

# Aspect based Sentiment Analysis in Hindi: Resource Creation and Evaluation

Md Shad Akhtar, Asif Ekbal and Pushpak Bhattacharyya

Department of Computer Science and Engineering

Indian Institute of Technology Patna

Patna, India

{shad.pcs15,asif,pb}@iitp.ac.in

## Abstract

Due to the phenomenal growth of online product reviews, sentiment analysis (SA) has gained huge attention, for example, by online service providers. A number of benchmark datasets for a wide range of domains have been made available for sentiment analysis, especially in resource-rich languages. In this paper we assess the challenges of SA in Hindi by providing a benchmark setup, where we create an annotated dataset of high quality, build machine learning models for sentiment analysis in order to show the effective usage of the dataset, and finally make the resource available to the community for further advancement of research. The dataset comprises of Hindi product reviews crawled from various online sources. Each sentence of the review is annotated with aspect term and its associated sentiment. As classification algorithms we use Conditional Random Field (CRF) and Support Vector Machine (SVM) for aspect term extraction and sentiment analysis, respectively. Evaluation results show the average F-measure of 41.07% for aspect term extraction and accuracy of 54.05% for sentiment classification.

**Keywords:** Aspect based Sentiment Analysis, Sentiment Dataset, Indian Languages.

## 1. Introduction

With the globalization of Internet, web generated contents are increasing at a tremendous pace. This huge amount of data has introduced several new challenges and opportunities in the research communities. These days customers or users are relying heavily on other user's opinion about a product or service before experiencing themselves. In order to get an unbiased opinion one has to extract and read all the reviews which is not an easy task to perform. Sentiment Analysis (Pang and Lee, 2008) refers to the problem of automatically determining the polarity of sentiment/opinion expressed by user in a chunk of text or review. In general, polarity of a review belongs to one of the four possible classes: positive, negative, neutral and conflict. Aspect based sentiment analysis (ABSA) is a fine-grained analysis of sentiments at the aspect or feature or attribute level. The term 'aspect' refers to an attribute or a component of the product/service that has been commented on in a review. Overall problem of aspect based sentiment analysis can be thought as a two-step process. The first step, i.e., aspect term extraction focuses on identifying various terms that denote aspects, and the second step, i.e. sentiment classification deals with classifying the sentiments with respect to the aspect. A review sentence, therefore, may contain more than one aspect term and the sentiment associated with each. Such a fine-grained analysis provides greater insight to the sentiments expressed in the written reviews. In recent times, there have been a growing trend for sentiment analysis at the more fine-grained level, i.e. for aspect based sentiment analysis (ABSA). Some of the recent systems that have emerged are (Toh and Wang, 2014; Chernyshevich, 2014; Wagner et al., 2014; Castellucci et al., 2014; Gupta et al., 2015). However, almost all these research are related to some specific languages, especially the English.

Sentiment analysis in Indian languages are still largely unexplored due to the non-availability of various resources and tools such as annotated corpora, lexicons, Part-of-Speech (PoS) tagger etc. Existing works (Joshi et al., 2010; Balamurali et al., 2012; Balamurali et al., 2011; Bakliwal et al., 2012; Mittal et al., 2013; Sharma et al., 2014; Das and Bandyopadhyay, 2010b; Das and Bandyopadhyay, 2010a; Das et al., 2012) involving Indian languages mainly discuss the problems of sentiment analysis at the coarse-grained level with the aims of classifying sentiments either at the sentence or document level.

In this work we describe our research on aspect based sentiment analysis in Hindi. Hindi is the national language in India, and ranks 5th in the world in terms of speaker population. As we already mentioned, the bottleneck for performing sentiment analysis involving Hindi is again due to the non-availability of benchmark datasets and the scarcity of various other resources and tools. Annotated dataset is certainly the foremost requirement for NLP tasks, irrespective of application and domain, and sentiment analysis is no exception. Therefore, a good dataset both in terms of quality and quantity has great impact on the overall system performance. Several benchmark datasets for sentiment analysis for resource-rich languages like English exist and these have been made freely available for research, e.g., SemEval 2014 datasets (Pontiki et al., 2014). However, Indian languages are still far behind in terms of such resources. Datasets, specific to Indian languages, which were created for the in-house developments by few of the research groups are very few in number, and there are three basic limitations: (i). smaller size (mostly in few 100s) (Joshi et al., 2010) (Balamurali et al., 2012) (Balamurali et al., 2011), (ii). low quality, as the data were generated by translating English reviews using Google translator (Bak-

liwal et al., 2012), and (iii). none of these was meant for aspect based sentiment analysis.

