

# Tagging Opinion Phrases and their Targets in User Generated Textual Reviews

**Narendra Gupta**

AT&T Labs - Research, Inc.  
Florham Park, NJ 07932 - USA  
ngupta@research.att.com

## Abstract

We discuss a tagging scheme to tag data for training information extraction models which can extract the features of a product/service and opinions about them from textual reviews, and which can be used across different domains with minimal adaptation. A simple tagging scheme results in a large number of domain dependent opinion phrases and impedes the usefulness of the trained models across domains. We show that by using minor modifications to this simple tagging scheme the number of domain dependent opinion phrases are reduced from 36% to 17%, which leads to models more useful across domains.

## 1 Introduction

A large number of opinion-rich reviews about most products and services are available on the web. These reviews are often summarized by star ratings to help consumers in making buying decisions. While such a summarization is very useful, often consumers like to know about specific features of the product/service. For example in the case of restaurants consumers might want to know what people think about their chicken dish. There are many research papers on both supervised (Li et al., 2010) and unsupervised (Liu et al., 2012), (Hu and Liu, 2004), (Popescu and Etzioni, 2005), (Baccianella et al., 2009) methods for extracting reviewer's opinions and their targets (features of products/services) from textual reviews. Unsupervised methods are preferred as they can be used across domains, however their performance is limited by the assump-

tions they make about lexical and syntactic properties of opinion and target phrases. We would like to use supervised methods to develop information extraction models that can also be used across domains with minimum adaptation. We hope to succeed in our goal because: a) even though there are domain specific opinion phrases, we believe a large proportion of opinion phrases can be used across the domains with the same semantic interpretation; b) target phrases mostly contain domain dependent words, but have domain independent syntactic relationships with opinion phrases. Obviously for a domain containing large number of domain dependent opinion phrases, our models will perform poorly and additional domain adaptation will be necessary.

In this paper we discuss a tagging scheme to manually tag the necessary training data. In section 2 we show that simply tagging opinion and target phrases, forces a large number of opinion phrases to contain domain dependent vocabulary. This makes them domain dependent, even when domain independent opinion words are used. In section 3 we propose a modification to the simple tagging scheme and show that this modification allows tagging of opinion phrases without forcing them to contain domain dependent vocabulary. We also identify many linguistic structures used to express opinions that cannot be captured even with our modified tagging scheme. In section 4 we experimentally show the improvement in the coverage of tagged domain independent opinion phrases due to our proposed modification. In section 5 we discuss the relationship with other work. We conclude this paper in section 6 by summarizing the contribution of this work.

## 2 A Simple Tagging Scheme

Our goal is to only extract author’s *current opinions* by using the smallest possible representation. *Past opinions* or those of other agents are not of interest.

As shown in Table 1, in this *simple tagging scheme* we tag opinions, their targets, and pronominal references in each sentence without considering the review or the domain the sentence is part of<sup>1</sup>. Opinion phrases are further categorized to represent their polarity and their domain dependence<sup>2</sup>.

There are two relations in this scheme viz.  $Target(Opinionphrase, Target|Referencephrase)$ , and  $Reference(Referencephrase, Targetphrase)$ .

Finally, we tag only the *contiguous non-overlapping spans* in the sentences.

Phrase Type		Domain Dependent	Tag Symbol
Opinion	Positive	No	P
		Yes	PD
	Negative	No	N
		Yes	ND
Neutral	Yes	UD	
Target		Yes	T
Pronominal Reference		No	R

Table 1: Different types of phrases to be tagged.

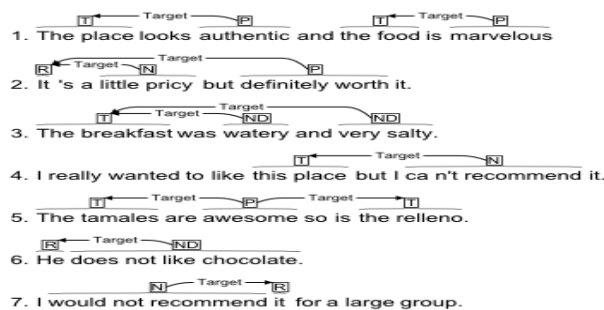


Figure 1: Examples tagged by the simple tagging scheme.

