Particular Requirements on Opinion Mining for the Insurance Business

Sven Rill, Johannes Drescher, Dirk Reinel, Jörg Scheidt, Florian Wogenstein
Institute of Information Systems (iisys)
University of Applied Sciences Hof
95028 Hof, Germany
{srill,jdrescher,dreinel,jscheidt,fwogenstein}@iisys.de

***Abstract*—In this paper, we discuss the work in progress of our current project focusing on opinion mining in the field of insurance business. The main purpose of this project is to improve Opinion Mining methods for the German language and optimize them with special regard to the insurance business. These improved methods make it possible to extract opinions from user-generated texts (in the insurance domain) in a better quality than today. We fetch the required text data for this study from a huge online community website for customers. There, we find a sufficient number of user reviews about insurance companies, which is necessary for our research. Besides the main purpose of this study, another aim is the development of a prototype. This could then be used to monitor the current "crowd's opinion" about insurance products and services. For this reason and in order to understand key aspects of the domain, we collaborate with the *nobisCum Deutschland GmbH*, a German company offering consulting and software development services for the insurance industry. Using one data source and limited evaluation sets, we obtained first results which look promising.**

Keywords-opinion mining; insurance; aspect-based opinion mining; domain-specific orientation.

I. INTRODUCTION

The steep rise of user-generated content over the last few years, especially user reviews, on the Web, together with its reliability and the wide public acceptance, makes it therefore important for many companies to observe these relevant sources. Since the amount of reviews, comments, posts, etc. (*user contributions* in short) from only one web platform is, sometimes, too large to be checked manually, methods are required to automate that. There are already some applications done by different groups or companies that claim to be able to reliably extract opinions from German texts. However, that does not always work very well yet. In many cases, the results are restricted to basic opinions (positive, neutral or negative) of certain texts, calculated by counting all opinion bearing words in this text (a kind of document-level sentiment classification, e.g., a bag of words approach).

Using textual data together with its meta information from different review portals, forums, blogs and/or social networks (e.g., Facebook, Twitter) together with advanced Opinion Mining methods such as *Aspect-based Opinion Mining* can help to solve this problem. Thereby, it is also very important to handle domain- and context-specific orientations on generated opinion words and phrases and, in the course of this, also to improve Opinion Mining methods for the German language. In our case, we have to improve them particularly for the insurance domain.

Beside this main purpose, another goal is to include the improved or developed methods and algorithms into a prototype. By using this prototype application, insurance companies will be able to monitor the opinion of their customers or of their whole peer group on certain aspects of their own products "in real time".

In Section II, some related work is introduced. Section III contains a brief description of the *OMVers* (Opinion Mining für die Versicherungsbranche - Opinion Mining for the insurance business) project, our identified issues and a short description about the input data. Section IV mainly depicts the different domain-specific opinion mining steps. In Section V, we want to explain our evaluation process. A short conclusion completes the paper in Section VI.

II. RELATED WORK

Especially in the last few years, a lot of research work has been done in the area of Opinion Mining and Sentiment Analysis. A detailed overview of the whole topic has been given from Pang and Lee [1] and just recently in a survey from Liu and Zhang [2].

A good overview of the *Aspect-based Opinion Mining* approach is also given in the work from Liu and Zhang [2]. In addition to this, Liu defined a model to describe aspects in a document, called *quintuple* [3]. A method to extract the required aspects is presented in research [4].

Furthermore, there are several lists of opinion words for multiple languages. For the English language such lists are SentiWordNet [5], the Subjectivity Lexicon [6], Semantic Orientations of Words [7] and two lists of positive and negative opinion words offered by [8].

O'Hare et al. [9] analyzed blogs in the financial domain to automatically determine the sentiment of the bloggers. Zhuang et al. [10] focused their research on the movie domain and proposed an approach for review mining and summarization.

 ISBN: 978-1-61208-227-1

There is a lot of preliminary work on several aspects in the field of Opinion Mining. As a summary of this work, one can say that a general approach to Opinion Mining, applicable in many domains, does not yet give satisfactory results. On the other hand, domain specific applications are already promising.

The focus of this project is to apply Opinion Mining techniques exclusively in the insurance business.

III. The OMVers Project

OMVers [11] is a collaborative project of the *nobisCum Deutschland GmbH (nobisCum)* and the *Institute of Information Systems (iisys)*. This synergetic cooperation is very important for the project's success. While nobisCum mainly works on developer tasks such as building a front- and a back-end, creating a suitable database, providing an interface for the opinion mining analyze module, etc., iisys can completely dedicate itself to research. During the entire time, the project can benefit from the knowledge as well as the experience of nobisCum about the insurance sector.

