May All Your Wishes Come True: A Study of Wishes

Abstract

A wish is "a desire or hope for something to happen." In December 2007, people from around the world offered up their wishes to be printed on confetti and dropped from the sky during the famous New Year's Eve "ball drop" in New York City's Times Square. We present an in-depth analysis of this collection of wishes. We then leverage this unique resource to conduct the first study on building general "wish detectors" for natural language text. Wish detection complements traditional sentiment analysis and is valuable for collecting business intelligence and insights into the world's wants and desires. We demonstrate the wish detectors' effectiveness on domains as diverse as consumer product reviews and online political discussions.

1 Introduction

Each year, New York City rings in the New Year with the famous "ball drop" in Times Square. In December 2007, the Times Square Alliance, co-producer of the Times Square New Year's Eve Celebration, launched a Web site called the Virtual Wishing Wall[1] that allowed people around the world to submit their New Year's wishes. These wishes were then printed on confetti and dropped from the sky at midnight on December 31, 2007 in sync with the ball drop.

We obtained access to this set of nearly 100,000 New Year's wishes, which we call the "WISH corpus." Table 1 shows a selected sample of the WISH corpus. Some are far-reaching fantasies and aspirations, while others deal with everyday concerns like economic and medical distress. We analyze this first-of-its-kind corpus in Section 2.

The New Oxford American Dictionary defines "wish" as "a desire or hope for something to happen." How wishes are expressed, and how such wishful expressions can be automatically recognized, are open questions in natural language processing. Leveraging the WISH corpus, we conduct the first study on building general "wish detectors" for natural language text, and demonstrate their effectiveness on domains as diverse as consumer product reviews and online political discussions. Such wish detectors have tremendous value in collecting business intelligence and public opinions. We discuss the wish detectors in Section 3, and experimental results in Section 4.

1.1 Relation to Prior Work

Studying wishes is valuable in at least two aspects:

1. Being a special genre of subjective expression, wishes add a novel dimension to sentiment analysis. Sentiment analysis is often used as an automatic market research tool to collect valuable business intelligence from online text (Pang and Lee, 2008; Shanahan et al., 2005; Koppel and Shtrimberg, 2004; Mullen and Malouf, 2008). Wishes differ from the recent focus of sentiment analysis, namely opinion mining, by revealing what people explicitly want to happen, not just what they like or dislike (Ding et al., 2008; Hu and Liu, 2004). For example, wishes in product reviews could contain new feature requests. Consider the following prod-

[1] http://www.timessquarenyc.org

514	*peace on earth*
351	*peace*
331	*world peace*
244	*happy new year*
112	*love*
76	*health and happiness*
75	*to be happy*
51	*i wish for world peace*
21	*i wish for health and happiness for my family*
21	*let there be peace on earth*
16	*i wish u to call me if you read this 555-1234*
16	*to find my true love*
8	*i wish for a puppy*
7	*for the war in iraq to end*
6	*peace on earth please*
5	*a free democratic venezuela*
5	*may the best of 2007 be the worst of 2008*
5	*to be financially stable*
1	*a little goodness for everyone would be nice*
1	*i hope i get accepted into a college that i like*
1	*i wish to get more sex in 2008*
1	*please let $name$ be healthy and live all year*
1	*to be emotionally stable and happy*
1	*to take over the world*

Table 1: Example wishes and their frequencies in the WISH corpus.

uct review excerpt: "*Great camera. Indoor shots with a flash are not quite as good as 35mm. I wish the camera had a higher optical zoom so that I could take even better wildlife photos.*" The first sentence contains positive opinion, the second negative opinion. However, wishful statements like the third sentence are often annotated as non-opinion-bearing in sentiment analysis corpora (Hu and Liu, 2004; Ding et al., 2008), even though they clearly contain important information. An automatic "wish detector" text-processing tool can be useful for product manufacturers, advertisers, politicians, and others looking to discover what people want.

2. Wishes can tell us a lot about people: their innermost feelings, perceptions of what they're lacking, and what they desire (Speer, 1939). Many psychology researchers have attempted to quantify the contents of wishes and how they vary with factors such as location, gender, age, and personality type (Speer, 1939; Milgram and Riedel, 1969; Ehrlichman and Eichenstein, 1992; King and Broyles, 1997). These studies have all been small scale with only dozens or hundreds of participants. The WISH corpus provides the first large-scale collection of wishes as a window into the world's desires.

