Intention Analysis for Sales, Marketing and Customer Service

Cohan Sujay Carlos[1] *Madhulika Yalamanchi*[1]
(1) Aiaioo Labs, India
cohan@aiaioo.com, madhulika@aiaioo.com

Abstract

In recent years, social media has become a customer touch-point for the business functions of marketing, sales and customer service. We aim to show that intention analysis might be useful to these business functions and that it can be performed effectively on short texts (at the granularity level of a single sentence). We demonstrate a scheme of categorization of intentions that is amenable to automation using simple machine learning techniques that are language-independent. We discuss the grounding that this scheme of categorization has in speech act theory. In the demonstration we go over a number of usage scenarios in an attempt to show that the use of automatic intention detection tools would benefit the business functions of sales, marketing and service. We also show that social media can be used not just to convey pleasure or displeasure (that is, to express sentiment) but also to discuss personal needs and to report problems (to express intentions). We evaluate methods for automatically discovering intentions in text, and establish that it is possible to perform intention analysis on social media with an accuracy of $66.97\% \pm 0.10\%$.

Keywords: intention analysis, intent analysis, social media, speech act theory, sentiment analysis, emotion analysis, intention.

1 Introduction

In this paper and the accompanying demonstration, we present and attempt to demonstrate the effectiveness of a method of categorization of intentions that is based on the needs of the marketing, sales and service functions of a business which are, according to Smith et al. (2011), the functions most impacted by social media. The categories of intention that we use are *purchase*, *inquire*, *complain*, *criticise*, *praise*, *direct*, *quit*, *compare*, *wish* and *sell*. We also use an *other* category consisting of sentences that do not express intentions.

In the demonstration, we show that the intention categories *purchase*, *sell* and *wish* are valuable to sales, that the *inquire* category can be used for outbound marketing, that *criticise*, *compare* and *praise* can be used for inbound marketing, and that *complain*, *direct* and *quit* can be used for customer service.

This does not mean that these categories are only of use to business. The intention to *complain* and the intention to *quit* have been studied extensively by Hirschman (1970) in the context of a wide range of social, political and economic phenomena. A game theoretic framework for the work of Hirschman (1970) has been proposed by Gehlbach (2006) and used to model the mechanism of collapse of communism in East Germany.

In Section 2 we describe the theoretical underpinnings of the present work and in Section 3 we go over related research. In Section 4 we discuss the quantity of social media messages that contain the categories of intentions that are the subject of the present study (we compare the quantities of intentions expressed with the quantities of expressions of sentiment). In Section 5 we describe and evaluate machine learning algorithms for automated intention analysis.

2 Background

2.1 Speech Act Theory

Austin (1975), in the theory of *speech acts*, distinguished between utterances that are statements (whose truth or falsity is verifiable) and utterances that are not statements. He observed that, "there are, traditionally, besides (grammarians') statements, also questions and exclamations, and sentences expressing commands or wishes or concessions."

In our work we deal with certain types of speech acts that can be called 'intentions' according to one common dictionary definition of the word 'intention', which is, "an aim or plan". In particular, we focus on the ten categories of intention (excluding *other*) in Table 1.

Another concept from speech act theory (Searle, 1983) is the 'direction of fit' of a speech act or *intentional state*. The direction of fit is said to be 'mind-to-world' if through the performance of the speech act, a mental state is established, revealed or altered. The direction of fit of a speech act or intentional state is said to be 'world-to-mind' if the performance of the speech act alters the state of the world.

Seven of the ten categories of intentions in our annotation scheme have the world-to-mind direction of fit (they are desires or intentions) and three have the mind-to-world direction of fit (beliefs). The three categories that have the mind-to-world direction of fit correspond to categories used in opinion mining (namely 'praise', 'criticize' and 'compare').

2.2 Discourse Theory

In the introducton to the collection "Intentions in Communication" Cohen et al. (1990) suggest that any theory that purports to explain communication and discourse "will have to place a strong emphasis on issues of *intention*". To illustrate the point, they offer a sample dialog between a customer looking for some meat and a butcher selling the same:

- Customer: "Where are the chuck steaks you advertised for 88 cents per pound?"
- Butcher: "How many do you want?"

