Using Verbs and Adjectives to Automatically Classify Blog Sentiment

Paula Chesley
Department of Linguistics
State University of New York at Buffalo
pchesley@buffalo.edu

Bruce Vincent, Li Xu, and **Rohini K. Srihari**
Department of Computer Science and Engineering
State University of New York at Buffalo
{bvincent,lixu}@buffalo.edu
rohini@cedar.buffalo.edu

Abstract

This paper presents experiments on subjectivity and polarity classifications of topic- and genre-independent blog posts, making novel use of a linguistic feature, verb class information, and of an online resource, the Wikipedia dictionary, for determining the polarity of adjectives. Each post from a blog is classified as *objective*, *positive*, or *negative*. Our method of determining the polarity of adjectives has an accuracy rate of 90.9%. Accuracy rates of two verb classes demonstrating polarity are 89.3% and 91.2%. Initial classifier results show blog-post accuracies with significant increases above the established baseline classification.

Introduction

This paper describes textual and linguistic features we extract and a classifier we develop to categorize blog posts with respect to sentiment. We ask whether a given blog post expresses subjectivity (vs. objectivity), and whether the sentiment a post expresses represents positive ("good") or negative ("bad") polarity. We simplify the expression of subjectivity vs. objectivity, as well as positive and negative polarity, into binary classification tasks.

In this research we address the following issues related to blog sentiment analysis:

1. How effectively can classes of verbs in blog posts categorize sentiment?
2. How can we utilize online lexical resources to automatically categorize adjectives expressing polarity?

We make use of verb-class information in the sentiment classification task, since exploiting lexical information contained in verbs has shown to be a successful technique for classifying documents (Klavans & Kan 1998). To obtain this information we use Semantex (formerly InfoXtract) (Srihari *et al.* 2006), an automatic text analyzer which groups verbs according to classes that often correspond to our polarity classification. Additionally, we utilize Wikipedia's online dictionary, the Wiktionary[1], to determine the polarity of adjectives seen throughout the posts. We then propagate this lexical information up to the post level by a machine learning classification algorithm in which a binary classification (e.g., objective/subjective) is made for each blog post.

	Objective	Positive	Negative	Total
Training	580	263	233	1076
Test	29	25	22	76

Table 1: A breakdown of our training and test datasets.

Building training and test datasets

Our classifier was trained using a dataset of objective, positive, and negative documents. Training text from the web was obtained by a two-step semi-automated, manually verified approach:

1. Obtain web documents via RSS and Atom web syndication feeds. Although the test dataset was manually chosen strictly from a set of blog posts of diverse topics, the training data were obtained from traditional websites as well as blogs. Objective feeds were from sites providing content such as world, national, and local news (CNN, NPR, Atlanta Journal and Constitution, etc.), and various sites focused on topics such as health, science, business, and technology. Subjective feeds included content from newspaper columns (e.g., Charles Krauthammer), letters to the editor (e.g., Washington Post), reviews (RottenTomatoes.com, etc.), and political blogs (Huffington Post, etc.).
2. Manually verify each document to confirm it strongly corresponds to an objective, positive, or negative categorization.

Due to varying degrees of sentiment expression in the feeds, we only considered documents that are categorically positive, negative, or objective. That is, for subjective texts, a given text could only express one opinion. The average rate of document retention of positive and negative feeds was 19%, a figure which illustrates the difficulty of obtaining quality subjective data at the document level given this criterion. For objective texts, we mandated that the sole goal of the author be to inform. Table 1 shows a breakdown of our datasets.

[1]http://en.wiktionary.org/wiki/.

Due to the ironically subjective nature of what constitutes subjectivity and polarity in a text, we carried out an agreement study on manual classification of our training dataset. We checked the agreement of our two raters on 75 randomly chosen documents in our training data, equally distributed across positive, negative, and objective categories. Categorical inter-rater agreement for judging a document objective, positive, or negative was K = .77. In our test set, we only used blog posts that both raters marked as objective, positive, or negative.

Feature-based Classification

In this work we make use of several features of three principal types for our classification task: textual features (exclamation points and question marks), part-of-speech features, and lexical semantic features. No part-of-speech features showed an impact on the classifications; hence for reasons of space we do not discuss them further. Each post is then represented as a feature vector in an SVM classification.