Beyond sentiment analysis, classifying sentences as wishes is an instance of non-topical classification. Tasks falling under this heading include computational humor (Mihalcea and Strapparava, 2005), genre classification (Boese and Howe, 2005), authorship attribution (Argamon and Shimoni, 2003), and metaphor detection (Krishnakumaran and Zhu, 2007), among others (Mishne et al., 2007; Mihalcea and Liu, 2006). We share the common goal of classifying text into a unique set of target categories (in our case, wishful and non-wishful), but use different techniques catered to our specific classification task. Our feature-generation technique for wish detection is inspired by template-based methods for information extraction (Brin, 1999; Agichtein and Gravano, 2000).

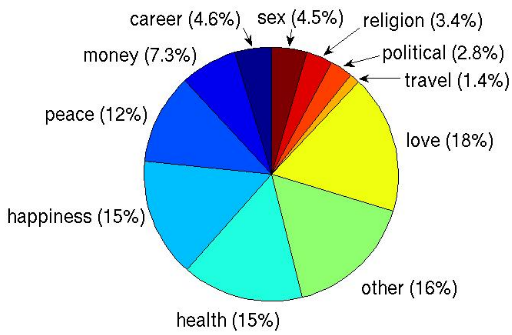
2 Analyzing the WISH Corpus

We analyze the WISH corpus with a variety of statistical methods. Our analyses not only reveal what people wished for on New Year's Eve, but also provide insight for the development of wish detectors in Section 3.

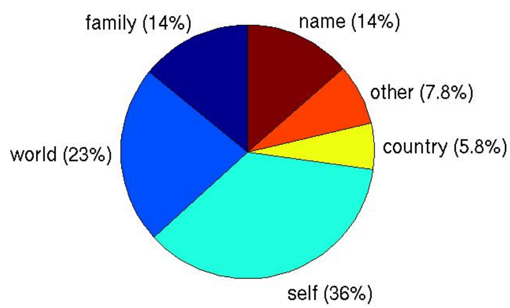
The complete WISH corpus contains nearly 100,000 wishes collected over a period of 10 days in December 2007, most written in English, with the remainder in Portuguese, Spanish, Chinese, French, and other languages. For this paper, we consider only the 89,574 English wishes. Most of these English wishes contain optional geographic meta data provided by the wisher, indicating a variety of countries (not limited to English-speaking) around the world. We perform minimal preprocessing, including TreeBank-style tokenization, downcasing, and punctuation removal. Each wish is treated as a single entity, regardless of whether it contains multiple sentences. After preprocessing, the average length of a wish is 8 tokens.

2.1 The Topic and Scope of Wishes

As a first step in understanding the content of the wishes, we asked five annotators to manually annotate a random subsample of 5,000 wishes. Sec-



(a) Topic of Wishes



(b) Scope of Wishes

Figure 1: Topic and scope distributions based on manual annotations of a random sample of 5,000 wishes in the WISH corpus.

tions 2.1 and 2.2 report results on this subsample.

The wishes were annotated in terms of two attributes: topic and scope. We used 11 pre-defined topic categories, and their distribution in this subsample of the WISH corpus is shown in Figure 1(a). The most frequent topic is *love*, while *health*, *happiness*, and *peace* are also common themes. Many wishes also fell into an *other* category, including specific individual requests ("i wish for a new puppy"), solicitations or advertisements ("call me 555-1234", "visit $website$.com"), or sinister thoughts ("to take over the world").

The 5,000 wishes were also manually assigned a scope. The scope of a wish refers to the range of people that are targeted by the wish. We used 6 pre-defined scope categories: *self* ("I want to be happy"), *family* ("For a cure for my husband"), specific person by *name* ("Prayers for $name$"), *country* ("Bring our troops home!"), *world* ("Peace to everyone in the world"), and *other*. In cases where multiple scope labels applied, the broadest scope was selected. Figure 1(b) shows the scope distribution. It is bimodal: over one third of the wishes are narrowly directed at one's self, while broad wishes at the world level are also frequent. The in-between scopes are less frequent.