Figure 1<sup>3</sup> shows examples of tagged sentences using the simple tagging scheme. It illustrates: a) a sentence can have multiple “Target” relationships b) pronominal references can be used as targets; c) many opinion phrases can have the same target and

<sup>1</sup>Once context independent information is extracted from individual sentences, post processing (not discussed in this paper) can aggregate information (e.g. resolve all identified pronominal references) within the context of the entire review.

<sup>2</sup>An opinion phrase is domain independent if its interpretation remains unchanged across domains

<sup>3</sup>Figures showing examples annotations are extracted from Brat Annotation Tool (<http://brat.nlpplab.org/>). In Brat tags are displayed as square boxes and relations as directed arcs.

vice versa; d) opinion phrases are not always adjectives and/or adverbs; e) in the sixth sentence “his” opinion about chocolate is not tagged instead, author’s opinion about the opinion holder is tagged; f) in the last sentence fragmented opinion phrase “not recommend for a large group” is not tagged.

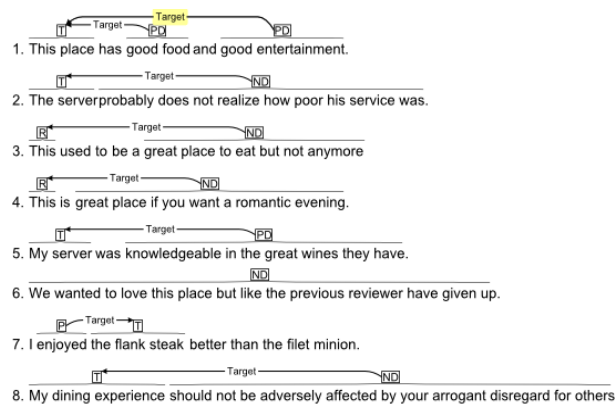


Figure 2: Examples where the simple tagging scheme is not discriminating enough.

Figure 2 shows examples where our simple tagging scheme is not discriminating enough. As a result majority of the opinion phrases are tagged as domain dependent. Example 1, 2 show that the tagging scheme cannot express attributes of a target. Therefore, they are lumped with the opinion phrases, making them domain dependent. In example 5 the opinion about “wines they have” is embedded in the tagged opinion phrase. In example 6 the fact that “we do not love this place” is not captured. Example 7 shows that our scheme can only tag one of the two targets of a comparative opinion expression. Example 8 shows a complex opinion expression involving multiple agents, opinions, expectations, analogies and modalities. To accurately represent opinions expressed in the infinitely many compositions, natural languages offer, a more complex representation is required. Instead of solving this knowledge representation problem, we introduce two additional tags and relations, and show that our modified tagging scheme is able to capture opinions expressed through some commonly used expressions.

## 3 A Modified Tagging Scheme

In our modified tagging scheme, we add 2 more tags and relations. We add an “Embedded Target”

(symbol ET) tag to represent attributes of the targets, embedded in the opinion phrases tagged by the *simple tagging scheme*. These attributes could have any relationship e.g. part-of, feature-of and instance-of, with the target of the opinion. More specifically in the *modified tagging scheme* we break the opinion phrases as tagged in the simple tagging scheme into opinion phrases and the embedded target phrases. We also add a “Negation” (symbol NO) tag to capture the negation of an opinion which often is located far from the opinion phrases (example 3 and 6 in Figure 2). The corresponding relations in our modified scheme are  $Embeddedtarget(OpinionPhrase, EmbeddedTarget)$  and  $Negation(NegationPhrase, OpinionPhrase)$ .



Figure 3: Example sentences tagged with modified tagging scheme.

Figure 3 shows the examples in Figure 2 tagged with the modified tagging scheme. From this tagging we can put together fragmented components of opinion and target phrases (Table 2) using the rule:  $Target(Op, Tp) \& Emb.Target(Op, ETp) \rightarrow Target(Op, Tp : ETp)$  i.e. if  $Tp$  is tagged as target phrase of the opinion phrase  $Op$  and  $ETp$  is tagged as its embedded target phrase then  $Tp : ETp$ <sup>4</sup> is the target of the opinion phrase  $Op$ . Similarly if a relation  $Negation(Np, Op)$  exists, the complete specification of the opinion is derived by adding the negation phrase  $Np$  to the opinion phrase  $Op$ .