The current work aims to tease out how verbs express sentiment in using the verb classes in Semantex. Our verb-class information is often more fine-grained than the Levin verb classes (Levin 1993) and can systematically incorporate polarity into a class. For example, *approving* verbs most often show positive orientation, while *doubting* verbs tend to show negative orientation. Hence we can capture sentiment expressed by statements like “I wholeheartedly *support* [Bush] for the same reason I did when I wrote the above posts.” Verbs such as *agree* and *support* are included in the *approving* verb class. Some verb classes, such as *approving* and *doubting* verbs, are included to elicit sentiment polarity, while others, such as *answering* and *suggesting* verbs, are meant to highlight subjective/objective factors. Table 3 lists the full set of verb classes used in the classification task. These manually developed verb classes are discussed in (Srihari *et al.* 2006) and (Li *et al.* 2003).

We use the online Wikipedia dictionary to determine the polarity of adjectives in the text. The English Wiktionary contains at least 96,054 entries (some entries actually being for words in other languages), an amount that increases by at least 3,000 entries per month[2]. A principal advantage of this dictionary is its coarse-grained content: adjectives in this resource are usually tersely defined, often by a list of synonyms. In brief, we find this resource similar to WordNet with its concepts of glosses and synsets, but the Wiktionary is not as specific. While (Hu & Liu 2004) yield positive results using WordNet for sentiment analysis, (Klavans & Kan 1998) note that their results are degraded by the ambiguity and fastidious detail of verb synsets and that the same phenomenon occurs with nouns. Not wishing to overgenerate adjective polarity where there is none, we see the Wikipedia concept as ideal for determining adjective polarity.

Our Wiktionary method assumes that an adjective will most likely have the same polarity as the words that define it. We query the Wiktionary page for each adjective in a given post, limiting our search to the adjectival portion of an entry, modulo antonyms and example usages. This query looks for the number of adjectives in two manually constructed lists of known polarity in the entry. We count the number of adjectives of manually established polarity and assign the adjective in question the polarity of the majority of these found adjectives. If an adjective of known polarity in the definition is negated (adjectival dictionary entries with negative orientation are often defined in terms of a negated positive adjective), we take its polarity to be the opposite of its established polarity. Table 2 shows an example classification of an adjective using the Wiktionary.

“lucky”	
positive	negative
fortunate	–
good	–
good	–
good	–
fortunate	–
$5 > 0$	
output: positive	

Table 2: Classification of the polar adjective *lucky* using the Wiktionary. The adjectives *fortunate* and *good* figure among our list of manually constructed adjectives of positive polarity and were respectively seen two and three times in the adjective entry. Since our system recognizes more adjectives of known positive polarity than of known negative polarity for *lucky*, it classifies this adjective as positive.

The adjectives in our lists were chosen for their lack of part-of-speech and polarity ambiguity. We have initially opted for smaller lists of 34 adjectives for both positive and negative polarities, a list size which can be changed in order to maximize F-measures of adjective polarity. As for negation of polarity, in these first experiments we simply leave out of the post’s sentiment classification any verb or adjective in a five-word window of a negation (within, of course, the same sentence).

Each blog post is then represented as a vector with count values for each feature in table 3. We then make a binary classification for each post based on the score a Support Vector Machine (SVM) classifier assigns to a post. Our LIBSVM implementation[3] gives the following binary classifications for all posts: objective/non-objective (i.e., subjective), positive/non-positive, and negative/non-negative. The SVM parameters are maximized for test accuracy.

Results

In the 76 test files there were a total of 3,460 adjective tokens and 836 adjective types. Of these 836 types 10.5% were assigned a polarity by our program. We consider type accuracy of our Wiktionary method to be the number of adjectives assigned a plausible polarity for that adjective, given the Wiktionary entry for each adjective; this figure is 90.9%. We then looked at two verb classes to see whether or not the verbs in the class do represent the a priori polarity we had at-

[2] `http://en.wiktionary.org/wiki/Wiktionary:Multilingual_statistics`.

[3] `http:www.csie.ntu.edu.tw/˜cjlin/libsvm`.

