 [10]
    - (iii) 多言語データの利用 [11]
  - (b) ドメインを横断した意見保有者の判定 [12]
  - (c) 観点の分類 [13], [14]
- (2) レビュー文書を対象としたもの (13 件)
  - (a) 意見性、極性 (肯定, 否定, 中立), 意見の強弱の分類
    - (i) 言語モデルに基づく定式化 [15]
    - (ii) ニューラルネットワークに基づく定式化 [16]
    - (iii) 語義の活用 [17]
    - (iv) ドメイン適応 [18], [19]
  - (b) レビューの意見構造の抽出
    - (i) 特徴 (属性) の抽出 [20], [21]
    - (ii) 対象, 意見表現の抽出 [22]
  - (c) 意見要約
    - (i) 根拠文の抽出 [23]
    - (ii) N-gram 単語列の抽出 [24]
  - (d) アプリケーション
    - (i) 偽装意見スパムの判定 [25], [26]
    - (ii) 協調フィルタリングへの応用 [27]
- (3) ブログを主な対象としたもの (2 件)
  - (a) 意見性、極性 (肯定, 否定, 中立), 意見の強弱の分類
    - (i) ニューラルネットワークに基づく半教師あり学習 [28]
  - (b) 言語モデルに基づくブログフィード検索 [29]
- (4) マイクロブログを対象としたもの (5 件)
  - (a) 意見性、極性 (肯定, 否定, 中立), 意見の強弱の分類
    - (i) マイクロブログ特有の表現の利用 [30]
    - (ii) 対象, 関連ツイートの利用 [31]
    - (iii) ハッシュタグの極性分類 [32]
  - (b) 顔文字やハッシュタグを利用した感情の分類 [33]
  - (c) ストリームデータの意見分析 [34]
- (5) その他 (5 件)
  - (a) 会話データを対象とした意見要約 [35]
  - (b) コミュニティ QA データを対象とした意見分析 [36]
  - (c) コミュニティ QA データを利用した辞書の構築 [37]

(d) SNS におけるクイズを利用した辞書の構築 [38]

(e) 色と感情を対応付ける辞書の構築 [39]

これらの研究から、意見分析についての最近の研究では、ニュース文書における観点の分類や、レビュー文書を対象とした偽装意見スパムの発見など、従来の文書ジャンルを対象とした新たな取り組みも進んでいる一方で、マイクロブログやコミュニティ QA など、新しいデータへの取り組みも増えていることがわかる。また、従来の意見分析研究で弱点とされてきた、コーパス作成のアノテーションのコストを軽減することや、ドメインに依存しない意見分析への取り組みも進んでいる。3 節では、上記の研究などでも利用されている、意見分析研究で利用されることの多い代表的な意見分析コーパスについて紹介する。

### 3. 意見分析コーパスの事例紹介

本節では、最近の研究でも用いられることの多い利用可能な意見分析コーパスについて紹介する。

#### 3.1 ニュース文書を主な対象とした意見分析コーパス

##### (1) MPQA 意見コーパス

最初に、ニュース文書を主な対象とした代表的な意見分析コーパスとして、MPQA (*Multi-Perspective Question Answering*) 意見コーパス\*3を紹介する。

本コーパスは、2002 年に、*Advanced Research and Development Activity (ARDA, 現, Disruptive Technology Office)* の *Northeast Regional Research Center (NRRC)* が主催した夏のワークショップにおいて、従来の質問応答システムとは異なる多観点の質問応答を対象としたコーパスを開発したことに端を発する [7]。この時点で、複数の研究グループに所属する 10 人の研究者の協力により、8 つのトピック (のちに、10 トピック) に関連する *World News Connection (WNC)* \*4 の 575 の記事 (Ver1.2 で、535 記事に選別) を対象としたコーパスが作成された。このコーパスは、Version 2.0 で、692 文書に拡張されており、その際には、*Wall Street Journal, American National Corpus, Enron* 社破たんに関する社員の電子メール、話し言葉の書き起こし、アラブ言語の翻訳などの文書が追加されている。この研究の貢献の一つは、意見情報のアノテーションのフレームワークを、多数の判定者による実験を通して厳密に定めた点にある [40]。アノテーションに際しては、シェフィールド大学による *GATE* システム\*5を使用している。意見情報については、以下の 2 つのタイプをもとに判定を行っている。なお、“心理状態 (*private states*)” とは、“ある経験をした人が、(ある

\*2 2012 年 4 月時点で最新の、自然言語処理・情報検索のトップコンファレンスに限定している。ショートペーパーは除外している。

\*3 <http://www.cs.pitt.edu/mpqa/>

\*4 <http://wnc.fedworld.gov/>

\*5 <http://gate.ac.uk/>

対象に対して、)ある態度を保持している状態”と定義される [40].