The focus of our work is to provide a benchmark setup for creating a dataset for aspect based sentiment analysis in Hindi, and then developing models for aspect term extraction and sentiment analysis for the effective usage of this dataset. This will surely open avenues for research in sentiment analysis involving Indian languages. To the best of our knowledge ours is the very first initiative, where we make an attempt to provide a benchmark setup for ABSA in Indian languages. Few of the examples are shown in Table 1 that gives an idea of what kind of datasets we have created. Three review text are listed in Devanagri script along with their Roman transliteration<sup>1</sup> and English translation form. The first review has one multi-word aspect term while the other two has one single-word aspect term. Sentiment towards respective aspect terms are listed in the last column of the table.

For the effective usage of the dataset we develop models for aspect term extraction and sentiment classification. We use Conditional Random Field (CRF) (Lafferty et al., 2001) as our learning algorithm for the aspect term extraction task and Support Vector Machine (SVM) (Cortes and Vapnik, 1995) for sentiment classification task. Evaluation shows the overall precision, recall and F-measure values of 61.96%, 30.72% and 41.07%, respectively for aspect term extraction and an accuracy of 54.05% for sentiment classification.

The rest of the paper is organised as follows: Various aspects of resource creation and its challenges are discussed in Section 2.1.. In Section 3., we describe a brief overview of aspect based sentiment analysis task. Experiments and evaluation results are furnished in Section 4.. Finally, we conclude in Section 5.

## 2. Benchmark Setup for ABSA in Hindi

To address the challenges as we pointed out earlier, we design web crawler, collect raw data, clean and annotate user generated reviews. In subsequent subsections we discuss these processes.

### 2.1. Data Crawling

We design a web crawler that downloads product reviews from various online sources. We have crawled more than 100 newspapers, blogs, e-commerce websites etc. <sup>2</sup> Following this process we have collected a total of 8,000 re-

<sup>1</sup>We use ITRANS tool for the transliteration

<sup>2</sup>List of few sources..

<http://www.jagran.com>

<http://www.gizbot.com>

<http://www.patrika.com>

<http://www.hi.themobileindian.com>

<http://www.mobilehindi.com>

<http://navbharattimes.indiatimes.com>

<http://hindi.starlive24.in/>

<http://www.amarujala.com>

<http://techjankari.blogspot.in>

<http://www.digit.in>

<http://khabar.ndtv.com/topic>

<http://www.hindi.mymobile.co.in/>

<http://www.bhaskar.com>

view sentences covering 12 domains. The set of domains comprises *Laptops*, *Mobiles*, *Tablets*, *Cameras*, *Home-appliances*, *Mobile apps*, *Smart watches*, *Headphones*, *Speakers*, *Television*, *Travel* and *Movies*. Some statistics of this dataset are presented in Section 2.4..

### 2.2. Pre-processing

We pre-process the crawled data to convert it into the desired forms. We perform the following steps to prepare the datasets for our use:

- At first, we identified and removed many irrelevant reviews from the dataset. This was done semi-automatically to ensure that the data we use finally was suitable for sentiment analysis.
- From the remaining reviews we dropped off many unprintable characters and various emoticons like ‘:)’ (smiley), ‘(Y)’ (thumbs up) etc.
- We scanned the reviews and corrected few of the obvious spelling mistakes by adding and/or deleting few characters or words.
- Many mismatched pairs (braces, quotes etc.) were corrected.
- Some sentences had missing sentence end marker (.), and hence we appended it.

A few examples of pre-processing are listed in Table 2.