2.2 Wishes Differ by Geographic Location

As mentioned earlier, wishers had the option to enter a city/country when submitting wishes. Of the manually annotated wishes, about 4,000 included valid location information, covering all 50 States in the U.S., and all continents except Antarctica.

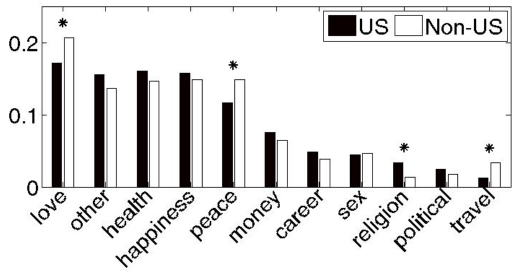
We noticed a statistically significant difference between wishes submitted from the United States (about 3600) versus non-U.S. wishes (about 400), both in terms of their topic and scope distributions. For each comparison, we performed a Pearson χ^2-test using location as the explanatory variable and either topic or scope as the response variable[2]. The null hypothesis is that the variables are independent. For both tests we reject the null hypothesis, with $p < 0.001$ for topic, and $p = 0.006$ for scope. This indicates a dependence between location and topic/scope. Asterisks in Figure 2 denote the labels that differ significantly between U.S. and non-U.S. wishes.[3]

In particular, we observed that there are significantly more wishes about *love*, *peace*, and *travel* from non-U.S. locales, and more about *religion* from the U.S.. There are significantly more *world*-scoped wishes from non-U.S. locales, and more *country*- and *family*-scoped wishes from the U.S..

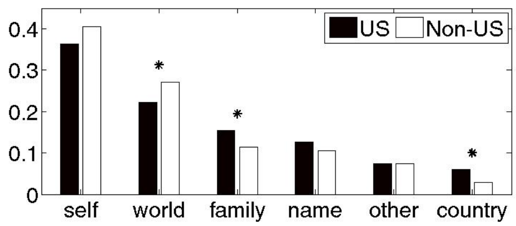
We also compared wishes from "red states" versus "blue states" (U.S. states that voted a majority

[2]The topic test examined a 2×11 contingency table, while the scope test used a 2×6 contingency table. In both tests, all of the cells in the tables had an expected frequency of at least 5, so the χ^2 approximation is valid.

[3]To identify the labels that differ significantly by location, we computed the standardized residuals for the cells in the two contingency tables. Standardized residuals are approximately $\mathcal{N}(0,1)$-distributed and can be used to locate the major contributors to a significant χ^2-test statistic (Agresti, 2002). The asterisks in Figure 2 indicate the surprisingly large residuals, i.e., the difference between observed and expected frequencies is outside a 95% confidence interval.



(a) Wish topics differ by Location



(b) Wish scopes differ by Location

Figure 2: Geographical breakdown of topic and scope distributions based on approximately 4,000 location-tagged wishes. Asterisks indicate statistically significant differences.

for the Republican and Democratic 2008 presidential candidates, respectively), but found no significant differences.

2.3 Wishes Follow Zipf's Law

We now move beyond the annotated subsample, and examine the full set of 89,574 English wishes. We noticed that a small fraction (4%) of unique wishes account for a relatively large portion (16%) of wish occurrences, while there are also many wishes that only occur once. The question naturally arise: do wishes obey Zipf's Law (Manning and Schütze, 1999)? If so, we should expect the frequency of a unique wish to be inversely proportional to its rank, when sorted by frequency. Figure 3 plots rank versus frequency on a log-log scale and reveals an approximately linear negative slope, thus suggesting that wishes do follow Zipf's law. It also shows that low-occurrence wishes dominate, hence learning might be hindered by data sparseness.

2.4 Latent Topic Modeling for Wishes

The 11 topics in Section 2.1 were manually predefined based on domain knowledge. In contrast,

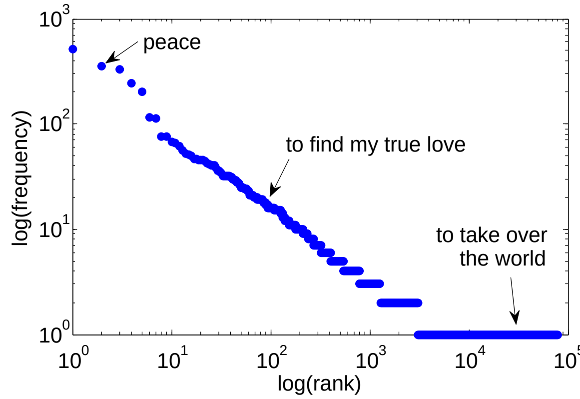


Figure 3: The rank vs. frequency plot of wishes, approximately obeying Zipf's law. Note the log-log scale.

in this section we applied Latent Dirichlet Allocation (LDA) (Blei et al., 2003) to identify the latent topics in the full set of 89,574 English wishes in an unsupervised fashion. The goal is to validate and complement the study in Section 2.1.