(a) 心理状態の明示的な言及 (direct subjective)

- 裁判官は、和解による早期解決を望んでいた.
- 財務省幹部は「今回の米国経済のマヒの影響は計り知れない」という.

このタイプの意見表現の、属性としては、以下のものが提案されている.

- (i) 意見のソース (意見を表明する人またはグループ)
- (ii) 意見の対象 (心理状態の対象または話題)
- (iii) 意見の強さ (弱い, 通常, 強い, 強烈)
- (iv) 意見の態度 (肯定, 否定, それ以外, なし)
- (v) 発話の暗黙性
- (vi) 条件節等に含まれる意見の不明瞭さ

(b) 主観に関連した手がかり表現の集合 (expressive subjectivity)

- 今後の情勢について専門家は、強権的なイメージが強く、国民に不評の新国王が、今後、どういう姿勢を打ち出すのか、見極めようとしている.

このタイプの意見表現の、属性としては、以下のものが提案されている.

- (i) 意見のソース (意見を表明する人またはグループ)
- (ii) 意見の強さ (弱い, 通常, 強い, 強烈)
- (iii) 意見の態度 (肯定, 否定, それ以外, なし)

また、もう一つの貢献は、これらのアノテーションをもとに意見を集約するなどのアプリケーションを提案することにあった [41]. その他、意見のソース (意見保有者) の判定 [42] や、文脈を考慮した極性判定 [43], 質問応答 [44] などの研究も進んでいる. 最近では、提案手法の効果を検証するデータセット [9], [12], [28], あるいはラベル付きデータのリソース [11] として利用されている.

(2) NTCIR 多言語意見分析コーパス

多言語を対象とした意見分析コーパスとして、NTCIR 多言語意見分析コーパス<sup>\*6</sup>[45], [46], [47] を紹介する. このコーパスは、日本語・英語・中国語の3か国語 (ただし、中国語は繁体字と簡体字の2種類の言語を対象とする) のニュース記事を対象として、共通のトピックについて集めた記事を対象としたコーパスを作成している. タスクは3回行われており、その変遷を表1に示す. 表中の+は新しい言語あるいはサブタスクが追加されたことを、-は変化がないことを示す.

ここでは、最新の NTCIR-8 のサブタスクについて説

明する. NTCIR-8 では、以下に示す5つの要素技術サブタスクと、1つの応用サブタスクが行われた.

(a) 意見文判定サブタスク

このサブタスクは、MPQA 意見コーパスと同様に、[40] の定義に基づき人手で判定を行ったものを正解として、自動抽出の精度を評価する.

(b) 適合意見文判定サブタスク

このサブタスクは、ある話題に適合する意見文について、自動抽出の精度を評価する.

(c) 極性分類サブタスク

このサブタスクは、1つの意見文中の極性が異なる節を分割することで、意見節を定義し、その節を、肯定意見、否定意見、中立意見のいずれかに分類する精度を評価する.

(d) 意見保有者判別サブタスク

意見を表明する人・組織・団体を判別する精度を評価する.

(e) 意見対象判別サブタスク

意見の対象・トピックを判別する精度を評価する.

(f) 言語横断意見質問応答サブタスク

英語の質問に対して、回答意見を異なる言語の文書集合から検索する精度を評価する.

これらのサブタスクの遂行のため、NTCIR-8 のサブタスクでは、意見を問う20の質問を用意し、毎日新聞 (日本語), *New York Times* (英語), *United Daily News* (繁体字中国語), *Xinhua Chinese* (簡体字中国語) の文書から回答となる意見を集約するため、サブタスクに関係する意見性、意見適合性、極性、意見保有者、意見対象、回答性をアノテートしたコーパスを作成した. コーパスのサイズを表2に示す.