The butcher's response would be perfectly natural in a scenario where the steaks are behind the counter where customers are not allowed, and the plausibility of this conversation shows that people infer intention, just as the butcher infers the intention of the customer to be a purchase intention (in this case, possibly as much from the context as from the language).

Georgeff et al. (1999) discuss the Belief-Desire-Intention (BDI) Model of Agency based on the work of Bratman (1987). In the present work, the term "intentions" loosely corresponds to the sense of "desire" as well as "intention" in the BDI model.

3 Related Research

3.1 Wishes in Reviews and Discussions

Goldberg et al. (2009) developed a corpus of wishes from a set of New Year's Day wishes and through evaluation of learning algorithms for the domains '*products*' and '*politics*', showed that even though the content of wishes might be domain-specific, the manner in which wishes are expressed is not entirely so. The definition of the word 'wish' used by Goldberg et al. (2009) is "a desire or hope for something to happen".

The wish to *purchase* and the wish to *suggest* improvements are studied in Ramanand et al. (2010). Ramanand et al. (2010) propose rules for identifying both kinds of wishes and test the collection of rules using a corpus that includes product reviews, customer surveys and comments from consumer forums. In addition, they evaluate their system on the *WISH corpus* of Goldberg et al. (2009). Wu and He (2011) also study the wish to suggest and the wish to purchase using variants of Class Sequential Rules (CSRs).

3.2 Requests and Promises in Email

Lampert et al. (2010) study the identification of requests in email messages and obtain an accuracy of 83.76%. A study of email communications by Carvalho and Cohen (2006) and Cohen et al. (2004) focuses on discovering speech acts in email, building upon earlier work on illocutionary speech acts (Searle, 1975; Winograd, 1987).

3.3 Speech Acts in Conversations

Bouchet (2009) describes the construction of a corpus of user requests for assistance, annotated with the illocutionary speech acts *assertive*, *commissive*, *directive*, *expressive*, *declarative*, and an *other* category for utterances that cannot be classified into one of those. Ravi and Kim (2007) use rules to identify threads that may have unanswered questions and therefore require instructor attention. In their approach, each message is classified as a *question*, *answer*, *elaboration* and *correction*.

3.4 Sentiment and Emotion

Three of the intentions in the present study, namely the intention to *praise* something, to *criticize* something, and to *compare* something with something else, have been studied by researchers in connection with sentiment analysis.

The detection of comparisons in text has been studied by Jindal and Liu (2006), and the use of comparative sentences in opinion mining has been studied by Ganapathibhotla and Liu (2008). Yang and Ko (2011) proposed a method to automatically identify 7 categories of comparatives in Korean. Li et al. (2010) used a weakly supervised method to identify comparative questions from a large online question archive. Different perspectives might be reflected in contrastive opinions, and these are studied by Fang et al. (2012) in the context of political texts using the Cross-Perspective Topic model.

The mining of opinion features and the creation of review summaries is studied in Hu and Liu (2006, 2004). A study of sentiment classification is reported in Pang et al. (2002), and the use of subjectivity detection in sentiment classification is reported in Pang and Lee (2004).

Studies to detect emotions in internet chat conversations have been described in Wu et al. (2002); Holzman and Pottenger (2003); Shashank and Bhattacharyya (2010). Minato et al. (2008) describe the creation of an emotions corpus in the Japanese language. Vidrascu and Devillers (2005) attempt to detect emotions in speech data from call center recordings.

4 Distribution of Intentions

Table 1 lists the categories of intentions that are the subject of the present study, their mapping to concepts from speech act theory, namely direction of fit, intentional state (desire/belief) and illocutionary point, and their counts in a corpus of sentences from social media.

