Mining User Experiences from Online Forums: An Exploration*

Valentin Jijkoun **Maarten de Rijke** **Wouter Weerkamp**
ISLA, University of Amsterdam
Science Park 107
1098 XG Amsterdam, The Netherlands
jijkoun,m.derijke,w.weerkamp@uva.nl

Paul Ackermans **Gijs Geleijnse**
Philips Research Europe
High Tech Campus 34
5656 AE Eindhoven, The Netherlands
paul.ackermans@philips.com
gijs.geleijnse@philips.com

1 Introduction

Recent years have shown a large increase in the usage of content creation platforms—blogs, community QA sites, forums, etc.—aimed at the general public.User generated data contains emotional, opinionated, sentimental, and personal posts. This characteristic makes it an interesting data source for exploring new types of linguistic analysis, as is demonstrated by research on, e.g., sentiment analysis [4], opinion retrieval [3], and mood detection [1].

We introduce the task of *experience mining*. Here, the goal is to gain insights into criteria that people formulate to judge or rate a product or its usage. These criteria can be formulated as the expectations that people have of the product in advance (i.e., the reasons to buy), but can also be expressed as reports of experiences while using the product and comparisons with other products. We focus on the latter: reports of experiences with products. In this paper, we define the task, describe guidelines for manual annotation and analyze linguistic features that can be used in an automatic experience mining system.

2 Motivation

Our main use-case is user-centered design for product development. User-centered design [2] is an innovation paradigm where users of a product are involved in each step of the research and development process. The first stage of the product design process is to identify unmet needs and demands of users for a specific product or a class of products. Forums, review sites, and mailing lists are platforms where people share experiences about a subject they care about. Although statements found in such platforms may not always be representative for the general user group, they can accelerate user-centered design.

Another use-case comes from online communities themselves. Users of online forums are often interested in other people's experiences with concrete products and/or solutions for specific problems. To quote one such user: *[t]he polls are the only information we have, though, except for individual [users] giving their own evaluations*. With the volume of online data increasing rapidly, users need improved access to previously reported experiences.

3 Experience mining

Experiences are particular instances of personally encountering or undergoing something. We want to identify experiences about a specific *target product*, that are *personal*, involve an *activity* related to the target and, moreover, are accompanied by *judgements or evaluative statements*. Experience mining is related to sentiment analysis and opinion retrieval, in that it involves identifying attitudes; the key difference is, however, that we are looking for *attitudes towards specific experiences* with products, not attitudes towards the products themselves.

4 An explorative study

To assess the feasibility of automatic experience mining, we carried out an explorative study: we asked human assessors to find experiences in actual forum data and then examined linguistic features likely to be useful for identifying experiences automatically.

*This research was supported by project STE-09-12 within the STEVIN programme funded by the Dutch and Flemish governments, and by the Netherlands Organisation for Scientific Research (NWO) under projects 640.001.501, 640.002.501, 612.066.512, 612.061.814, 612.061.815, 640.004.802.

Proceedings of the NAACL HLT 2010 Workshop on Computational Linguistics in a World of Social Media, pages 17–18,
Los Angeles, California, June 2010.

Feature	Mean and deviation in posts with exper.	without exper.
subjectivity score[2]	0.07 ±0.23	0.17 ±0.35
polarity score[2]	0.87 ±0.30	0.77 ±0.38
#words per post	102.57 ±80.09	52.46 ±53.24
#sentences per post	6.00 ±4.16	3.34 ±2.33
# words per sentence	17.07 ±4.69	15.71 ±7.61
#questions per post	0.32 ±0.63	0.54 ±0.89
p(post contains question)	0.25 ±0.43	0.33 ±0.47
#*I*'s per post	5.76 ±4.75	2.09 ±2.88
#*I*'s per sentence	1.01 ±0.48	0.54 ±0.60
p(sentence in post contains *I*)	0.67 ±0.23	0.40 ±0.35
#non-modal verbs per post	19.62 ±15.08	9.82 ±9.57
#non-modal verbs per sent.	3.30 ±1.18	2.82 ±1.37
#modal verbs per sent.	0.22 ±0.22	0.26 ±0.36
fraction of past-tense verbs	0.26 ±0.17	0.17 ±0.19
fraction of present tense verbs	0.42 ±0.18	0.41 ±0.23

Table 1: Comparison of surface text features for posts with and without experience; $p(\cdot)$ denotes probability.

We acquired data by crawling two forums on shaving,[1] with 111,268 posts written by 2,880 users.

Manual assessments Two assessors (both authors of this paper) were asked to search for posts on five specific target products using a standard keyword search, and label each result post as:

- reporting no experience, or
- reporting an off-target experience, or
- reporting an on-target experience.

Moreover, posts should be marked as reporting an experience only if (i) the author explicitly reports his or someone else's (a concrete person's) use of a product; and (ii) the author makes some conclusions/judgements about the experience.

In total, 203 posts were labeled by the two assessors, with 101 posts marked as reporting an experience by at least one assessor (71% of those an on-target experience). The inter-annotator agreement was 0.84, with Cohen's $\kappa = 0.71$. If we merge on- and off-target experience labels, the agreement is 0.88, with $\kappa = 0.76$. The high level of agreement demonstrates the validity of the task definition.

Features for experience mining We considered a number of linguistic features and compared posts reporting experience (on- or off-target) to the posts with no experience. Table 1 lists the features and the comparison results. Remarkably, the subjectivity score is lower for experience posts: this indicates that our task is indeed different from sentiment retrieval. Experience posts are on average twice as long as non-experience posts and contain more sentences with pronoun *I*. They also contain more content (non-modal) verbs, especially past tense verbs. Table 2 presents a more detailed analysis of the verb use. Experience posts appear to contain more verbs referring to concrete actions rather than to attitude and perception. It is still to be seen, though, whether this informal observation can be quantified using resources such as standard semantic verb classification (*state, process, action*), WordNet verb hierarchy or FrameNet semantic frames.

[1] www.shavemyface.com, www.menessentials.com/community

[2] Computed using LingPipe: http://alias-i.com/lingpipe

With experience	Without experience
used 0.15, found 0.09, bought 0.07, tried 0.07, got 0.07, went 0.07, started 0.05, switched 0.04, liked 0.03, decided 0.03	got 0.09, thought 0.09, switched 0.06, meant 0.06, used 0.06, went 0.06, ignored 0.03, quoted 0.03, discovered 0.03, heard 0.03

Table 2: Most frequent past tense verbs following *I* in posts with and without experience, with rel. frequencies.

5 Conclusions

We introduced the novel task of experience mining. Users of products share their experiences, and mining these could help define requirements for next-generation products. We developed annotation guidelines for labeling experiences, and used them to annotate data from online forums. An initial exploration revealed multiple features that might prove useful for automatic labeling via classification.

References

[1] K. Balog, G. Mishne, and M. de Rijke. Why are they excited?: identifying and explaining spikes in blog mood levels. In *EACL '06*, pages 207–210, 2006.

[2] B. Buxton. *Sketching User Experiences: Getting the Design Right and the Right Design*. Morgan Kaufmann Publishers Inc., 2007.

[3] I. Ounis, C. Macdonald, M. de Rijke, G. Mishne, and I. Soboroff. Overview of the TREC 2006 Blog Track. In *TREC 2006*, 2007.

[4] B. Pang and L. Lee. Opinion mining and sentiment analysis. *Found. Trends Inf. Retr.*, 2(1-2):1–135, 2008.