To apply LDA to the wishes, we treated each individual wish as a short document. We used 12 topics, Collapsed Gibbs Sampling (Griffiths and Steyvers, 2004) for inference, hyperparameters $\alpha = 0.5$ and $\beta = 0.1$, and ran Markov Chain Monte Carlo for 2000 iterations.

The resulting 12 LDA topics are shown in Table 2, in the form of the highest probability words $p(\text{word}|\text{topic})$ in each topic. We manually added summary descriptors for readability. With LDA, it is also possible to observe which words were assigned to which topics in each wish. For example, LDA assigned most words in the wish "world(8) peace(8) and my friends(4) in iraq(1) to come(1) home(1)" to two topics: peace and troops (topic numbers in parentheses). Interestingly, these LDA topics largely agree with the pre-defined topics in Section 2.1.

3 Building Wish Detectors

We now study the novel NLP task of wish detection, i.e., classifying individual sentences as being wishes or not. Importantly, we want our approach to work on target domains other than New Year's wishes, including consumer product reviews and online political discussions. It should be pointed out that wishes are highly domain dependent. For example, "I wish for world peace" is a common wish on New Year's

Topic	Summary	Top words in the topic, sorted by $p(\text{word}\mid\text{topic})$
0	New Year	year, new, happy, 2008, best, everyone, great, years, wishing, prosperous, may, hope
1	Troops	all, god, home, come, may, safe, s, us, bless, troops, bring, iraq, return, 2008, true, dreams
2	Election	wish, end, no, more, 2008, war, stop, president, paul, not, ron, up, free, less, bush, vote
3	Life	more, better, life, one, live, time, make, people, than, everyone, day, wish, every, each
4	Prosperity	health, happiness, good, family, friends, all, love, prosperity, wealth, success, wish, peace
5	Love	love, me, find, wish, true, life, meet, want, man, marry, call, someone, boyfriend, fall, him
6	Career	get, wish, job, out, t, hope, school, better, house, well, want, back, don, college, married
7	Lottery	wish, win, 2008, money, want, make, become, lottery, more, great, lots, see, big, times
8	Peace	peace, world, all, love, earth, happiness, everyone, joy, may, 2008, prosperity, around
9	Religion	love, forever, jesus, know, loves, together, u, always, 2, 3, 4, much, best, mom, christ
10	Family	healthy, happy, wish, 2008, family, baby, life, children, long, safe, husband, stay, marriage
11	Health	com, wish, s, me, lose, please, let, cancer, weight, cure, mom, www, mother, visit, dad

Table 2: Wish topics learned from Latent Dirichlet Allocation. Words are sorted by $p(\text{word}\mid\text{topic})$.

Eve, but is exceedingly rare in product reviews; and vice versa: "I want to have instant access to the volume" may occur in product reviews, but is an unlikely New Year's wish. For this initial study, we do assume that there are some labeled training data in the target domains of interest.

Our focus, however, is on transferring the knowledge learned from the out-of-domain WISH corpus to other domains. The key insight is the following: while the content of wishes (e.g., "world peace") may not transfer across domains, the ways wishes are expressed (e.g., "I wish for ___") may. We call these expressions *wish templates*. Our novel contribution is an unsupervised method for discovering candidate templates from the WISH corpus which, when applied to other target domains, improve wish detection in those domains.

3.1 Two Simple Wish Detectors

Before describing our template discovery method, we first describe two simple wish detectors, which serve as baselines.