### 2.3. Data Annotation

For annotation of dataset, we follow the guidelines, which are in line with SemEval 2014 (Pontiki et al., 2014) shared task. We convert datasets in xml form. Each aspect term that appears in a review sentence is annotated along with its associated sentiment. The sentiment is classified into four categories, viz., *positive*, *negative*, *neutral* and *conflict*.

Two instances of the dataset along with their xml structure are presented in Table 3. At the top of the table, two example sentences are given in its original form, i.e., in Devanagri script along with its Roman transliteration and English translation. The first review contains only one aspect term and its polarity is positive. However, the second review does not have any aspect term. The other half of Table 3 defines the annotation structure in xml format. The `<sentences>` node represents root node of the xml that contains every sentence of the review as its children i.e. `<sentence>`. To uniquely identify each `<sentence>`, an ‘*id*’ is associated with it as an attribute. Each `<sentence>` node has two children, namely `<text>` and `<aspectTerms>`. The `<text>` node holds one review sentence, whereas `<aspectTerms>` contains *n* `<aspectTerm>` nodes as its children if a review sentence has *n* aspect terms. For the example at hand *n* equals to 1 and 0 for the two respective reviews. Each `<aspectTerm>` node holds four attributes: ‘*term*’, ‘*from*’, ‘*to*’ & ‘*polarity*’. Attribute ‘*term*’ defines aspect term represented by current node while ‘*polarity*’ stores the sentiment towards

<http://www.howws.com>

	Review Text	Aspect Term	Polarity
Devanagri	इसकी ऑडियो क्वालिटी शानदार है।		
Transliterated	Isakee AWDiyo kvaaliTee shaanadaara hai.	ऑडियो क्वालिटी	positive
Translated	Its audio quality is superb.		
Devanagri	इस लैपटॉप का वजन 2.38 किलोग्राम है जो भारी है।		
Transliterated	Is laipaTawp kaa vajan 2.38 kilograam hai jo bhaaree hai.	वजन	negative
Translated	Weight of this laptop is 2.38 kilogram which is heavy.		
Devanagri	इसकी स्क्रीन 15.6 इंच की है।		
Transliterated	Isakee skreen 15.6 INch kee hai.	स्क्रीन	neutral
Translated	Its has 15.6 inch screen.		

Table 1: Example of Aspect Based Sentiment Analysis.

	Review Text
Original (Devanagari)	स्क्रीन का रिज़ोल्यूशन 1024 गुणा 600 है, जो काफी अच्छ है
Original (Transliterated)	skreen kaa riZolyooshan 1024 guNNaa 600 hai , jo kaaphee Achchh hai
Corrected	स्क्रीन का रिज़ोल्यूशन 1024 गुणा 600 है , जो काफी अच्छा है।
Corrected	skreen kaa riZolyooshan 1024 guNNaa 600 hai , jo kaaphee Achchhaa hai.
Original (Devanagri)	सिंगल माउस बटन साधारण साउंड क्वालिटी...:(
Original (Transliterated)	siNgal maaUs baTan saadhaaraNN saaUND kvaaliTee...:(
Corrected	सिंगल माउस बटन और साधारण साउंड क्वालिटी है।
Corrected	siNgal maaUs baTan <b>AOr</b> a saadhaaraNN saaUND kvaaliTee <b>hai</b> .

Table 2: Pre-processing examples.

the ‘term’ which is ‘positive’. Position of the aspect term in the review text is determined by attributes ‘from’ and ‘to’ which store the index of first and last character, respectively in the review text.

All the review documents collected are presented to three different annotators, who were native speakers of Hindi language. To check the goodness of annotations by different annotators we calculate inter-rater agreement. Cohen’s Kappa coefficient (Cohen, 1960) is a statistical measure to analyse the inter-rater agreement and defined as

$$K = \frac{Pr(a) - Pr(e)}{1 - Pr(e)} \quad (1)$$

where  $Pr(a)$  &  $Pr(e)$  are the observed and by chance agreement among raters. We perform Cohen’s Kappa coefficient on the datasets and an average agreement of 95.18% was obtained. This shows that the annotated dataset we generated is of acceptable quality. Finally, majority voting based technique was employed to merge three annotated versions of datasets. Agreement/disagreement matrices among annotators for aspect term extraction task are listed in Figure 1.