Intention	Direction of fit	Des/Bel	Illocution	Business Fn	Count
wish	world-to-mind	desire	directive	marketing	543
purchase	world-to-mind	desire	directive	sales	2221
inquire	world-to-mind	desire	directive	marketing	2972
compare	mind-to-world	belief	representative	research	508
praise	mind-to-world	belief	representative	research	1574
criticize	mind-to-world	belief	representative	research	2031
complain	world-to-mind	desire	representative	service	2107
quit	world-to-mind	desire	commissive	service	744
direct	world-to-mind	desire	directive	service	706
sell	world-to-mind	desire	directive	procurement	524
other					2775

Table 1: Categories annotated in the corpus.

Only 4113 sentences belonged to categories related to opinion (praise, criticize and compare), demonstrating that other speech acts are prevalent on social media in certain contexts.

5 Experimental Evaluation

A set of experiments was performed using naive bayes classification, maximum entropy classification, and support vector machine classification to see if intention analysis could be automated, and to see what features might be used to tell categories of intentions apart.

5.1 Corpus Slices

The experiments were performed using three slices of categories from the corpus. The first slice (Slice 1) consisted of the categories purchase, inquire, complain, criticize, praise and other, (6 categories) all of which number greater than 1500 in the corpus. The second slice (Slice 2) consisted of direct and quit (both of which have more than 700 each in the corpus) in addition to the above categories, for a total of 8 categories. The last slice (Slice 3) consisted of sell, compare and wish (which have more than 500 occurrences each in the corpus) in addition to the 8 categories mentioned above, for a total of 11 categories.

5.2 Automatic Classification

Naive bayesian (NB) classifiers, maximum entropy (ME) classifiers, and support vector machine (SVM) classifiers were evaluated on the corpus of intentions. The features used were n-grams (all n-grams containing keywords used to crawl the social media text were discarded).

Features	**NB**	**ME**	**SVM (RBF)**
unigrams	60.97 ± 0.01	68.24 ± 0.02	68.96 ± 0.02
bigrams	60.07 ± 0.02	65.38 ± 0.01	65.19 ± 0.01
unigrams+bigrams	64.07 ± 0.02	70.43 ± 0.02	69.37 ± 0.02

Table 2: Average five-fold cross-validation accuracies on Slice 1 (sentence order randomized).

Features	**NB**	**ME**	**SVM (RBF)**
unigrams	51.18 ± 0.02	53.06 ± 0.01	58.96 ± 0.02
bigrams	52.14 ± 0.02	54.89 ± 0.01	52.96 ± 0.01
unigrams+bigrams	56.66 ± 0.02	60.71 ± 0.02	57.95 ± 0.01

Table 3: Average five-fold cross-validation accuracies on Slice 2 (sentence order randomized).

Features	**NB**	**ME**	**SVM (RBF)**
unigrams	46.40 ± 0.01	53.06 ± 0.01	52.99 ± 0.02
bigrams	46.94 ± 0.01	50.01 ± 0.01	48.18 ± 0.02
unigrams+bigrams	51.45 ± 0.01	55.43 ± 0.02	52.62 ± 0.02

Table 4: Average five-fold cross-validation accuracies on Slice 3 (sentence order randomized).

Accuracy scores for Slices 1, 2 and 3 are listed in Table 2, Table 3 and Table 4 and Table 5.

6 Demonstration

We will demonstrate the use of intention analysis in a number of usage scenarios to establish its value to sales, marketing and customer service.

6.1 Identifying Leads for Sales

The ability to find customers who have a need for a particular product or service is valuable to the sales function of a business. We demonstrate how customers who wish to buy certain products may be identified by monitoring conversations on social media.

Features	NB	ME	SVM (RBF)
unigrams	57.91 ± 0.10	65.27 ± 0.11	65.96 ± 0.09
bigrams	56.61 ± 0.06	62.22 ± 0.08	61.78 ± 0.09
unigrams+bigrams	59.97 ± 0.08	66.97 ± 0.10	65.57 ± 0.09

Table 5: Average 5-fold cross-validation accuracies on Slice 1 of the unshuffled corpus.