As can be seen in Table 2, the modified tagging scheme is able to capture the opinions and their targets more precisely than the simple tagging scheme. In addition, opinion phrases become mostly domain independent. Still, there is some information loss. For example in sentence 4 “for” rela-

<sup>4</sup>The colon in this expression is intended to join specifications of the target.

Sentence	Opinion phrase	Target Phrase
1	good	This place: food
	good	This place: entertainment
2	does not realize how poor	The server: service
3	not anymore: a great	This: place to eat
4	great	This: place
	romantic	This: evening
5	knowledgeable	My server: wine
	great	they: wine
6	have given up: love	this place

Table 2: Tagged information in Figure 3.

tionship between the two opinion phrases is ignored (“place is great for romantic evening”), instead we extract “This:place” is “great” and “This<sup>5</sup>:evening” is “Romantic”. This although not exact, captures the essence of the reviewer’s opinion without additional complexity in the tagging scheme. In the rest of this section, we describe other natural language structures used to express opinions and also show how they are handled in our tagging scheme.

### 3.1 Ambiguous Targets

In many situations it is difficult to distinguish between the target and the embedded target. In Figure 4 two possible tags on a sentence are shown. In the

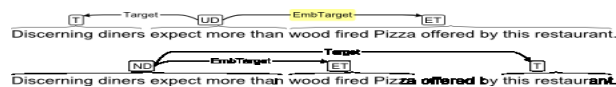


Figure 4: Examples sentences showing ambiguity in tagging targets and embedded targets.

first version, the neutral opinion about the “discerning diners” is tagged. In the second version, domain dependent negative opinion about “this restaurant” is tagged. If the context of tagging i.e. interest in opinions about the restaurants, was known, this ambiguity is resolved. In our context free tagging, we resolve this ambiguity by preferring the subject of the sentence as the main target of the opinions.

### 3.2 Conditional Opinions

Opinions are also expressed in conditional form and sometimes, like in example 1 and 2 in Figure 5, it is difficult to separate the opinion phrases from the target/embedded target phrases and the only choice is to tag entire sentence/segment as domain dependent neutral. Even though in the first sentence opinion about the loud music is expressed, and in the sec-

<sup>5</sup>Anaphora resolution will bind “This” to the reviewed restaurant.

ond sentence opinion about the food of the restaurant is expressed, they cannot be tagged as such even with our modified scheme. Examples 3 and 4 as

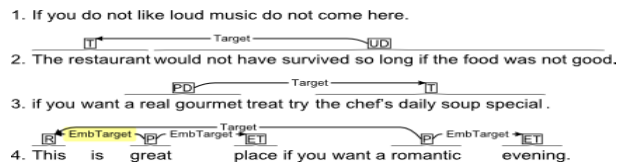


Figure 5: Examples of conditional opinion phrases.

shown in Figure 5 however, can be segmented into opinions and their targets/embedded targets. These examples illustrate that when there are no negations in the conditional opinions they typically can be segmented into opinion and target phrases.

### 3.3 Opinion Referencing Other Opinion Phrases

Figure 6 shows examples where opinions about other opinions are expressed. In the first example, the opinion “most impressive” reinforces other opinions; such reinforcement cannot be represented in our tagging scheme. In the second example, however, the pronoun “it” references the magazine’s opinion, which is ignored in our tagging scheme.



Figure 6: Examples where opinion expressions reference other opinion expressions.

### 3.4 Implicit Target Switch

In the first part of the sentence shown in Figure 7 an opinion about “this: steak” is expressed and in the second part an opinion about “this: ambiance” is expressed. Clearly if “this” refers to a steak, it cannot have ambiance. It must be the ambiance of the restaurant serving the steak. Our tagging scheme does not capture this implicit target switching.



Figure 7: Example of implicit target switching.

## 4 Coverage experiment

Table 3 shows the counts of domain dependent opinion phrases tagged on a small sample of data from 3 different domains, using both simple and modified schemes. The number of domain dependent opinion phrases in case of the modified tagging scheme is reduced by more than half. Even for the MP3 players with a large domain dependent vocabulary, 73% of opinion phrases are tagged as domain independent. This will make models trained on different domains useful even for MP3 players.