## 2.4. Dataset Statistics

After pre-processing, our dataset contains 5,417 review sentences across 12 domains. There are a total of 2,290 positive, 712 negative, 2,226 neutral and 189 conflict reviews (sentence-level). Overall it contains 100,279 and 4,509 tokens and aspect terms, respectively. Polarity classification of these aspect terms count to 1,986 positive, 569 negative, 1,914 neutral and 40 conflict sentiments. Overall and domain-wise details of this dataset are reported in Table 4.

## 3. Method for Aspect Term Extraction and Sentiment Analysis

In this section we describe the models that we develop for aspect term extraction and sentiment classification. For both the tasks we identify and implement a set of language independent features which are implemented without using any domain-specific external resources and/or tools.

### 3.1. Aspect Term Extraction

Aspect term extraction task is cast as a sequence learning problem. Each token of the review is marked with the BIO encoding scheme, where B, I and O denote the beginning,

Id	Format	Review Text
1.	Devanagri	इसकी ऑडियो क्वालिटी शानदार है।
	Transliterated	Isakee AWDiyo kvaaliTee shaanadaara hai.
	Translated	Its audio quality is superb.
2.	Devanagri	यह बहुत महंगा है।
	Transliterated	yah bahut mahaNgaai hai.
	Translated	It is very costly.

  

Annotation Structure	
<pre> &lt;sentences&gt;   &lt;sentence id= "1" &gt;     &lt;text&gt; इसकी ऑडियो क्वालिटी शानदार है।&lt; \text&gt;     &lt;aspectTerms&gt;       &lt;aspectTerm from="5" to="18" term="ऑडियो क्वालिटी" polarity="positive" /&gt;     &lt; \aspectTerms&gt;   &lt; \sentence&gt;   &lt;sentence id= "2"&gt;     &lt;text&gt; यह बहुत महंगा है।&lt; \text&gt;   &lt; \sentence&gt;   &lt;sentence id= "3"&gt;     ...   &lt; \sentence&gt; &lt; \sentences&gt; </pre>	

Table 3: Dataset annotation structure.

intermediate and outside entities of aspect terms. For example, in the following aspect term “*Audio Quality*”, the first and second tokens have associated classes ‘B’ and ‘I’, respectively and rest of the tokens corresponds to ‘O’.

Its audio quality is superb.					
इसकी	ऑडियो	क्वालिटी	शानदार	है	।
Isakee	AWDiyo	kvaaliTee	shaanadaara	hai	.
O	B	I	O	O	O

We develop a model based on Conditional Random Field (CRF) (Lafferty et al., 2001) that has shown success in solving similar kinds of problems. The classifier is trained with the following set of features:

- Word & local context:** Surface word and its local context which lies within a context window of size 3 are used as features for training the model.
- Part-of-Speech (PoS):** PoS tag information of the cur-

rent and/or the surrounding tokens are used as features in the proposed method,

- Chunk information:** Chunk information is a prominent feature for identifying the multi-word aspect terms.
- Prefixes and suffixes:** Prefix and suffix strings of surface word are used as the features. They are extracted by stripping off a fixed length character sequences from the beginning or end positions of words.

We have used LTRC shallow parser<sup>3</sup> for tokenization, POS and chunk information.

### 3.2. Sentiment Classification

Once the aspect terms are identified, sentiment analysis is performed to predict the polarity of sentiment expressed towards this aspect term in the given review. We choose Support Vector Machine (SVM) (Cortes and Vapnik, 1995) as the preferred classifier because of its efficiency in solving

<sup>3</sup>[http://ltrc.iiit.ac.in/showfile.php?filename=downloads/shallow\\_parser.php](http://ltrc.iiit.ac.in/showfile.php?filename=downloads/shallow_parser.php)

