Wishful Thinking
Finding suggestions and 'buy' wishes from product reviews

J. Ramanand
BFS Innovations
Cognizant Technology Solutions
Pune, India
ramanand.janardhanan@cognizant.com

Krishna Bhavsar
BFS Innovations
Cognizant Technology Solutions
Pune, India
krishna.bhavsar@cognizant.com

Niranjan Pedanekar
BFS Innovations
Cognizant Technology Solutions
Pune, India
niranjan.pedanekar@cognizant.com

Abstract

This paper describes methods aimed at solving the novel problem of automatically discovering 'wishes' from (English) documents such as reviews or customer surveys. These wishes are sentences in which authors make suggestions (especially for improvements) about a product or service or show intentions to purchase a product or service. Such 'wishes' are of great use to product managers and sales personnel, and supplement the area of sentiment analysis by providing insights into the minds of consumers. We describe rules that can help detect these 'wishes' from text. We evaluate these methods on texts from the electronic and banking industries.

1 Introduction

Various products and business services are used by millions of customers each day. For the makers of these products & services, studying these customer experiences is critical to understanding customer satisfaction and making decisions about possible improvements to the products. Thanks to the advent of weblogs, online consumer forums, and product comparison sites, consumers are actively expressing their opinions online. Most of these reviews are now available on the web, usually at little or no cost. Moreover, these are available for a variety of domains, such as financial services, telecom services, consumer goods etc.

Automated analysis of opinions using such reviews could provide a cheaper and faster means of obtaining a sense of such customer opinions, thus supplementing more traditional survey methods. In addition, automated analysis can significantly shorten the time taken to find insights into the customer's mind and actions.

Sentiment analysis of texts such as product reviews, call center notes, and customer surveys aims to automatically infer opinions expressed by people with regards to various topics of interest. A sentiment analysis exercise classifies the overall opinion of a review document into positive, neutral, or negative classes. It may also identify sentiments at a finer granularity, i.e. recognizing the mix of opinions about the topic(s) expressed in the text. However, industry analysts (Strickland, 2009) report some common problems with the results of these exercises:

1. The results (usually numerical scores split across positive, negative, neutral classes) are hard to meaningfully interpret.
2. These results are more useful to certain roles and domains. Brand, reputation, and service managers in media and retail industries find sentiment analysis more useful than product managers or sales teams in various industries.
3. The results do not *'indicate user action'* i.e. opinions do not help identify a future action of the author based on the comments. An example of this is: does the consumer indicate that he intends to stop using a service after a negative experience?
4. The reader of the report often asks *"what do I do next?"* i.e. the results are not always 'actionable'. There is a gap between understanding the results and taking an appropriate action.

Proceedings of the NAACL HLT 2010 Workshop on Computational Approaches to Analysis and Generation of Emotion in Text, pages 54–61,
Los Angeles, California, June 2010. ©2010 Association for Computational Linguistics

This has led to interest in identifying aspects indirectly related to sentiment analysis, such as gauging possible loss of clientele or tapping into desires to purchase a product. Many of these methods attempt to identify *'user intent'*.

In this paper, we propose rule-based methods to identify two kinds of 'wishes' – one, the desire to see improvement in a product, and the other to purchase a product. These methods have been designed & tested using a variety of corpora containing product reviews, customer surveys, and comments from consumer forums in domains such as electronics and retail banking. From our reading, there has been only one published account of identifying 'wishes' (including suggestions) and no known work on identifying purchasing wishes. We hope to build approaches towards more comprehensive identification of such content.

The paper is organized as follows. We begin by discussing some of the work related to this upcoming area. Section 3 details our characterization of wishes. Section 4 describes the corpora used for these methods. We discuss our proposed algorithms and rules in Sections 5 & 6, including a discussion of the results. Finally, we wrap up with our conclusions and directions for future work.