Domain	Num. Sentences	Number of Tagged Opinion Phrases		
		Total	Domain Dependent Simple	Domain Dependent Modified
Restaurant	68	101	31(30%)	13(13%)
Hotels	147	111	39(35%)	15(14%)
MP3 Plyr.	350	287	103(36%)	48(17%)

Table 3: Comparison of simple and modified scheme.

## 5 Relationship to other work

Kessler et al. (2010)<sup>6</sup> have tagged automobile data (JDPA Corpus) with sentiment expressions (our opinion phrases) and mentions (our target and embedded target phrases). JDPA representation is more extensive than ours. It explicitly represents many relationships among mentions and a number of modifiers of sentiment expressions. The strength of our scheme however, is in the way we choose the targets. In JDAP, mentions are tagged as targets of their *modifying* sentiment expressions. In our scheme we tag the main object as the target of opinions. For some cases both JDPA and our schemes result in equivalent representations, but for others we believe our scheme results in a more accurate representation.

As can be verified for Example 1 in Figure 2 both schemes result in an equivalent representation. For example in Figure 8, on the other hand our scheme represents the opinion expressions more accurately than JDPA. This example contains an opinion about any good camera. Therefore, the target of the opinion in our scheme is “good camera” and not “camera” by itself, and the opinion is “must have a great zoom”, “zoom” being embedded target we can drive *Target(must have a great, good camera:zoom)*. In JDPA this will be represented as *Target(good,camera), Target(must have a great,zoom)*,

<sup>6</sup>Author is thankful to the reviewers of the paper to point out this reference.

part-of(camera, zoom). Notice that JDPa explicitly represents that the “camera is good”, which is not true, and is not represent in our scheme.



Figure 8: Example where our scheme captures the opinions more accurately than JDPa.

The tagged data by Hu and Liu (2004)(H-L data) is the another data that has opinions and their targets labeled. It has been used by many researchers to benchmark their work. We randomly selected reviews from the H-L data and tagged them with our modified tagging scheme (Figure 9). Several observations can be made from Table 4, showing information tagged by our scheme and by the labels in the H-L data. First, not all opinion and targets are tagged in the H-L data. Instead of tagging the opinion phrases directly, the H-L data relies on labeler’s assessments for polarity strengths of the opinion. In the H-L data even the targets may or may not be present in the sentence (example 2). Again the H-L data relies on the labeler’s assessment of what the target is. Clearly in the H-L data the labeling is performed with a specific context in mind while our scheme makes no such assumption. The main reason for this difference is that Hu and Liu (2004) used this data only to test their un-supervised technique, while our motivation is to use the tagged data for supervised training of models that could be used across domains. With the contextual assumptions made in the labeling, the models trained by using the H-L data will perform very poorly when used across domains.

Modified Tagging Scheme			H-L Label	
Opinion	Pol.	Target	Pol.	Target
incredibly overpriced	neg	apple i-pod		
not(regret)	pos	the purchase	3	player
Not(any doubts)	pos	this player		
easy	pos	software: to use	2	software
much cheaper	pos	player	2	price
good looking	pos	player		
beautiful	pos	blue back-lit screen		
good	pos	this→lack of a viewing hole for ..		
not(damaged/scratched)	pos	the face		
fast	pos	transfer rate		
suck	neg	the stock ones→headphones	-3	headphone
will out sell	pos	this player		

Table 4: Side by side comparison of tagged information with our modified tagging scheme and H-L data

Wiebe et al. (2005) describe the MPQA tagging

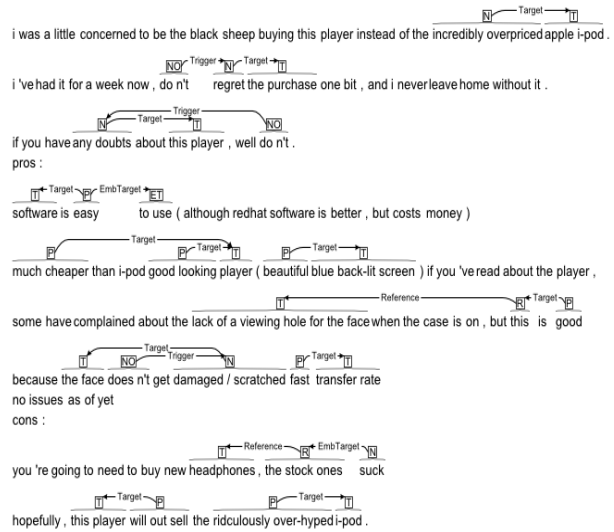


Figure 9: Tagging a review from Hu and Liu (2004) data using the our modified tagging scheme.