6.2 Identifying Needs for Marketing

Marketing can use inquiries on social media to identify interested persons and educate them about pertinent offerings. Political teams can use inquiries to educate voters. They can also use intentions expressed on social media to identify needs and wants. In this segment of the demonstration, we show how inquiries about a product or service, and expressions of interest may be detected.

6.3 Identifying Issues for Customer Service

Customer service might be able to better respond to criticism and complaints if it can spot customers who are dissatisfied or have problems. In this segment of the demonstration, we show how complaints and criticism of a product or a service may be detected.

7 Conclusion

In this study, we have proposed a way of categorizing text in terms of the intentions expressed. We have argued that such a set of categories might be useful to numerous business functions. We have shown that these categories are encountered frequently on social media, and demonstrated the value of using intention analysis in marketing, sales and customer service scenarios. Furthermore, we have shown that it is possible to achieve an accuracy of $66.97\% \pm 0.10\%$ at the task of classifying sentence-length texts into the intention categories described in this paper.

Acknowledgements

We are very grateful to the team of Chakri J. Prabhakar, Jonas Prabhakar, Noopura Srihari and Shachi Ranganath for their patient, careful and painstaking (and underpaid and very under-rewarded) development of the corpus that was used in these experiments. We are also in the debt of Vijay Ramachandran and Rohit Chauhan, the founders of WisdomTap.com, for paying us to start working on intention analysis, for sharing with us a number of novel ideas on the subject of purchase intention analysis and its applications, and for their help and support of our research work and our efforts to build a corpus for intention analysis. We are also sincerely grateful to the anonymous reviewers of an earlier and longer version of this paper for their valuable comments and suggestions.

References

Austin, J. L. (1975). *How to Do Things With Words*. Harvard University Press, Cambridge, MA.

Bouchet, F. (2009). Characterization of conversational activities in a corpus of assistance requests. In Icard, T., editor, *Proceedings of the 14th Student Session of the European Summer School for Logic, Language, and Information (ESSLLI)*, pages 40–50.

Bratman, M. E. (1987). *Intention, Plans, and Practical Reason*. Harvard University Press, Cambridge, MA.

Carvalho, V. R. and Cohen, W. W. (2006). Improving email speech act analysis via n-gram selection. In *Proceedings of the HLT/NAACL 2006 (Human Language Technology conference - North American chapter of the Association for Computational Linguistics) - ACTS Workshop*, New York City, NY.

Cohen, P. R., Pollack, M. E., and Morgan, J. L. (1990). *Intentions in Communication*. The MIT Press.

Cohen, W. W., Carvalho, V. R., and Mitchell, T. M. (2004). Learning to classify email into "speech acts". In Lin, D. and Wu, D., editors, *Proceedings of EMNLP 2004*, pages 309–316, Barcelona, Spain. Association for Computational Linguistics.

Fang, Y., Si, L., Somasundaram, N., and Yu, Z. (2012). Mining contrastive opinions on political texts using cross-perspective topic model. In Adar, E., Teevan, J., Agichtein, E., and Maarek, Y., editors, *WSDM*, pages 63–72. ACM.

Ganapathibhotla, M. and Liu, B. (2008). Mining opinions in comparative sentences. In Scott, D. and Uszkoreit, H., editors, *COLING*, pages 241–248.

Gehlbach, S. (2006). A formal model of exit and voice. *Rationality and Society*, 18(4):1043–4631.

Georgeff, M., Pell, B., Pollack, M., Tambe, M., and Wooldridge, M. (1999). The belief-desire-intention model of agency. pages 1–10. Springer.

Goldberg, A. B., Fillmore, N., Andrzejewski, D., Xu, Z., Gibson, B., and Zhu, X. (2009). May all your wishes come true: a study of wishes and how to recognize them. In *NAACL '09 Proceedings of Human Language Technologies: The 2009 Annual Conference of the North American Chapter of the Association for Computational Linguistics*, pages 263–271, Stroudsburg, PA. Association for Consumer Research.

Hirschman, A. O. (1970). *Exit, Voice, and Loyalty - Responses to Decline in Firms, Organizations, and States*. Harvard University Press, Cambridge, MA.