最近では、これらのデータは、提案手法の有効性の評価 [10] や、複数言語のリソースを活用した意見分析 [11] などの研究に用いられている.

- (3) 上記2つ以外で、ニュース文書を対象とした意見分析コーパスとしては、*SemEval-2007 Task 14: Affective Text* [48]<sup>\*7</sup>や、*Sentiment-annotated quotation set* [49], [50]<sup>\*8</sup>がある. 前者は、1,250のニュースサイトおよび新聞記事から収集した見出しについて、6つの感情ラベルと2つの極性ラベルを付与している. 後者は、*European Commission* の *Joint Research Centre* が開発した *Europe Media Monitor (EMM)* [51]<sup>\*9</sup> でクロールしたニュース記事のうち、意見の保有者と対象が明確な1,590の引用意見 (quotation) について、意見情報をアノテートした結果を公開している.

<sup>\*7</sup> <http://www.cs.unt.edu/~rada/affectivetext>

<sup>\*8</sup> [http://langtech.jrc.ec.europa.eu/JRC\\_Resources.html](http://langtech.jrc.ec.europa.eu/JRC_Resources.html)

<sup>\*9</sup> <http://emm.newsbrief.eu/overview.html>

<sup>\*6</sup> <http://research.nii.ac.jp/ntcir/data/data-ja.html>

表 1 NTCIR-6, 7, 8 における多言語意見分析タスクの変遷

NTCIR	言語	サブタスク	判定単位	想定アプリケーション	コーパス	新聞の期間
NTCIR-6	英語, 日本語, 繁体字中国語	意見文判定 適合文判定 極性判定 意見保有者判別	文	情報検索	毎日新聞, 読売新聞 Xinhua English, Hong Kong Standar, CIRB 等	1998-2001
NTCIR-7	+簡体字中国語	+意見対象判別	意見節	質問応答	+Xinhua Chinese	-
NTCIR-8	-	+言語横断意見質問応答	-	意見質問応答	+NYT, UDN	2002-2005

表 2 NTCIR-8 MOAT のテストコレクションのサイズ

言語	トピック数 (意見質問数)			文書数			文数			意見節数		
	合計	サンプル	テストデータ	合計	サンプル	テストデータ	合計	サンプル	テストデータ	合計	サンプル	テストデータ
日本語	21	1	20	178	8	170	6,992	322	6,670	7,894	341	7,553
英語	21	1	20	150	12	138	6,564	399	6,165	6,723	500	6,223
簡体字中国語	20	1	19	410	25	385	4,720	228	4,492	4,744	232	4,512
繁体字中国語	21	1	20	787	12	775	9,684	160	9,524	N/A	N/A	N/A

表 3 TREC 2006/2008 ブログテストコレクションのサイズ

サイズ	Blogs06	Blogs08
重複のないブログの数	100,649	1,303,520
クロールの期間	2005年12月6日～2006年2月21日	2008年1月14日～2009年2月10日
パーマネントリンクの数	3,215,171	28,488,766

### 3.2 レビュー文書を対象とした意見分析コーパス

本節では、レビュー文書を対象とした4つのコーパスを中心に紹介する。

#### (1) 映画レビューを対象とした意見分析データセット

まず、映画レビューの意見分析のためのデータセット<sup>\*10</sup>を紹介する。本コーパスは、*Internet Movie Database (IMDb)*<sup>\*11</sup>を対象とした意見分析コーパスである[5]。具体的には、肯定・否定それぞれ1,000のレビュー文書[52]や、肯定・否定それぞれ5,331のレビュー文[53]を提供している。その他、レビューの強さ[53]のデータセットや、5,000ずつの主観文・客観文[52]のデータセットも提供している。最近では、提案手法の効果を検証するデータセット[15], [28]として、利用されている。

#### (2) Amazon レビューを対象とした製品の属性に基づく要約データセット

次に、Amazon レビューを対象としたデータセットとして、製品の属性に基づく要約 (Feature-based Summarization)<sup>\*12</sup>を紹介する。本コーパスは、最初に5つの製品[54]、のちに9つの製品を追加して[55]、Amazon のレビュー文に、その文に現れる製品の属性と、それに対する評価値を+ (肯定) / - (否定) のかたちで付与している。その他、意見スパム (偽装レビュー) [56]や、製品を比較する意見文[57], [58]のアノテーションデータも公開している。最近では、アプリケーションのリソース[23], [26]として用いられている。