2 Related Work

The principal context of our work is in the area of sentiment analysis, which is now a widely researched area because of the abundance of commentaries from weblogs, review sites, and social networking sites. In particular, we are interested in the analysis of product reviews (Dave et al., 2003; Hu and Liu, 2004), as well as its application to more service-oriented industries such as banks.

We have built a sentiment analyzer that can analyze product and service reviews from a variety of domains. This also accepts social networking commentaries, customer surveys and news articles. The implementation follows a lexicon-based approach, similar to the one described in Ding et al. (2008), using lexicons for product attributes and opinion words for basic sentiment analysis.

Our work is not a sub-task of sentiment analysis, but supplements the area. A similar example of a classification task that works on the sentence level and is also related to sentiment analysis is Jindal and Liu (2006) which aims to identify comparisons between two entities in texts such as product reviews.

Goldberg et al. (2009) introduced the novel task of identifying wishes. This used a "WISH" corpus derived from a web site that collected New Year's wishes. Goldberg et al. (2009) studied the corpus in detail, describing the nature, geography, and scope of topics found in them. The paper also looked at building 'wish detectors', which were applied on a corpus of political comments and product reviews. A mix of manual templates and SVM-based text classifiers were used. A method to identify more templates was also discussed.

Our task, though similar to the above problem, has some novel features. In particular, there are two significant differences from Goldberg et al. (2009). We are interested in two specific kinds of wishes: sentences that make suggestions about existing products, and sentences that indicate the writer is interested in purchasing a product. (These are described in detail in Section 3.) Secondly, our interest is limited to product reviews, and not to social or political wishes.

In Requirements Engineering, some methods of analyzing requirement documents have used linguistic techniques to understand and correlate requirements. These are somewhat related to our task, aiming to detect desired features in the project to be executed. och Dag et al. (2005) has some useful discussions on this topic.

Kröll and Strohmaier (2009) study the idea of Intent Analysis, noting a taxonomy of Human Intentions, which could be useful in future discussions on the topic.

3 What are Wishes

3.1 Defining Wishes

A dictionary definition (Goldberg et al. (2009)) of a *"wish"* is *"a desire or hope for something to happen."* Goldberg et al. (2009) discuss different types of wishes, ranging from political to social to business. In our case, we limit our interest to comments about products and services. In particular, we are interested in two specific kinds of wishes.

3.2 Suggestion Wishes

These are sentences where the commenter wishes for a change in an existing product or service. These range from specific requests for new product features and changes in existing behaviour, or an indication that the user is unhappy with the current experience. Examples[1]:

1. *I'd love for the iPod shuffle to also mirror as a pedometer.*
2. *It would be much better if they had more ATMs in my area.*

We also include sentences that do not fully elaborate on the required change, but could serve as a pointer to a nearby region that may contain the required desire. Examples of these:

1. *I wish they'd do this.*
2. *My wish list would be as follows:*

It is important to note the difference between our definition of wishes and that in Goldberg et al. (2009). That study seeks to discover any sentence expressing any desire. For instance, Goldberg et al. (2009) marks the following as wishes:

1. *I shouldn't have been cheap, should have bought a Toshiba.*
2. *hope to get my refund in a timely manner.*

In our approach, we do not treat these as wishes since they do not suggest any improvements.

In some cases, improvements could be inferred from a negative opinion about the product. The implication is that the customer would be happier if the problem could be fixed. Examples:

1. *"My only gripe is the small size of the camera body"* which implies *"I wish the camera was bigger"*.
2. *"The rubber flap that covers the usb port seems flimsy"* which implies *"I wish the rubber flap was more robust"*.

We do not address such implicit wishes.

3.3 Purchasing Wishes

These are sentences where the author explicitly expresses the desire to purchase a product. In some cases, a preferred price range is also indicated.