A. Major Issues

Before we will be able to analyze user written texts in the insurance domain in such a way that suitable results can be achieved, several subtasks have to be defined and finished. In preparation for this project we identify the following major issues as such tasks (partly based on survey [2]). The final aim is to create opinion quintuples (see [3]) for every analyzed user written text in this domain.

1) Generate Opinion List: Existing lists containing opinion bearing adjectives, nouns or verbs (and phrases) for the German language are not sufficient for our project. Therefore, we need to produce an own list. Our research group considered this issue separately and published a generic approach to generate such lists [12].

2) Improve Opinion Mining for German: Currently, there are some weaknesses in Opinion Mining methods for the German language, e.g., identifying compound nouns. An example for such a noun is *"Versicherungsbetrug"* - *"insurance fraud"*.

3) Identify, extract and group aspects: The automatic extraction of aspects (also known as *features*) from a text is one of the most important parts of this project. Hu and Liu [4] present an interesting two-step-method to perform that. The accuracy of the first step was already improved by Popescu and Etzioni [13]. After extracting, the aspects have to be combined to groups. For that, the OpenThesaurus (see http://www.openthesaurus.de/) will be useful for looking up synonyms.

4) Handle domain-specific opinion words: Since we have our self-generated general opinion list, another important issue is how to handle domain-specific opinion words. For example, in the sentence *"I will change my insurance company"* the verb *change* in the insurance domain expresses a really strong opinion (negative in this case). Whereas in the sentence *"I will change my clothes"* the same verb expresses no opinion (objective sentence). Furthermore, we assume that about 80 % of the entries of our opinion list are universal, in other words domain-independent. However, this assumption has not been verified yet.

5) Handle context-dependent opinion words: Similar to the domain-specific issues mentioned above, the problem of context-dependent opinion words is really relevant for our project, too. Let us have a look at the following two sentences: *"I will change to this insurance company"* and *"I will change to my previous insurance company"*. The verb *"change"* expresses opposite opinions, positive in the first sentence, negative in the second one. There is another special case we can see in the sentence *"In any case, I will change"*. In such a case it would be impossible to determine the opinion expressed by *"change"* without looking at previously written sentences.

Ding et al. [14] proposed an approach to handle opinion words that are context-dependent.

6) Map opinion words to aspects: As soon as we have identified the aspects we can use our list of opinion bearing words to bring them together. This is a huge and important step towards our aim to create opinion quintuples. However, there are already methods for this officially called *aspect sentiment classification*, which we can test and adapt to our specific requirements [4] [14].

B. Additional Issues

In addition to the major issues there are some additional issues, such as extracting and grouping the entity (the insurance company), extracting the opinion holder and time, mining comparative opinions, handling coreference resolutions and extending the approach for multilingualism.

Although the extraction of the entity, opinion holder and time are not minor issues, we want to simplify this aspect for now. Therefore, we define that

- There is only one well-known entity per text,
- The opinion holder is always the author of the text and
- The time is always the publication time of the text.

For that reason, our opinion quintuple at the beginning will look as follows: (e, a_j, oo_j, h, t). Entity *e*, opinion holder *h* and time *t* are unique per text, while aspect *a* and the opinion value *oo* are not. In a later phase of the project, this current restriction has to be improved. Please note furthermore that, instead of a basic opinion orientation [3], we work with continuous opinion values between -1 and +1 [12].

After having finished all subtasks described in Section III-A, we expect very satisfactory results. This is the reason why the additional issues are not mandatory at the moment.

C. *Input Data*

For first investigations, we use user-generated reviews about automobile insurances from an online community for customers called *Ciao* [15]. So far, no method has been developed that would enable us to crawl such reviews automatically, so we have to fetch them manually. Besides the text, the main information of a typical *Ciao* review is a title, the author's alias plus additional user information, a rating value, the publication date, pros and cons defined by the author, an "advisable flag" as well as an average review rating retrieved from other *Ciao* members. We also use some of this additional information for this project.

IV. Domain-Specific Opinion Mining

The aim of our data analysis is to go down to the aspect level of a specific text and create the appropriate opinion quintuples. Thus, it will be possible to see a quantitative as well as qualitative summary of the opinions of different texts.

However, to see how good or bad other methods for this application case work, we have decided to use an iterative approach. This implies that we start our research with a basic *document-level sentiment classification* [1], meaning that the smallest unit is the whole review text. This method is only suitable for documents that contain just one entity (as they do in our case). Nevertheless, the granularity of this technique is probably not fine enough for our application.