#### (3) 複数のドメインを対象とした意見分析データセット

また、Amazon レビューの複数のドメインを対象としたデータセット<sup>\*13</sup>を紹介する[59]。このデータセットでは、書籍、DVD、電気製品、台所用品の4つのドメインに関する製品データを取り扱っている。それぞれのドメインでは、1,000の肯定・否定のラベルが付いたレビュー文書と、ラベルの付いていない、より多くのレビュー文書が含まれている。最近では、ドメイン適応の検証[18], [19]や、提案手法の有効性の評価[16]に用いられている。

#### (4) 楽天データセット

楽天技術研究所では、楽天市場の約5,000万件の商品についての約1,660万件のレビュー、楽天トラベルの11,468件のホテル・旅館についての約34万件のレビュー、楽天GORAの1,669件のゴルフ場の約32万件のレビューを、研究利用目的のデータセットと

して、国立情報学研究所<sup>\*14</sup>および高度言語情報融合フォーラム (ALAGIN)<sup>\*15</sup> から公開している。また、そのデータを利用して、2008年<sup>\*16</sup>、2009年<sup>\*17</sup>、2010年<sup>\*18</sup>、2011年<sup>\*19</sup>と4回にわたり、研究開発シンポジウムを開催している。

### 3.3 ブログを対象とした意見分析コーパス

本節では、まず、TREC ブログコーパスと関連研究について紹介し、最後にその他のコーパスについて紹介する。

#### (1) TREC 2006/2008 ブログコーパス

情報検索に関する国際評価会議 TREC (*Text REtrieval Conference*) において、2006年から2010年にわたり、ブログトラック[60]が開催された。このためのテストコレクションが、グラスゴー大学より配布されている<sup>\*20</sup>。テストコレクションは、2006年と2008年に作成したものの2種類があり、その内容を表3に示す。このデータを用いて、意見発見 (opinion finding) タスク、ブログ抽出タスク、トップニュース検出タスクなどが行われている。意見発見タスクでは、ある対象 (人名、地名、製品名など) に対する意見を表明するブログポストの検索技術を評価しており、従来の情報検索に基づく適合性 (relevance) と、意見性 (opinionatedness) に基づくランキングを評価している。本タスクでは、3年間の評価により、150のトピックに適合する175,818のブログポストに対して、人手評価が提供されている。本タスクのデータを用いて、最近では、ブログフィード (同一サイトに属するブログポストの集合) についての意見性判定モデル[29]の提案が行われている。

#### (2) TAC 2008 意見質問応答・要約タスク

TAC (*Text Analysis Conference*) では、2008年度に、意見質問応答タスクならびに意見要約タスクが開催された[61]<sup>\*21</sup>。このタスクでは、TREC Blog Track で使用されたトピックを参照しつつ、ある対象 (製品名、人名、組織名等) について書かれた意見を問う質問に対して、その回答を TREC 2006 ブログテストコレクションから抽出し (意見質問応答タスク)、回答を集約して流暢なテキストを生成する技術 (意見要約タスク) を評価している。意見質問応答タスクでは、50の対象に対して、90の厳密な回答を持つリスト型質問 (rigid list question) と87の緩い回答を持つリスト型質問 (squishy list question) に対する回答抽出技術を

<sup>\*10</sup> <http://www.cs.cornell.edu/people/pabo/movie-review-data>

<sup>\*11</sup> 当初は、ニュースグループ rec.arts.movies.reviews のアーカイブ、最近では、Amazon.com により運営されている。  
<http://www.imdb.com/>