Examples:

1. *I have a Canon digital rebel xt, I am looking for a lens that will take sports actions football shots at night.*
2. *I want to purchase a cell phone range 12-15000/-... please suggest me some good and stylish phones?*
3. *We are also thinking of buying a condo in a few months…*

4 Corpora for Design and Evaluation

4.1 Suggestion Wishes

As part of building and testing our in-house sentiment analyzer, we collected a variety of texts from different sources such as popular consumer review sites (such as Epinions.com and MouthShut.com) and weblogs. These primarily belonged to the domains of electronics and retail banking. Of these, we chose reviews about the Apple iPod and a collection of banking reviews about five leading US banks. We also used customer surveys conducted for two products of a financial services company[2]. The sizes of the corpora are summarized in Table 1.

Some observations about these texts:

1. The texts are in American or British English and are largely well-formed.

2. They cover both reviews of products and descriptions of customer service.

3. The customer surveys consisted of sections for positives and negatives feedback, with an optional 'suggestions' section.

4. Wish sentences in the reviews were infrequent (on average, less than 1% of the total sentences). The surveys had a much larger presence of wishes (about 5% on average).

In addition, Goldberg et al. (2009) has made available a WISH corpus, which is a sample of 7614 sentences consisting of sentences from political discussions and product reviews. Since we are only interested in the latter, we evaluated our algorithm only on the product review sentences (1235 in number). 3% (41 sentences[3]) of these have been labeled as wishes.

[1] All sentences are taken from review sites such as epinions.com

[2] Anonymous for confidentiality reasons

[3] In the WISH corpus, 149 (12%) are marked as wishes; however we only chose those wishes that suggest improvements.

In a pre-processing step, individual sentences in the corpora were identified using GATE's (Cunningham, 2002) sentence splitter.

4.2 Purchasing Wishes

Similar to our collection of sentences for suggestions, we collected texts from review sites and consumer forums (such as Alibaba.com and Yahoo! Answers) that not only reviewed products and shared complaints but also allowed users to post requests for purchases.

The corpus consisted of 1579 sentences about the following products: Apple iPhone, Cameras, Desktop PCs, and a mix of Credit Cards from four leading Indian and American banks.

5 Finding Suggestions

5.1 Approach

The input to our system consists of the following:

1. Datasets containing sentences.
2. ATTRLEX[4]: A lexicon of product attributes for each of the domains. (e.g. the iPod attributes were words like *'battery'*, *'interface'* etc.)
3. POSLEX: A lexicon of positive opinions (words such as *'good'*, *'better'*, *'fast'*).
4. NEGLEX: A lexicon of negation words (these are words that invert the opinion of a sentence. e.g: *'not'*, *'wouldn't'*)

We began by manually classifying sentences in samples from each of the corpora as 'wishes' or 'non-wishes'. We then looked for common phrases and words across all these wishes to derive patterns and rules.

Initial analysis led to some proto-rules. These rules were then refined by using further analysis and in some cases, decision trees. The rules are grouped as follows.

5.1.1 Rules based on modal verbs

A majority of the wishes had pivotal phrases involving modal verbs such as *"would"*, *"could"*, *"should"* etc. Examples:

1. *It <u>would</u> be a much more valuable service if they would fix this flaw.*
2. *It <u>might be</u> nice if one could drag-and-drop music files and have the iPod reconstruct its index on-the-fly.*
3. *I <u>would</u> prefer the unit to have a simple on off switch.*

This led to the following rules:

a. modal verb + auxiliary verb + positive opinion word

Match sentences which contain the pattern:
<modal verb> <auxiliary verb> {window of size 3} <positive opinion word>

Where

Modal verb belongs to {*may, might, could, would, will, should*}
Auxiliary verb belongs to {*be, have been*}
Positive Opinion word belongs to POSLEX

The positive word should appear to the right of the modal verb in a pre-defined window size (usually 3 to 5).

b. modal verb + preference verb

Match sentences which contain the pattern:
<modal verb> {window of size 3} <preference verb>

Where

Modal verb belongs to {*may, might, would, will*}
Preference verb belongs to {*love, like, prefer, suggest*}

c. Other rules

Match sentences containing:
"should be able" or
"should come with" or
"could come with"

5.1.2 The *"needs to"* rule

Sentences containing the phrase *"needs to"* are candidate wishes, such as in the examples:
1. Apple <u>needs to</u> step it up and get better longer lasting batteries.
2. Their customer service representatives <u>need to</u> be educated in assisting customers.
3. <u>need to</u> be able to configure the boxes.