Domains	# Tokens	# Sentences	Aspect Terms				
			# Pos	# Neg	# Neu	# Con	Total
Laptops	6419	348	185	33	169	1	388
Mobiles	21923	1141	600	210	578	28	1416
Tablets	25323	1244	418	157	479	2	1056
Cameras	3097	150	107	11	64	1	183
Headphones	835	43	20	8	19	0	47
Home appliances	1746	84	10	0	34	0	44
Speakers	726	47	20	3	25	0	48
Smart watches	5709	330	47	22	149	2	220
Televisions	2179	135	41	3	99	1	144
Mobile apps	4577	229	98	20	46	0	164
Travels	14157	776	273	19	98	0	390
Movies	13588	890	167	83	154	5	409
Overall	100279	5417	1986	569	1914	40	4509

Table 4: Dataset statistics. Here, POS-positive, Neg-negative, Neu-Neutral and Con-Conflict

text analytics problems. The classifier is trained with the following features:

1. **Target aspect term and local context:** Sentiment bearing words usually occur closer to the target aspect term. We extract target term along with its preceding and following few tokens, and use as features for training. For the proposed method we fix context window size to 5.
2. **Word Bigrams:** Pair of two consecutive tokens are used as features to capture the co-occurrence behavior of the tokens.
3. **Semantic Orientation (SO):** Semantic Orientation (SO) (Hatzivassiloglou and McKeown, 1997) is a measure of association of a token towards positive or negative sentiments and can be defined as:

$$SO(t) = PMI(t, posRev) - PMI(t, negRev) \quad (2)$$

where  $PMI(t, posRev)$  stands for point-wise mutual information of a token  $t$  towards positive sentiment reviews.

## 4. Experiments and Evaluation

As a base learning algorithm we make use of Conditional Random Field (CRF) and Support Vector Machine (SVM) for the aspect term extraction and sentiment classification tasks respectively. We use CRF++<sup>4</sup> and TinySVM<sup>5</sup> based packages for our experiments.

<sup>4</sup><http://taku910.github.io/crfpp/>

<sup>5</sup><http://chasen.org/taku/software/TinySVM/>

To evaluate the performance of the system, we use the evaluation scripts made available by the SemEval 2014 shared task organizers. We perform 3-fold cross-validation to report the final evaluation results. We obtain average F-measure of 41.07% for aspect term extraction and accuracy of 54.05% for sentiment analysis. We also evaluate the models of aspect term extraction and sentiment analysis for each domain separately. Evaluation results are reported in Table 5. It is evident that the performance is greatly influenced by the amount of reviews available for a particular domain. The only exception is the *Travel* domain which reports merely 15.03% for aspect term extraction despite having relatively good amount of reviews. This could be because *Travel* contains reviews from a diverse set of places e.g. 'religious', 'hill stations', 'beaches', 'monuments' etc. and most of the aspect terms belong to its specific category.

### 4.1. Error Analysis

Confusion matrices for both aspect term extraction and sentiment classification are shown in Figure 2. For aspect term extraction task, confusion matrix suggests that more than 50% of aspect terms are not detected at all, which in turn, reduces the recall value. The classifier faces problem in detecting multi-word aspect terms. Confusion matrix shows that only 25% multi-word aspect terms are correctly predicted. The system of sentiment classification suffers most due to problems associated with the 'conflict' class. Accuracy for 'negative' class is lower as compared to 'positive' and 'neutral'. This could be because of relatively few instances of 'negative' and 'conflict' classes. Prediction performance of conflict

	B-ASP	I-ASP	O
B-ASP	4037	2	290
I-ASP	13	2844	435
O	72	64	92522

(a) Annotator 1 vs Annotator 2

	B-ASP	I-ASP	O
B-ASP	4258	5	66
I-ASP	28	203	93
O	16	28	92614

(b) Annotator 1 vs Annotator 3

	B-ASP	I-ASP	O
B-ASP	4007	14	101
I-ASP	4	2803	103
O	276	402	92567

(c) Annotator 2 vs Annotator 3

Figure 1: Agreement/Disagreement matrix for aspect term extraction task. Here, B-ASP, I-ASP and O denote the beginning, inetermediate and outside of aspect term, respectively.