<sup>\*12</sup> <http://www.cs.uic.edu/~liub/FBS/sentiment-analysis.html>

<sup>\*13</sup> <http://www.cs.jhu.edu/~mdredze/datasets/sentiment/>

<sup>\*14</sup> <http://www.nii.ac.jp/cscenter/idr/rakuten/rakuten.html>

<sup>\*15</sup> <http://alaginrc.nict.go.jp/>

<sup>\*16</sup> <http://rit.rakuten.co.jp/conf/rrds2008/>

<sup>\*17</sup> <http://rit.rakuten.co.jp/conf/rrds2009/>

<sup>\*18</sup> <http://rit.rakuten.co.jp/conf/rrds3/>

<sup>\*19</sup> <http://rit.rakuten.co.jp/conf/rrds4/>

<sup>\*20</sup> [http://ir.dcs.gla.ac.uk/test\\_collections/](http://ir.dcs.gla.ac.uk/test_collections/)

<sup>\*21</sup> <http://www.nist.gov/tac/data/past/2008/OpSummQA08.html>

評価している。前者は文字列の完全一致を、後者はナゲット [62] と呼ばれる情報単位を回答の単位として定義し、概念的な一致をもとに評価を行う。意見要約タスクでは、意見質問応答タスクのうち、22 の対象についての緩い回答を持つリスト型質問を選択し、要約を生成する技術を、内容、読みやすさ、回答性の観点から評価している。

### (3) TREC 2011 マイクロブログコーパス

TREC では、2011 年度よりマイクロブログに焦点を当てたトラック<sup>\*22</sup>を開始した。このためのテストコレクションは、NIST<sup>\*23</sup>から配布されている。このデータを用いて、リアルタイム検索に焦点を当てたタスクが行われている。

### (4) その他のコーパス

日本語については、京都大学と NTT が協力して、4 つのトピックを対象とした解析済みブログコーパス<sup>\*24</sup>を公開している。また、京都観光ブログデータ (1,041 記事) が、高度言語情報融合フォーラム (ALAGIN)<sup>\*25</sup>から公開されている。

## 4. おわりに

本稿では、最近の意見分析に関する研究を対象として、コーパスの利用状況を明らかにした。また、言語処理研究で、利用されることの多い意見分析コーパスを、文書ジャンル別に紹介した。今後、SNS、マイクロブログ、コミュニティ QA など、利用者が意見を発信するメディアが発展するに伴い、意見分析研究およびそのためのコーパス開発の重要性は引き続き増していくと考えられる。

謝辞 本研究の一部は、科学研究費補助金基盤研究 C (課題番号 24500291) ならびに筑波大学図書館情報メディア系プロジェクト研究の助成を受けて遂行された。