[4] These lexicons were built by semi-automated means using components built for our in-house sentiment analyzer which help detect opinions and attributes for a domain from related texts

For this pattern, we created a decision tree model with the following features:
1. Presence of negation word to the left of *"needs to"*
2. Presence of a *'product attribute'* word to the left
3. Whether the sentence is interrogative
4. Subject of the sentence from the list: {*I, you, s/he, we, this, that, those, it, they, one, someone, somebody, something*}

Based on analysis and the combination suggested by the decision tree experiments, we formulated rules. Some of these rules are as follows:
1. Interrogative sentences or those with a negation word to the left of *"need to"* are not wishes.
2. If the product attribute is present (usually as the subject), the sentence is a wish.
3. If the subject of the sentence is one of *"this, that, these"*, the sentence is likely to be a wish. When the subject is one of *"I, you, one"*, the sentence is not a wish.

5.1.3 Other rules

Sentences containing the patterns:
1. *"I wish"*: along with filters such as the subject (*"they, you, product"*) etc. can be matched as wishes.
2. *"hopefully"* or *"I hope"*
3. *"should be able to"* or *"should come with"*

These rules match very infrequently in the dataset. A summary of rule accuracy can be seen in Table 3.

5.2 Results

5.2.1 Precision of Rules

Type	Total sentences	No. of predicted wishes	No. of correct wishes	Precision
iPod	21147	90	53	58.89%
Banking	15408	75	23	30.67%
Product 1	4240	224	187	83.48%
Product 2	6850	355	284	80.00%
WISH corpus	1236	28	16	57.14%

Table 1 Precision of wish identification for various data sets

5.2.2 Recall of Rules

Recall was calculated on a 10% random sample from each data set, except in case of the WISH corpus, where all sentences were taken into account.

Type	No. of correctly predicted wishes in the sample	No. of actual wishes in the sample	Recall
iPod	7	14	50.0%
Banking	3	5	60.0%
Product 1	24	45	53.3%
Product 2	28	70	40.0%
WISH corpus	16	41	39.0%

Table 2 Recall of wish identification

5.2.3 Rule Analysis

This table analyses performance of the top 3 most frequently matched rules. For each type of data, the first row shows the number of wishes predicted by each rule. The succeeding row shows the corresponding precision.

Type/Rule	Modal, aux, positive opinion	Modal, preference	"Needs to"	Others
iPod	24	8	7	14
	57%	53%	43%	82%
Banking	14	17	7	2
	37%	85.0%	50%	28.5%
Product 1	89	56	25	17
	87%	83.6%	71%	85%
Product 2	146	25	50	30
	90%	71.4%	71%	90.9%
WISH Corpus	7	2	3	4
	63.6%	50%	50%	57.1%

Table 3 Rule Analysis

5.3 Comments on Results

Wishes occur very infrequently in reviews, where authors may or may not choose to talk about improvements. Surveys produced more wishes because of the design and objectives of the survey. Also, the language used in suggesting improvements was more consistent across authors, making it easier to catch them. Wishes could be made about existing product attributes, but several wish-

wishes were about newer aspects. This could help product managers envisage features that their customers are asking for.

Experiments on the banking reviews showed the worst results. The dataset had very few wishes and the language used was usually part of a narrative, which threw up a lot of false positives. It could also be that the nature of the collected dataset was such that it did not contain sufficient number of wishes.