1. **[Manual]**: It may seem easy to locate wishes. Perhaps looking for sentences containing the phrases "i wish," "i hope," or some other simple patterns is sufficient for identifying the vast majority of wishes in a domain. To test this hypothesis, we asked two native English speakers (not the annotators, nor affiliated with the project; no exposure to any of the wish datasets) to come up with text patterns that might be used to express wishes. They were shown three dictionary definitions of "to wish (v)" and "wish (n)". They produced a ranked list of 13 templates; see Table 3. The underscore matches any string. These templates can be turned into a simple rule-based classifier: If part of a sentence matches one of the templates, the sentence is classified as a wish. By varying the depth of the list, one can produce different precision/recall behaviors. Overall, we expect [Manual] to have relatively high precision but low recall.

i wish ___
i hope ___
i want ___
hopefully ___
if only ___
would be better if ___
would like ___ if ___
___ should ___
would that ___
can't believe ___ didn't ___
don't believe ___ didn't ___
___ do want
i can has ___

Table 3: Manually-created templates for identifying wishes.

2. **[Words]**: Another simple method for detecting wishes is to train a standard word-based text classifier using the labeled training set in the target domain. Specifically, we represent each sentence as a binary word-indicator vector, normalized to sum to 1. We then train a linear Support Vector Machine (SVM). [Words] may have higher recall, but precision may suffer. For instance, the sentence "Her wish was carried out by her husband" is not a

wish, but could be misclassified as one because of the word “wish”.

Note that neither of the two baseline methods uses the WISH corpus.

3.2 Automatically Discovering Wish Templates

We now present our method to automatically discover high quality wish templates using the WISH corpus. The key idea is to exploit redundancy in how the same wish content is expressed. For example, as we see in Table 1, both “world peace” and “i wish for world peace” are common wishes. Similarly, both “health and happiness” and “i wish for health and happiness” appear in the WISH corpus. It is thus reasonable to speculate that “i wish for ___” is a good wish template. Less obvious templates can be discovered in this way, too, such as “let there be ___” from “peace on earth” and “let there be peace on earth.”

We formalize this intuition into a bipartite graph, illustrated in Figure 4. Let $W = \{w_1, \ldots, w_n\}$ be the set of unique wishes in the WISH corpus. The bipartite graph has two types of nodes: content nodes C and template nodes T, and they are generated as follows. If a wish w_j (e.g., “i wish for world peace”) contains another wish w_i (e.g., “world peace”), we create a content node $c = w_i$ and a template node t (e.g., “i wish for ___”). We denote this relationship by $w_j = c + t$. Note the order of c and t is insignificant, as how the two combine is determined by the underscore in t, and $w_j = t + c$ is just fine. In addition, we place a directed edge from c to t with edge weight $\text{count}(w_j)$, the frequency of wish w_j in the WISH corpus. Then, a template node appears to be a good one if many heavy edges point to it.

On the other hand, a template is less desirable if it is part of a content node. For example, when $w_j =$“health and happiness” and $w_i =$“health”, we create the template $t =$“___ and happiness” and the content node $c = w_i$. If there is another wish $w_k =$“i wish for health and happiness”, then there will be a content node $c' = w_j$. The template t thus contains some content words (since t matches c'), and may not generalize well in a new domain. We capture this by backward edges: if $\exists c' \in C$, and $\exists$ string s (s not necessarily in C or W) such that $c' = s + t$, we add a backward edge from t to c' with edge weight $\text{count}(c')$.

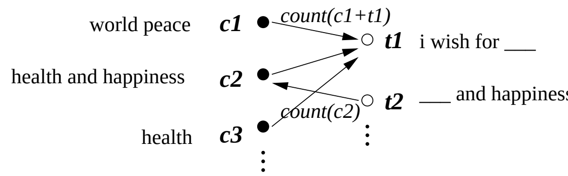


Figure 4: The bipartite graph to create templates.

Based on such considerations, we devised the following scheme for scoring templates:

$$\text{score}(t) = \text{in}(t) - \text{out}(t), \tag{1}$$

where $\text{in}(t)$ is the in-degree of node t, defined as the sum of edge weights coming into t; $\text{out}(t)$ is the out-degree of node t, defined similarly. In other words, a template receives a high score if it is “used” by many frequent wishes but does not match many frequent content-only wishes. To create the final set of template features, we apply the threshold $\text{score}(t) \geq 5$. This produces a final list of 811 templates. Table 4 lists some of the top templates ranked by $\text{score}(t)$. While some of these templates still contain time- or scope-related words (“for my family”), they are devoid of specific topical content. Notice that we have automatically identified several of the manually derived templates in Table 3, and introduce many new variations that a learning algorithm can take advantage of.