## 参考文献

- [1] Pang, B. and Lee, L.: *Opinion Mining and Sentiment Analysis*, Now Publishers Inc (2008).
- [2] Liu, B.: Sentiment Analysis and Subjectivity, *Handbook of Natural Language Processing* (Indurkha, N. and Damerau, F. J., eds.), CRC Press, pp. 627-664 (2010).
- [3] 大塚裕子, 乾 孝司, 奥村 学: 意見分析エンジン, コロナ社 (2007).
- [4] 乾 孝司, 奥村 学: テキストを対象とした評価情報の分析に関する研究動向, *自然言語処理*, Vol. 13, No. 3, pp. 201-241 (2006).
- [5] Pang, B., Lee, L. and Vaithyanathan, S.: Thumbs up? Sentiment Classification using Machine Learning Techniques, *Proc. of the 2002 Conf. on Empirical Methods in Natural Language Processing (EMNLP 2011)*, pp. 79-86 (2002).
- [6] Turney, P.: Thumbs Up or Thumbs Down? Semantic Orientation Applied to Unsupervised Classification of Reviews, *Proc. of the 40th Annual Meeting on Association for Computational Linguistics*, Philadelphia, USA, pp. 417-424 (2002).
- [7] Wiebe, J., Breck, E., Buckley, C., Cardie, C., Davis, P., Fraser, B., Litman, D., Pierce, D., Riloff, E. and Wilson, T.: NRRRC Summer Workshop on Multiple-Perspective Question Answering Final Report (2002).
- [8] Shanahan, J. G., Qu, Y. and Wiebe, J. M.: *Computing Attitude and Affect in Text: Theory and Applications*, Springer-Verlag (2006).
- [9] Yessenalina, A. and Cardie, C.: Compositional Matrix-Space Models for Sentiment Analysis, *Proc. of the 2011 Conf. on Empirical Methods in Natural Language Processing (EMNLP 2011)*, Edinburgh, Scotland, UK, pp. 172-182 (2011).
- [10] Zhou, L., Li, B., Gao, W., Wei, Z. and Wong, K.-F.: Unsupervised Discovery of Discourse Relations for Eliminating Intra-sentence Polarity Ambiguities, *Proc. of the 2011 Conf. on Empirical Methods in Natural Language Processing (EMNLP 2011)*, Edinburgh, Scotland, UK, pp. 162-171 (2011).
- [11] Lu, B., Tan, C., Cardie, C. and Tsou, B. K.: Joint Bilingual Sentiment Classification with Unlabeled Parallel Corpora, *Proc. of the 49th Annual Meeting of the Association for Computational Linguistics (ACL 2011)*, Portland, Oregon, pp. 320-330 (2011).
- [12] Wiegand, M. and Klakow, D.: Generalization Methods for In-Domain and Cross-Domain Opinion Holder Extraction, *Proc. of the 13th Conf. of the European Chapter of the Association for Computational Linguistics (EACL 2012)*, Avignon, France, pp. 325-335 (2012).
- [13] Park, S., Lee, K. S. and Song, J.: Contrasting Opposing Views of News Articles on Contentious Issues, *Proc. of the 49th Annual Meeting of the Association for Computational Linguistics (ACL 2011)*, Portland, Oregon, pp. 340-349 (2011).
- [14] Fang, Y., Si, L., Somasundaram, N. and Yu, Z.: Mining Contrastive Opinions on Political Texts using Cross-Perspective Topic Model, *Proc. of the Fifth International Conf. on Web Search and Web Data Mining (WSDM 2012)*, Seattle, WA, USA, pp. 63-72 (2012).
- [15] Maas, A. L., E. Daly, R., Pham, P. T., Huang, D., Ng, A. Y. and Potts, C.: Learning Word Vectors for Sentiment Analysis, *Proc. of the 49th Annual Meeting of the Association for Computational Linguistics (ACL 2011)*, Portland, Oregon, pp. 142-150 (2011).
- [16] Beshpalov, D., Bai, B., Qi, Y. and Shokoufandeh, A.: Sentiment Classification Based on Supervised Latent n-gram Analysis, *Proc. of the 20th ACM Conf. on Information and Knowledge Management (CIKM 2011)*, Glasgow, Scotland, UK, pp. 375-382 (2011).
- [17] Balamurali, A. R., Joshi, A. and Bhattacharyya, P.: Harnessing WordNet Senses for Supervised Sentiment Classification, *Proc. of the 2011 Conf. on Empirical Methods in Natural Language Processing (EMNLP 2011)*, Edinburgh, Scotland, UK, pp. 1081-1091 (2011).
- [18] He, Y., Lin, C. and Alani, H.: Automatically Extracting Polarity-Bearing Topics for Cross-Domain Sentiment Classification, *Proc. of the 49th Annual Meeting of the Association for Computational Linguistics (ACL 2011)*, Portland, Oregon, pp. 123-131 (2011).
- [19] Bollegala, D., Weir, D. and Carroll, J.: Using Multiple Sources to Construct a Sentiment Sensitive Thesaurus for Cross-Domain Sentiment Classification, *Proc. of the 49th Annual Meeting of the Association for Computational Linguistics (ACL 2011)*, Portland, Oregon, pp. 132-141 (2011).