After that, we test a method called *sentence-level sentiment classification*, which is an intermediate step before finally reaching the "supreme discipline" of *aspect-based sentiment analysis* or *aspect-based opinion mining*, respectively.

After each step, we perform an evaluation of the results produced by the respective algorithm with a self-defined set of sentences (see Section V). This allows us to measure the quality of every method used as well as to compare them. Furthermore it will help us to improve the methods in an iterative way.

Since the implementation of the first two main steps (*document-level sentiment classification* and *sentence-level sentiment classification*) has almost been completed, we want to describe the ongoing and future work on the *aspect-based sentiment analysis* below, which is partly based on the five tasks (except IV-A) to be performed to build opinion quintuples [2].

A. *Domain-specific Opinion List Generation*

As already mentioned in Section III-A1, we have published a generic approach to generate opinion lists of phrases [12]. By using this approach, we have already created such a list.

As described in Section III-A4, we now have to deal with domain-dependent opinion words. These words are currently not contained in the list or have the wrong opinion value for the insurance domain (remember the example of the verb *change*). Therefore, in the next step we add such words to our existing list manually or change their opinion value if they already exist.

B. *Entity Extraction and Grouping*

Currently, this task is treated as an additional issue (see Section III-B). Therefore, we start by using the *Ciao* hierarchy (*Ciao* > insurances > automobile insurances > [list of all insurances]) to determine the (unique) entity of review texts. Later on, we plan to extract entities automatically. This is still required in order to analyze blogs and forums (unstructured sources) and could be handled by using Named Entity Recognition (NER) techniques [16].

C. *Aspect Extraction and Grouping*

As mentioned in Section III-A3, aspect extraction is a crucial task of our project. Currently, we highly simplify this issue. That means we are currently using a manually produced list of grouped aspects as well as their synonyms prepared by *nobisCum*. Although this approach works quite well, our aim is to extract and group aspects automatically as soon as possible. In addition to that, a method to split compound nouns (a common occurrence in the German language) is needed as well (see Section III-A2).

D. *Opinion Holder and Time Extraction*

Similar to the entity extraction part, this task is also treated as an additional issue. As described in Section III-C, every user review from *Ciao* contains the author's alias and the publication date, among others. This meta information is presently used to determine the opinion holder as well as the time of the whole review. After finishing the major issues (see Section III-A), this task will be automated, too.

E. *Aspect Sentiment Classification*

As we now have our list with opinion bearing words (not yet domain-specific) and the required, at the moment static, list of aspects, we can start a simple analysis of reviews to determine the opinion on different aspects. Since our opinion list also contains opinion values for phrases like *"nicht gut"* - *"not good"*, we do not have to handle valence shifters.

We are still at the beginning of this task, i.e., by now we have neither managed context-dependent opinion words (see Section III-A5) nor "but phrases", comparative opinions, etc.

 ISBN: 978-1-61208-227-1

Currently, we use a basic approach to aggregate opinions on various aspects (see Section III-A6).

Later on, we plan to significantly improve this approach. After generating and using a domain-specific opinion list, the next intermediate aim of this task is to handle context-dependent opinion words, which are indispensable to get satisfactory results.

F. Opinion Quintuple Generation

After finishing all previous tasks, we will be able to produce simplified opinion quintuples *(e, a_j, oo_j, h, t)* (see Section III-B). As already mentioned, the quality as well as the complexity of these quintuples are to be improved iteratively by improving the other tasks.

V. Evaluation

To evaluate the quality of the applied and adjusted opinion mining methods, we create a set of sentences related to the insurance domain.

We have considered using a three-class model. The first class, which is the basic one, contains simple sentences such as *"The claim settlement of that insurance company is very good"*. Sentences of this category should be easy to analyze. The second class includes more difficult sentences (that means one subordinate clause, "but phrases", etc.), e.g., *"The employee of the customer service was friendly but not really helpful"*. The third and most intricate class contains sentences with many subordinate clauses as well as irony, e.g., *"They raised the yearly subscription again, I really love this company"*.

Every whole sentence as well as the aspects inside, if any, must be tagged in a machine-readable schema. In order to determine the opinion value of a sentence or aspect we decided to use the following six categories: strong positive (sp), weak positive (wp), neutral (n), weak negative (wn), strong negative (sn) and objective (o).

This set of tagged sentences enables us to measure the quality of our methods. Thus, we can see improvements as well as possible deterioration.