Some of the false positives were difficult to avoid. Take an example such as:

I wish it will be a better year.

Though it is a 'wish' in general, this does not fit our definition of product suggestion though it fits a rule well. More semantic or contextual analysis may be required in this case. We do not filter out sentences that do not refer to already published product attributes since authors may be talking about adding completely new features, such as in the case:

I wish it will be in magazine form next year.

Of the rules, the first rule (modal + auxiliary + positive opinion word) had the highest contribution to make. The second rule was more consistent in detecting correct wishes. Incidentally, the "needs to" rule for banking reviews outperforms the results for iPod sentences – the only time this happens.

Different patterns may be applicable for different domains and types of texts. A possible approach to improving results would be to have a 'rule selection' phase were rules that fall below a certain threshold are discarded.

6 Finding Buy Wishes

6.1 Approach

Similar to finding suggestions, we assembled a corpus of sentences for various products and services, this time from forums that also contain buy-sell sections. These may contain comments like:

1. I am trying to find where I can purchase the complete 1st season of Army Wives-can you help me?
2. I am seriously looking for a new bank...
3. I want to give a new year's present to my 5 year old nephew. My budget is 1500 Rupees.

We derived proto-rules and refined them by manual analysis and decision trees. The pattern of each rule is:

…<rule phrase> <common sub-rule>…

If a sentence contains such a pattern, it is deemed to be a buy wish.

To begin, we describe a common sub-rule that is used with all rules.

6.1.1 Buy Identification common sub-rule

This depends on the following three aspects:

a. A *'buy verb'* from among {*find, buy, purchase, get, acquire*} should be present

b. Absence of a negation word (from NEGLEX) to the left of rule phrase

c. Subjects:

The subject should not be one of these: {*you, one, they, someone, those*}

The subject could be one of these: {*I, we, me*}

6.1.2 Rule phrases

Rule phrases are one of the following

1. *"want to"*
2. *"desire to"*
3. *"would like to"*
4. *"where can/do I"*
5. *"place to"*
6. *"going to"*
7. *"looking to/for"*
8. *"searching to/for"*
9. *"interested in"*

Of these, in rules involving phrases 7, 8, and 9, we also check if there are any past tense verbs preceding rule phrase. In such cases, we do not classify the sentence as a wish. For phrase 5, interrogative sentences are also ignored.

6.2 Results

6.2.1 Precision

Type	Total sentences	No. of predicted wishes	No. of correct wishes	Precision
iPhone	193	43	41	95.34%
iPod	176	48	37	79.54%
Credit Cards	865	6	4	66.67%

Canon Cameras	170	40	39	97.50%
Desktop PCs	175	36	34	94.44%

Table 4 Precision of wish identification for various data sets

6.2.2 Recall[5]

Type	No. of expected wishes	No. of correctly predicted wishes	Recall
iPhone	80	41	51.25%
iPod	54	37	68.51%
Canon Camera	65	39	60.00%
Desktop PCs	66	34	51.52%

Table 5 Recall of wish identification

6.2.3 Rule Analysis

This table analyses the precision of the tope three rules that matched the most sentences.

Rule Phrase	No. of matched sentences	No. of correct matches	*Precision*
Looking for	98	85	86.73%
Want to	24	22	91.67%
Interested In	6	6	100%

Table 6 Rule Analysis

6.3 Comments on Results

Buy wishes tend to occur only in forums where buyers can advertise their search and hope to receive advice or meet prospective sellers. In addition to sites dedicated to specific products, social networks such as Twitter[6] also provide such a platform. This is in contrast to regular weblogs.

The results for all the electronic products showed a precision of about 80% or more. As in the case of suggestion wishes, wishes were very rare in the credit cards postings.

The recall in all cases was above 50%. Buy wish sentences matching The *"looking for"* and *"want to buy/purchase"* rules were common. An observation was that in some cases, people would simply list the expected attributes of the product they were looking for. Because of the nature of the forum, other users would interpret it as a buy/sell request. We could not separate these sentences from other kinds of sentences in the data set.