Top 10	Others in Top 200
___ in 2008	i want to ___
i wish for ___	___ for everyone
i wish ___	i hope ___
i want ___	my wish is ___
___ this year	please ___
i wish ___ in 2008	wishing for ___
i wish to ___	may you ___
___ for my family	i wish i had ___
i wish ___ this year	to finally ___
___ in the new year	for my family to have ___

Table 4: Top templates according to Equation 1.

3.3 Learning with Wish Template Features

After discovering wish templates as described above, we use them as features for learning in a new domain (e.g., product reviews). For each sentence in the new domain, we assign binary features indicating which templates match the sentence. Two types

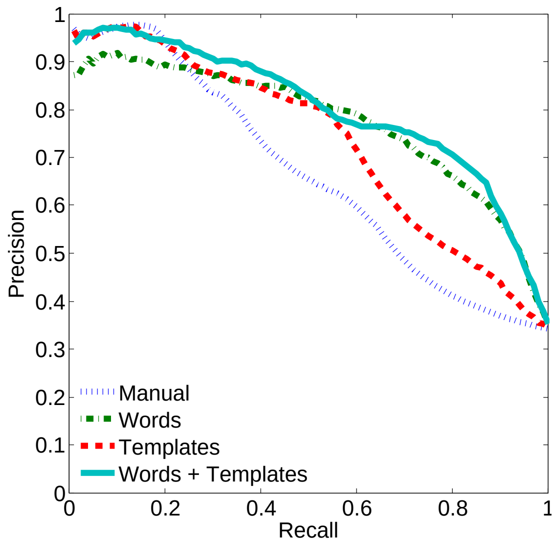


Figure 5: Politics domain precision-recall curves.

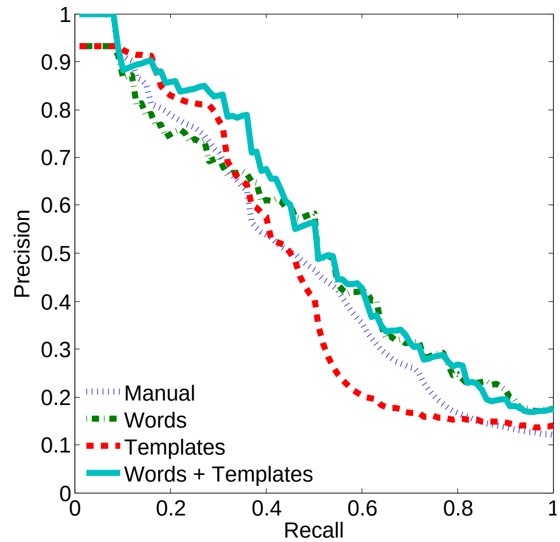


Figure 6: Products domain precision-recall curves.

of matching are possible. *Strict matching* requires that the template must match an entire sentence from beginning to end, with at least one word filling in for the underscore. (All matching during the template generation process was strict.) *Non-strict matching* requires only that template match somewhere within a sentence. Rather than choose one type of matching, we create both strict and non-strict template features (1622 binary features total) and let the machine learning algorithm decide what is most useful.

Our third wish detector, **[Templates]**, is a linear SVM with the 1622 binary wish template features. Our fourth wish detector, **[Words + Templates]**, is a linear SVM with both template and word features.

4 Experimental Results

4.1 Target Domains and Experimental Setup

We experimented with two domains, manually labeled at the sentence-level as wishes or non-wishes. Example wishes are listed in Table 6.

Products. Consumer product reviews: 1,235 sentences selected from a collection of amazon.com reviews (Hu and Liu, 2004; Ding et al., 2008). 12% of the sentences are labeled as wishes.

Politics. Political discussion board postings: 6,379 sentences selected from politics.com (Mullen and Malouf, 2008). 34% are labeled as wishes.

We automatically split the corpora into sentences using MxTerminator (Reynar and Ratnaparkhi, 1997). As preprocessing before learning, we tokenized the text in the Penn TreeBank style, downcased, and removed all punctuation.