VI. Conclusion

Although user-generated content provides an enormous potential in the area of Opinion Mining, which makes it attractive for companies to pursue real time customer monitoring without relying on the usual polling techniques, it has been used little so far. The main reason for this is that Opinion Mining methods, which are available to companies (especially German ones) still do not work satisfactorily.

We aim to improve Opinion Mining methods for the German language in the course of this project. Early experiments have already shown that a domain- and, of course, a context-dependent approach is indispensable for this. As a next step, we have to check whether the assumption that about 80% of our self-generated opinion list is domain-independent is correct.

At first glance, the extraction of opinions with our adapted methods and our opinion list works quite well, but until now we have only worked with user reviews from a single source as well as within limited evaluation sets. In the near future, we would include blogs and forums and we should then see how different writing styles, text structure and a mix of topics affect our methods.

The first audited results of the project are expected in the fourth quarter of 2012.

Acknowledgment

The *OMVers* project is publicly sponsored by the Bavarian Ministry of Economic Affairs, Infrastructure, Transport and Technology (BStMWIVT).

The Institute of Information Systems is supported by the Foundation of Upper Franconia and by the State of Bavaria.

References

[1] B. Pang and L. Lee, "Opinion mining and sentiment analysis," vol. 2, p. 1, 2008.

[2] B. Liu and L. Zhang, *Mining Text Data*, 2012, ch. A Survey of Opinion Mining and Sentiment Analysis, pp. 415–463.

[3] B. Liu, *Handbook of Natural Language Processing*, 2010, vol. 2, ch. Sentiment Analysis and Subjectivity.

[4] M. Hu and B. Liu, "Mining and summarizing customer reviews," in *Proceedings of the tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD-2004)*, 2004, pp. 168–177.

[5] S. Baccianella, A. Esuli, and F. Sebastiani, "Sentiwordnet 3.0: An enhanced lexical resource for sentiment analysis and opinion mining," in *Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC'10)*, vol. 25, May 2010, pp. 2200–2204.

[6] T. Wilson, J. Wiebe, and P. Hoffmann, *Human Language Technology Conference - Conference on Empirical Methods in Natural Language Processing*, 2005, ch. Recognizing Contextual Polarity in Phrase-Level Sentiment Analysis.

[7] H. Takamura, T. Inui, and M. Okumura, "Extracting semantic orientations of words using spin model," in *Proceedings of the 43rd Annual Meeting of the ACL*, 2005, pp. 133–140.

[8] B. Liu, M. Hu, and J. Cheng, "Opinion observer: Analyzing and comparing opinions on the web," in *Proceedings of the 14th International World Wide Web conference (WWW-2005)*, 2005, pp. 342–351.

[9] N. O'Hare, M. Davy, A. Bermingham, P. Ferguson, P. Sheridan, C. Gurrin, and A. F. Smeaton, "Topic-dependent sentiment analysis of financial blogs," in *Proceedings of the 1st International CIKM Workshop on Topic-Sentiment Analysis for Mass Opinion Measurement (TSA-09)*, 2009, p. 9.

[10] L. Zhuang, F. Jing, and X.-Y. Zhu, “Movie review mining and summarization,” in *Proceedings of the 15th ACM International Conference on Information and Knowledge Management (CIKM-2006)*, 2006, pp. 43–50.

[11] Institute of Information Systems (iisys), “Analytical Information Systems: Projects,” http://www.iisys.de/en/research/research-groups/analytical-information-systems/projects.html, (accessed September 21, 2012).

[12] S. Rill, J. Drescher, D. Reinel, J. Scheidt, O. Schütz, F. Wogenstein, and D. Simon, “A Generic Approach to Generate Opinion Lists of Phrases for Opinion Mining Applications,” in *Proceedings of the First International Workshop on Issues of Sentiment Discovery and Opinion Mining (WISDOM-2012)*, 2012.

[13] A.-M. Popescu and O. Etzioni, “Extracting product features and opinions from reviews,” in *HLT ’05: Proceedings of the conference on Human Language Technology and Empirical Methods in Natural Language Processing*, Oct 2005, pp. 339–346.

[14] X. Ding, B. Liu, and P. S. Yu, “A holistic lexicon-based approach to opinion mining,” *WSDM08*, pp. 231–240, 2008.

[15] Ciao GmbH, “Preisvergleich und Testberichte bei Ciao,” http://www.ciao.de/, (accessed June 4, 2012).

[16] D. Nadeau and S. Sekine, “A survey of named entity recognition and classification,” vol. 30, pp. 3–26, 2007.