	B	I	O
B	1210	128	2477
I	142	754	2105
O	1065	887	91511

(a) Aspect Term Extraction

	positive	negative	neutral	conflict
positive	1416	30	540	0
negative	376	51	142	0
neutral	917	27	970	0
conflict	30	4	6	0

(b) Sentiment Classification

Figure 2: Confusion matrix for ABSA.

class may be addressed with more training instances as well as by implementing class-specific features. We also perform qualitative analysis of errors that the system incurs.

### 1. Aspect term extraction:

- We observe that the presence of prepositions and conjunctions inside an aspect entity confuses the

Domain	Aspect term extraction			Sentiment
	Pre	Rec	F-measure	Accuracy
Laptops	74.59	56.87	64.53	50.98
Mobiles	67.48	44.42	53.57	54.07
Tablets	61.50	33.67	43.52	57.19
Cameras	60.0	31.76	41.53	59.06
Headphones	100.0	27.78	43.47	46.15
Home appl.	100.0	16.67	28.57	79.23
Speakers	83.33	22.72	35.71	53.84
Smart watches	50.0	41.50	45.36	64.70
Televisions	75.60	42.46	54.38	65.47
Mobile apps	50.0	18.0	26.47	61.53
Travels	32.60	9.77	15.03	68.78
Movies	70.14	58.02	63.51	39.23
Overall	61.96	30.72	41.07	54.05

Table 5: Result of aspect term extraction and sentiment classification

system to correctly mark its boundry. For example, in the example review given below, “डिस्प्ले की व्यूइंग एंगल और ब्राइटनेस” (Disple kee vyooING ENgal AOra braaITanes) is an aspect term but our system only predicts “डिस्प्ले की व्यूइंग एंगल” (Disple kee vyooING ENgal) as an aspect term and leaves “और ब्राइटनेस” (AOra braaITanes) umarked.

**Review:** “डिवाइस के अनुसार डिस्प्ले की व्यूइंग एंगल और ब्राइटनेस अच्छी और प्रभावित करने वाली है।”

**Transliteration:** “DivaaIs ke Anusaara Disple kee vyooING ENgal AOra braaITanes Achchhee AOra prabhaavit karane vaalee hai.

- When a noun phrase precedes or succeeds an aspect term, the system marks each token of the neighboring phrase as a part of the aspect term.

### 2. Sentiment Classification:

- When the sentiment bearing words occur far away from the aspect term, its sentiment is not correctly captured by the system. In the following sentence “रिज्यूम नोटिफिकेशन” (rijyoom noTiphikeshan) is an aspect term and its polarity is positive. The sentiment bearing word “खास” (khaas) is located 7 words away from the aspect term. Beacuse of this long distance, classifier was not able to capture the sentiment properly.

**Review:** “पेरिस्कोप एप में रिज्यूम नोटिफिकेशन फीचर दिया गया है जो बहुत ही खास है।”

**Transliteration:** “periskop Ep meN rijyoom noTiphikeshan pheechara diyaa gayaa hai jo bahut hee khaas hai.”

It should be noted that, for the tasks, we make use of a very basic language and domain independent features. Performance of both these models can be improved by defining more domain-specific features. Systematic feature engineering might be useful to pickup the best set of features.

## 5. Conclusion

In this paper we propose a benchmark setup for aspect based sentiment analysis in Hindi. We have crawled various online sources, performed pre-processing to clean the data, and annotated the dataset with aspect terms and polarity classes. The dataset comprises of Hindi product reviews crawled from the various online sources across 12 domains. Based on this dataset we build supervised classifiers for aspect term extraction and sentiment classification. Evaluation results on 3-fold cross-validation show the overall precision, recall and F-measure values of 61.96%, 30.72% and 41.07%, respectively for aspect term extraction and an accuracy of 54.05% for sentiment classification. We also make the dataset available to the community for the advancement of further research involving Indian languages.

In future, we would like to investigate domain-specific features for both the tasks. We would also like to explore deep learning methods for aspect based sentiment analysis.

## 6. Acknowledgements

We acknowledge undergraduate students Siddharth Kumar (IEST, Shibpur), Kumari Geeta (BIT Patna) and Pooja Sharma (Manipal University) for contributing in collection and annotation of datasets.