In most cases, the sentences were terse and used phrases like *"we need"* and *"seeking"*. Further expanding the rule phrases & sub-phrases to include their synonyms is likely to improve recall.

7 Conclusions and Future Work

This paper described two novel problems in the world of opinion and intention mining, that of identifying 'wishes' relating to improvements in products and for purchasing them. These are likely to be directly useful to business users. We build approaches towards such detections, by the use of English-language patterns. To the best of our knowledge, this is the first attempt at solving such problems.

The approach for identifying suggestions works best for texts that contain explicit wishes, especially customer surveys. They work reasonably well for (electronic) product reviews. In contrast, reviews about banking services tend to contain narratives and have more implicit opinions and wishes. Similarly, the algorithm to detect buy wishes works well for electronic product reviews in comparison to banking products.

Wish statements appear very infrequently in reviews. Existing sentiment analysis corpora may not be sufficient to use in creating wish detectors. Augmenting corpora such as the WISH dataset or creating even more robust and representative corpora would be a must for such exercises. A possible source could be the "Make A Wish" foundation.

One of the possible future directions could be to look at tense and mood analysis of sentences. Wish sentences come under the 'optative' mood. Techniques that help identify such a mood could provide additional hints to the nature of the sentence. More features related to parts of speech and semantic roles could be explored.

We also plan to look at machine learning approaches, but the availability of good quality training data is a limiting factor.

The emergence of social networking sites may provide more challenges for such detectors. Sites like Twitter are already being used to advertise

[5] The credit cards set had very few actual wishes (less than 10) with which to carry out a meaningful recall exercise

[6] http://twitter.com

intentions to buy or sell. However, the nature of discourse in these media is markedly different to regular reviews and forums due to size restrictions.

Any system that helps business users to identify new customers or engage with existing ones would need to tap into all these emerging channels. The need for such detectors is likely to increase in the future, thus providing further motivation to study this nascent area.

References

Hamish Cunningham, Diana Maynard, Kalina Bontcheva, and Tablan, Valentin. *GATE: A framework and graphical development environment for robust NLP tools and applications.* 2002

Kushal Dave, Steve Lawrence, and David M. Pennock. *Mining the peanut gallery: Opinion extraction and semantic classification of product reviews.* Proceedings of the 12th international conference on World Wide Web. 2003.

Minqing Hu and Bing Liu. *Mining and summarizing customer reviews.* Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining. 2004.

Andrew B. Goldberg, Nathanael Fillmore, David Andrzejewski, Zhiting Xu, Z, Bryan Gibson, and Xiaojin Zhu. *May all your wishes come true: A study of wishes and how to recognize them.* Proceedings of Human Language Technologies: The 2009 Annual Conference of the North American Chapter of the Association for Computational Linguistics. 2009.

Nitin Jindal and Bing Liu. *Identifying comparative sentences in text documents.* Proceedings of the 29th annual international ACM SIGIR conference on Research and development in information retrieval. 2006.

Mark Kröll and Markus Strohmaier, M. *Analyzing human intentions in natural language text.* Proceedings of the fifth international conference on Knowledge capture. 2009.

Johan Natt och Dag, Vincenzo Gervasi, Sjaak Brinkkemper, and Björn Regnell, B. *A linguistic engineering approach to large-scale requirements management.* Managing Natural Language Requirements in Large-Scale Software Development. Vol 22-1. 2005.

Marta Strickland. *Five Reasons Sentiment Analysis Won't Ever Be Enough.* http://threeminds.organic.com/2009/09/five_reasons_sentiment_analysi.html. 2009.

Xiaowen Ding, Bing Liu, and Philip S.Yu. *A holistic lexicon-based approach to opinion mining.* Proceedings of the international conference on Web search and web data mining. 2008.