For all four wish detectors, we performed 10-fold cross validation. We used the default parameter in SVM^{light} for all trials (Joachims, 1999). As the data sets are skewed, we compare the detectors using precision-recall curves and the area under the curve (AUC). For the manual baseline, we produce the curve by varying the number of templates applied (in rank order), which gradually predicts more sentences as wishes (increasing recall at the expense of precision). A final point is added at recall 1 corresponding to applying an empty template that matches all sentences. For the SVM-based methods, we vary the threshold applied to the real-valued margin prediction to produce the curves. All curves are interpolated, and AUC measures are computed, using the techniques of (Davis and Goadrich, 2006).

4.2 Results

Figure 5 shows the precision-recall curves for the Politics corpus. All curves are averages over 10 folds (i.e., for each of 100 evenly spaced, interpolated recall points, the 10 precision values are averaged). As expected, [Manual] can be very precise with low recall—only the very top few templates achieve high precision and pick out a small number of wishes with “i wish” and “i hope.” As we

Corpus	[Manual]	[Words]	[Templates]	[Words + Templates]
Politics	0.67 ± 0.03	0.77 ± 0.03	0.73 ± 0.03	0.80 ± 0.03
Products	0.49 ± 0.13	0.52 ± 0.16	0.47 ± 0.16	0.56 ± 0.16

Table 5: AUC results (10-fold averages $\pm$ one standard deviation).

Products:
the only area i wish apple had improved upon would be the screen
i just want music to eminate from it when i want how i want
the dial on the original zen was perfect and i wish it was on this model
i would like album order for my live albums and was just wondering
Politics:
all children should be allowed healthcare
please call on your representatives in dc and ask them to please stop the waste in iraq
i hope that this is a new beginning for the middle east
may god bless and protect the brave men and that we will face these dangers in the future

Table 6: Example target-domain wishes correctly identified by [Words + Templates].

introduce more templates to cover more true wishes, precision drops off quickly. [Templates] is similar, with slightly better precision in low recall regions. [Words] is the opposite: bad in high recall but good in low recall regions. [Words + Templates] is the best, taking the best from both kinds of features to dominate other curves. Table 5 shows the average AUC across 10 folds. [Words + Templates] is significantly better than all other detectors under paired t-tests ($p = 1 \times 10^{-7}$ vs. [Manual], $p = 0.01$ vs. [Words], and $p = 4 \times 10^{-7}$ vs. [Templates]). All other differences are statistically significant, too.

Figure 6 shows the precision-recall curves for the Products corpus. Again, [Words + Templates] mostly dominates other detectors. In terms of average AUC across folds (Table 5), [Words + Templates] is also the best. However, due to the small size of this corpus, the AUC values have high variance, and the difference between [Words + Templates] and [Words] is not statistically significant under a paired t-test ($p = 0.16$).

Finally, to understand what is being learned in more detail, we take a closer look at the SVM models' weights for one fold of the Products corpus (Table 7). The most positive and negative features make intuitive sense. Note that [Words + Templates] seems to rely on templates for selecting wishes and words for excluding non-wishes. This partially explains the synergy of combining the feature types.

Sign	[Words]	[Templates]	[Words + Templates]
+	wish	i hope ___	hoping ___
+	hope	i wish ___	i hope ___
+	hopefully	hoping ___	i just want ___
+	hoping	i just want ___	i wish ___
+	want	i would like ___	i would like ___
-	money	family ___	micro
-	find	___ forever	about
-	digital	let me ___	fix
-	again	___ d	digital
-	you	___ for my dad	you

Table 7: Features with the largest magnitude weights in the SVM models for one fold of the Products corpus.

5 Conclusions and Future Work

We have presented a novel study of wishes from an NLP perspective. Using the first-of-its-kind WISH corpus, we presented a technique for generating domain-independent wish template features that improve wish detection performance across product reviews and political discussion posts. Much work remains in this new research area, including the careful examination of more types of features, as well as methods that do not require as much labeled training data for domain transfer.

References

Eugene Agichtein and Luis Gravano. 2000. Snowball: Extracting relations from large plain-text collections. In *In Proceedings of the 5th ACM International Conference on Digital Libraries*, pages 85–94.

Alan Agresti. 2002. *Categorical Data Analysis*. Wiley-Interscience, second edition.