Holzman, L. E. and Pottenger, W. M. (2003). Classification of emotions in internet chat: An application of machine learning using speech phonemes. 2003. available on www: www.lehigh.edu/ leh7/papers/emotionclassification.p df. Technical report, Lehigh University.

Hu, M. and Liu, B. (2004). Mining and summarizing customer reviews. In Kim, W., Kohavi, R., Gehrke, J., and DuMouchel, W., editors, *KDD*, pages 168–177. ACM.

Hu, M. and Liu, B. (2006). Opinion feature extraction using class sequential rules. In *AAAI Spring Symposium: Computational Approaches to Analyzing Weblogs*, pages 61–66.

Jindal, N. and Liu, B. (2006). Mining comparative sentences and relations. In *AAAI*, pages 1331–1336. AAAI Press.

Lampert, A., Dale, R., and Paris, C. (2010). Detecting emails containing requests for action. In *HLT '10 Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics*, Stroudsburg, PA.

Li, S., Lin, C.-Y., Song, Y.-I., and Li, Z. (2010). Comparable entity mining from comparative questions. In Hajic, J., Carberry, S., and Clark, S., editors, *ACL*, pages 650–658. The Association for Computer Linguistics.

Minato, J., Bracewell, D. B., Ren, F., and Kuroiwa, S. (2008). Japanese emotion corpus analysis and its usefor automatic emotion word identification. *Engineering Letters*, 16(1):172–177.

Pang, B. and Lee, L. (2004). A sentimental education: Sentiment analysis using subjectivity summarization based on minimum cuts. In Scott, D., Daelemans, W., and Walker, M. A., editors, *ACL*, pages 271–278. ACL.

Pang, B., Lee, L., and Vaithyanathan, S. (2002). Thumbs up? sentiment classification using machine learning techniques. *CoRR*, cs.CL/0205070.

Ramanand, J., Bhavsar, K., and Pedanekar, N. (2010). Wishful thinking: finding suggestions and 'buy' wishes from product reviews. In *CAAGET '10 Proceedings of the NAACL HLT 2010 Workshop on Computational Approaches to Analysis and Generation of Emotion in Text*, Stroudsburg, PA.

Ravi, S. and Kim, J. (2007). Profiling student interactions in threaded discussions with speech act classifiers. In *Proceedings of the 2007 Conference on Artificial Intelligence in Education: Building Technology Rich Learning Contexts That Work*, pages 357–364, Amsterdam, The Netherlands. IOS Press.

Searle, J. R. (1975). A taxonomy of illocutionary acts. *Language, Mind and Knowledge*, pages 344–369.

Searle, J. R. (1983). *Intentionality: An Essay in the Philosophy of Mind*. Cambridge University Press, Cambridge, MA.

Shashank and Bhattacharyya, P. (2010). Emotion analysis of internet chat. In *In the Proceedings of the ICON Conference 2010*.

Smith, N., Wollen, R., and Zhou, C. (2011). *The Social Media Management Handbook*. John Wiley and Sons, Inc., Hoboken, NJ.

Vidrascu, L. and Devillers, L. (2005). Detection of real-life emotions in call centers. In *INTERSPEECH*, pages 1841–1844. ISCA.

Winograd, T. (1987). A language/action perspective on the design of cooperative work. *Human-Computer Interaction*, 3(1):3–30.

Wu, T., Khan, F. M., Fisher, T. A., Shuler, L. A., and Pottenger, W. M. (2002). Posting act tagging using transformation-based learning. In *In the Proceedings of the Workshop on Foundations of Data Mining and Discovery, IEEE International Conference on Data Mining*.

Wu, X. and He, Z. (2011). Identifying wish sentence in product reviews. *Journal of Computational Information Systems*, 7(5):1607–1613.

Yang, S. and Ko, Y. (2011). Extracting comparative entities and predicates from texts using comparative type classification. In Lin, D., Matsumoto, Y., and Mihalcea, R., editors, *ACL*, pages 1636–1644. The Association for Computer Linguistics.