<sup>\*22</sup> <http://sites.google.com/site/microblogtrack/home>

<sup>\*23</sup> <http://trec.nist.gov/data/tweets/>

<sup>\*24</sup> <http://nlp.ist.i.kyoto-u.ac.jp/kuntt/>

<sup>\*25</sup> <http://alaginrc.nict.go.jp/>



- 8 - A Step Toward Cross Lingual Opinion Analysis, *Proc. of the Eighth NTCIR Workshop Meeting*, NII, Japan, pp. 209–220 (2010).
- [46] Seki, Y., Evans, D. K., Ku, L. W., Sun, L., Chen, H. H. and Kando, N.: Overview of Multilingual Opinion Analysis Task at NTCIR-7, *Proc. of the Seventh NTCIR Workshop Meeting*, NII, Japan, pp. 185–203 (2008).
- [47] Seki, Y., Evans, D. K., Ku, L. W., Chen, H. H., Kando, N. and Lin, C. Y.: Overview of Opinion Analysis Pilot Task at NTCIR-6, *Proc. of the Sixth NTCIR Workshop Meeting*, NII, Japan, pp. 265–278 (2007).
- [48] Strapparava, C. and Mihalcea, R.: SemEval-2007 Task 14: Affective Text, *Proc. of the 4th International Workshop on Semantic Evaluations (SemEval-2007)*, Prague, pp. 70–74 (2007).
- [49] Balahur, A. and Steinberger, R.: Rethinking Sentiment Analysis in the News: from Theory to Practice and Back, *Proc. of Workshop on Opinion Mining and Sentiment Analysis' (WOMSA)*, Sevilla, Spain, pp. 1–12 (2009).
- [50] Balahur, A., Steinberger, R., Kabadjov, M., Zavarella, V., van der Goot, E., Halkia, M., Pouliquen, B. and Belyaeva, J.: Sentiment Analysis in the News, *Proc. of the Seventh International Conf. on Language Resources and Evaluation (LREC 2010)*, Valletta, Malta, pp. 2216–2220 (2010).
- [51] Steinberger, R., Pouliquen, B. and van der Goo, E.: An introduction to the Europe Media Monitor, *Proc. of ACM SIGIR 2009 Workshop: Information Access in a Multilingual World*, Boston, MA, USA (2009).
- [52] Pang, B. and Lee, L.: A Sentimental Education: Sentiment Analysis Using Subjectivity Summarization Based on Minimum Cuts, *Proc. of the 42nd Ann. Meeting of the Assoc. for Computational Linguistics (ACL2004)*, Barcelona, Spain, pp. 271–278 (2004).
- [53] Pang, B. and Lee, L.: Seeing stars: Exploiting class relationships for sentiment categorization with respect to rating scales, *Proc. of the 43rd Annual Meeting of the Association for Computational Linguistics (ACL 2005)*, Ann Arbor, Michigan, pp. 115–124 (2005).
- [54] Hu, M. and Liu, B.: Mining and summarizing customer reviews, *Proc. of the ACM SIGKDD International Conf. on Knowledge Discovery and Data Mining (KDD-2004)*, Seattle, Washington, USA (2004).
- [55] Ding, X., Liu, B. and Yu, P. S.: A Holistic Lexicon-Based Approach to Opinion Mining, *Proc. of First ACM International Conf. on Web Search and Data Mining (WSDM-2008)*, Stanford University, California, USA (2008).
- [56] Jindal, N. and Liu, B.: Opinion Spam and Analysis, *Proc. of First ACM International Conf. on Web Search and Data Mining (WSDM-2008)*, Stanford University, California, USA (2008).
- [57] Jindal, N. and Liu, B.: Identifying Comparative Sentences in Text Documents, *Proc. of the 29th Annual International ACM SIGIR Conf. on Research and Development on Information Retrieval (SIGIR-06)*, Seattle, USA (2006).
- [58] Ganapathibhotla, M. and Liu, B.: Mining Opinions in Comparative Sentences, *Proc. of the 22nd International Conf. on Computational Linguistics (Coling-2008)*, Manchester (2008).
- [59] Blitzer, J., Dredze, M. and Pereira, F.: Biographies, Bollywood, Boom-boxes and Blenders: Domain Adaptation for Sentiment Classification, *Proc. of the 45th Annual Meeting of the Association of Computational Linguistics (ACL 2007)*, Prague, Czech Republic, pp. 440–447 (2007).
- [60] Macdonald, C., Santos, R. L., Ounis, I. and Soboroff, I.: Blog Track Research at TREC, *ACM SIGIR Forum*, Vol. 44, No. 1, pp. 58–75 (2010).
- [61] Dang, H. T.: Overview of the TAC 2008 Opinion Question Answering and Summarization Tasks, *Proc. of Text Analysis Conf. 2008 (TAC 2008)*, pp. 24–35 (2008).
- [62] Lin, J. and Demner-Fushman, D.: Will Pyramids Built of Nuggets Topple Over?, *Proc. of the 2006 Human Language Technology Conf. and the North American Chapter of the Association for Computational Linguistics Annual Meeting (HLT/NAACL 2006)*, New York City, USA, pp. 383–390 (2006).