## 7. References

- Bakliwal, A., Arora, P., and Varma, V. (2012). Hindi subjective lexicon: A lexical resource for hindi polarity classification. *Proceedings of the Eight International Conference on Language Resources and Evaluation (LREC)*.
- Balamurali, A. R., Joshi, A., and Bhattacharyya, P. (2011). Harnessing wordnet senses for supervised sentiment classification. In *Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing, EMNLP 2011, 27-31 July 2011, John McIntyre Conference Centre, Edinburgh, UK, A meeting of SIGDAT, a Special Interest Group of the ACL*, pages 1081–1091.
- Balamurali, A. R., Joshi, A., and Bhattacharyya, P. (2012). Cross-lingual sentiment analysis for indian languages using linked wordnets. In *COLING 2012, 24th International Conference on Computational Linguistics, Proceedings of the Conference: Posters, 8-15 December 2012, Mumbai, India*, pages 73–82.
- Castellucci, G., Filice, S., Croce, D., and Basili, R. (2014). Unitor: Aspect based sentiment analysis with structured learning. In *Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014)*, pages 761–767, Dublin, Ireland, August. Association for Computational Linguistics and Dublin City University.
- Chernyshevich, M. (2014). IHS R&D Belarus: Cross-domain Extraction of Product Features using Conditional Random Fields. pages 309–313.
- Cohen, J. (1960). A Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement*, 20(1):37.
- Cortes, C. and Vapnik, V. (1995). Support-vector networks. *Mach. Learn.*, 20(3):273–297, September.
- Das, A. and Bandyopadhyay, S. (2010a). Phrase-level polarity identification for bangla. *Int. J. Comput. Linguist. Appl. (IJCLA)*, 1(1-2):169–182.
- Das, D. and Bandyopadhyay, S. (2010b). Labeling emotion in bengali blog corpus—a fine grained tagging at sentence level. In *Proceedings of the 8th Workshop on Asian Language Resources*, page 47.
- Das, A., Bandyopadhyay, S., and Gambäck, B. (2012). Sentiment analysis: what is the end user’s requirement? In *Proceedings of the 2nd International Conference on Web Intelligence, Mining and Semantics*, page 35. ACM.
- Gupta, D. K., Reddy, K. S., Ekbal, A., et al. (2015). Psoasent: Feature selection using particle swarm optimization for aspect based sentiment analysis. In *Natural Language Processing and Information Systems*, pages 220–233. Springer.
- Hatzivassiloglou, V. and McKeown, K. R. (1997). Predicting the Semantic Orientation of Adjectives. In *Proceedings of the ACL/EACL*, pages 174–181.
- Joshi, A., Balamurali, A., and Bhattacharyya, P. (2010). A fall-back strategy for sentiment analysis in hindi: a case study. *Proceedings of the 8th ICON*.
- Lafferty, J. D., McCallum, A., and Pereira, F. C. N. (2001). Conditional Random Fields: Probabilistic Models for Segmenting and Labeling Sequence Data. In *ICML*, pages 282–289.
- Mittal, N., Agarwal, B., Chouhan, G., Bania, N., and Pareek, P. (2013). Sentiment analysis of hindi review based on negation and discourse relation. In *proceedings of International Joint Conference on Natural Language Processing*, pages 45–50.
- Pang, B. and Lee, L. (2008). Opinion mining and sentiment analysis. *Foundations and trends in information retrieval*, 2(1-2):1–135.
- Pontiki, M., Galanis, D., Pavlopoulos, J., Papageorgiou, H., Androutopoulos, I., and Manandhar, S. (2014). Semeval-2014 task 4: Aspect based sentiment analysis. In *Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014)*, Dublin, Ireland, August.
- Sharma, R., Nigam, S., and Jain, R. (2014). Polarity detection movie reviews in hindi language. *CoRR*, abs/1409.3942.
- Toh, Z. and Wang, W. (2014). Dlirec: Aspect term extraction and term polarity classification system. In *Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014)*, pages 235–240.
- Wagner, J., Arora, P., Cortes, S., Barman, U., Bogdanova, D., Foster, J., and Tounsi, L. (2014). Dcu: Aspect-based polarity classification for semeval task 4. In *Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014)*, pages 223–229.