Subjective vs. Objective Features		Polarity Features	
Objective Verb Classes answering, asking, asserting, explaining **Subjective Verb Classes** believing suggesting mental sensing	**Textual Features** exclamation points question marks	**Adjectives (Wiktionary)** positive adjectives negative adjectives	**Positive Verb Classes** positive mental affecting, approving, praising **Negative Verb Classes** abusing, accusing, arguing, doubting, negative mental affecting

Table 3: Features used in the sentiment classification task.

tributed them. Here accuracy is the number of verbs having a primary sense expressing the a priori polarity; for *doubting* verbs in our training data, accuracy is 89.3%. Accuracy on *positive mental affecting* verbs is 91.2%.

The baseline accuracy for each post classification is guessing the majority category, which is the opposite of the category elicited. For example, in the positive/non-positive classification we assume a post is either objective or negative, for a total of 51/76 posts, a 67.1% baseline. Table 4 shows that every classification using our textual and linguistic features yields results that represent a significant increase above the baseline figures (paired t-test, all $p < .01$).

To get an idea of which individual features were contributing to the accuracy of our classifier, we experimented in holding out each feature. Holding out a feature such as the positive adjectives we obtained from the Wiktionary significantly decreases system accuracy, which implies that this feature plays a large role in the correct classification of posts. The effects of holding out individual features are given in table 4. Significant effects of holding out features on a given classification task (paired t-test, all $p < .01$ except except *approving* verbs, for which $p = .013$) are given in bold.

The hold-out experiments show initial results of what features prove useful in a given classification. For objective and positive posts, positive adjectives acquired from Wikipedia's Wiktionary play a key role in increasing overall accuracy. Thus our strategy for using an online dictionary in a sentiment classification task appears to have been successful. The *asserting* and *approving* verb classes also contribute significantly to accuracy of the positive classification. These results show that verb classes can improve results on sentiment classification of blogs. These preliminary findings would also seem to indicate that we can substantially reduce our feature space for the classification task.

The principal challenge remaining for us is classifying a blog post for sentiment with the same high levels of accuracy we obtain at the lexical level. In improving recall rates at the lexical level, we aim to increase overall post-level accuracy: while adjective accuracy in test data is high, the corresponding recall rates are low. Hopefully the high rate of lexical entries added monthly to the Wiktionary will help to improve adjective recall, which will in turn increase accuracy at the post level.

Held-out feature	Objective	Positive	Negative
All (baseline)	61.8%	67.1%	71.1%
None (system accuracy)	**72.4%**	**84.2%**	**80.3%**
exclamation points	69.7%	82.9%	80.3%
question marks	72.4%	84.2%	80.3%
answering verbs	72.4%	84.2%	80.3%
asking verbs	72.4%	84.2%	80.3%
asserting verbs	72.4%	**76.3%**	80.3%
explaining verbs	76.3%	86.8%	80.3%
believing verbs	72.4%	84.2%	80.3%
suggesting verbs	76.3%	85.5%	80.3%
mental sensing verbs	72.4%	84.2%	80.3%
positive mental affecting verbs	75%	84.2%	80.3%
approving verbs	73.7%	**77.6%**	80.3%
praising verbs	73.7%	84.2%	81.6%
abusing verbs	73.7%	84.2 %	80.3%
accusing verbs	72.4%	84.2%	80.3%
arguing verbs	72.4%	81.6%	80.3%
doubting verbs	71.1%	84.2%	80.3%
negative mental affecting verbs	71.1%	82.9%	81.6%
positive adjectives	**61.8%**	**67.2%**	80.3%
negative adjectives	75%	82.8%	78.9%

Table 4: Effects of holding out an individual feature on accuracy, against system accuracy using all features. Verb classes are grouped according to a priori subjectivity and polarity judgements–e.g., *positive mental affecting*, *approving*, and *praising* verbs all express positive polarity.

References

Hu, M., and Liu, B. 2004. Mining and Summarizing Customer Reviews. In *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 168–177.

Klavans, J., and Kan, M. 1998. Role of Verbs in Document Analysis. In *Proceedings of the Conference COLING-ACL, Montreal*, 680–686.

Levin, B. 1993. *English Verb Classes and Alternations*. U. Chicago Press.

Li, W.; Zhang, X.; Niu, C.; Jiang, Y.; and Srihari, R. 2003. An Expert Lexicon Approach to Identifying English Phrasal Verbs. In *Proceedings of ACL 2003*, 513–520.

Srihari, R.; Li, W.; Niu, C.; and Cornell, T. 2006. InfoXtract: A Customizable Intermediate Level Information Extraction Engine. *Journal of Natural Language Engineering* 12:4.