Shlomo Argamon and Anat Rachel Shimoni. 2003. Automatically categorizing written texts by author gender. *Literary and Linguistic Computing*, 17:401–412.

David M. Blei, Andrew Y. Ng, and Michael I. Jordan. 2003. Latent dirichlet allocation. *Journal of Machine Learning Research*, 3:993–1022.

Elizabeth Sugar Boese and Adele Howe. 2005. Genre classification of web documents. In *Proceedings of the 20th National Conference on Artificial Intelligence (AAAI-05), Poster paper*.

Sergey Brin. 1999. Extracting patterns and relations from the world wide web. In *WebDB '98: Selected papers from the International Workshop on The World Wide Web and Databases*, pages 172–183. Springer-Verlag.

Jesse Davis and Mark Goadrich. 2006. The relationship between precision-recall and roc curves. In *ICML '06: Proceedings of the 23rd international conference on Machine learning*, New York, NY, USA. ACM.

Xiaowen Ding, Bing Liu, and Philip S. Yu. 2008. A holistic lexicon-based approach to opinion mining. In *WSDM '08: Proceedings of the international conference on Web search and web data mining*, pages 231–240. ACM.

Howard Ehrlichman and Rosalind Eichenstein. 1992. Private wishes: Gender similarities and difference. *Sex Roles*, 26(9):399–422.

Thomas Griffiths and Mark Steyvers. 2004. Finding scientific topics. *Proceedings of the National Academy of Sciences*, 101(suppl. 1):5228–5235.

Minqing Hu and Bing Liu. 2004. Mining and summarizing customer reviews. In *Proceedings of KDD '04, the ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 168–177. ACM Press.

Thorsten Joachims. 1999. Making large-scale svm learning practical. In B. Schölkopf, C. Burges, and A. Smola, editors, *Advances in Kernel Methods - Support Vector Learning*. MIT Press.

Laura A. King and Sheri J. Broyles. 1997. Wishes, gender, personality, and well-being. *Journal of Personality*, 65(1):49–76.

Moshe Koppel and Itai Shtrimberg. 2004. Good news or bad news? let the market decide. In *AAAI Spring Symposium on Exploring Attitude and Affect in Text*, pages 86–88.

Saisuresh Krishnakumaran and Xiaojin Zhu. 2007. Hunting elusive metaphors using lexical resources. In *Proceedings of the Workshop on Computational Approaches to Figurative Language*, pages 13–20, Rochester, New York, April. Association for Computational Linguistics.

Christopher D. Manning and Hinrich Schütze. 1999. *Foundations of Statistical Natural Language Processing*. The MIT Press, Cambridge, Massachusetts.

Rada Mihalcea and Hugo Liu. 2006. A corpus-based approach to finding happiness. In *Proceedings of AAAI-CAAW-06, the Spring Symposia on Computational Approaches to Analyzing Weblogs*.

Rada Mihalcea and Carlo Strapparava. 2005. Making computers laugh: Investigations in automatic humor recognition. In *Empirical Methods in Natural Language Processing*.

Norman A. Milgram and Wolfgang W. Riedel. 1969. Developmental and experiential factors in making wishes. *Child Development*, 40(3):763–771.

Gilad Mishne, Krisztian Balog, Maarten de Rijke, and Breyten Ernsting. 2007. Moodviews: Tracking and searching mood-annotated blog posts. In *Proceedings International Conf. on Weblogs and Social Media (ICWSM-2007)*, pages 323–324.

Tony Mullen and Robert Malouf. 2008. Taking sides: User classification for informal online political discourse. *Internet Research*, 18:177–190.

Bo Pang and Lillian Lee. 2008. Opinion mining and sentiment analysis. *Foundations and Trends in Information Retrieval*, 2(1-2):1–135.

Jeffrey C. Reynar and Adwait Ratnaparkhi. 1997. A maximum entropy approach to identifying sentence boundaries. In *Fifth Conference on Applied Natural Language Processing*.

James Shanahan, Yan Qu, and Janyce Wiebe, editors. 2005. *Computing attitude and affect in text*. Springer, Dordrecht, The Netherlands.

George S. Speer. 1939. Oral and written wishes of rural and city school children. *Child Development*, 10(3):151–155.