scheme for identifying private states of agents, including those of the author and any other agent referred in the text. The MPAQ tags *direct subjective expressions* (DSE) e.g. “faithful” and “criticized”, and *expressive subjective elements* (ESE) e.g. “highway robbery” and “illegitimate”, to identify the private states. We only tag author’s opinions. For example in ““The US fears a spill-over,” said Xirao-Nima” the MPQA will identify the private states of “US” and of “Xirao-Nima”. We, however, will not tag this sentence since the author is not expressing any opinion.

Opinions are part of an agent’s private state, but not all private states are opinions. For example in the sentence “I am happy” the author is describing his private state and not an opinion. In the MPQA the author’s private state will be identified by “happy” but, in our tagging scheme this sentence will not be tagged. However, in the sentence “I am happy with their service” author is expressing an opinion about “their service” and will be tagged in our scheme.

Another difference between MPQA and our scheme is that MPQA tags only the private states of agents, causing some inconsistencies as illustrated by the following example. In the sentence “The U.S. is full of absurdities”, “absurdities” is tagged as a private state of the U.S. At the same time in sentence “The report is full of absurdities”, “absurdities” is tagged as a private state of the author, and

“the report” is relegated to its target. In our tagging scheme both “the US” and “the report” are consistently tagged as targets of the opinion phrase “absurdities”. Because of these differences we believe that the MPQA data is less suitable for opinion mining research.

## 6 Conclusion

We discussed a tagging scheme to tag data for training information extraction models to extract from textual reviews the features of a product/service and opinions about them, and which can be used across domains with minimal adaptation. We demonstrated that a) by using a simple tagging scheme a large proportion of opinion phrases are tagged as domain dependent, defeating our goal to train models usable across domains; b) even when a domain independent vocabulary is used, a more complex tagging scheme is needed to fully disambiguate opinion and target phrases. Instead of addressing this complex representation problem, we show that by introducing two additional tags the number of domain dependent opinion phrases is reduced from 36% to 17%. This will lead to information extraction models that perform better when used across domains.

## 7 Acknowledgments

Author is thankful to Amanda Stent and to the anonymous reviewers for their comments and suggestions, which significantly contributed to improving the quality of the publication.

## References

- Baccianella, Stefano, Andrea Esuli, and Fabrizio Sebastiani. 2009. Multi-facet rating of product reviews. In *Proceedings of the 31th European Conference on IR Research on Advances in Information Retrieval (ECIR 09)*. pages 462–472.
- Hu, Minqing and Bing Liu. 2004. Mining and summarizing customer reviews. In *Proceedings of the ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD)*. pages 168–177.
- Kessler, Jason S., Miriam Eckert, Lyndsay Clark, and Nicolas Nicolov. 2010. The icwsm 2010 jdpa sentiment corpus for the automotive domain. In

*Proceedings of the 4th International AAAI Conference on Weblogs and Social Media Data Challenge Workshop (ICWSM-DCW 2010)*.

- Li, Fangtao, Chao Han, Minlie Huang, Xiaoyan Zhu, Yingju Xia, Shu Zhang, and Hao Yu. 2010. Structure-aware review mining and summarization. In *Proceedings of the 23rd International Conference on Computational Linguistics (COLING 2010)*. pages 653–661.
- Liu, Kang, Liheng Xu, and Jun Zhao. 2012. Opinion target extraction using word-based translation model. In *Proceedings of the 2012 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning, EMNLP-CoNLL 2012*. pages 1346–1356.
- Popescu, Ana-Maria and Oren Etzioni. 2005. Extracting product features and opinions from reviews. In *Proceedings of the Human Language Technology Conference and the Conference on Empirical Methods in Natural Language Processing (HLT/EMNLP)*.
- Wiebe, Janyce, Theresa Wilson, and Claire Cardie. 2005. Annotating expressions of opinions and emotions in language. *Language Resources and Evaluation* 39(2-3):165–210.