From Genre Classification to Aspect Extraction: New Annotation Schemas for Book Reviews

Tamara Álvarez-López, Patrice Bellot, Milagros Fernandez-Gavilanes, Enrique Costa-Montenegro

From Genre Classification to Aspect Extraction: New Annotation Schemas for Book Reviews

Tamara Álvarez-López
Aix Marseille Univ, Université de Toulon, CNRS, LIS
University of Vigo
talvarez@gti.uvigo.es

Patrice Bellot
Aix Marseille Univ, Université de Toulon, CNRS, LIS
Marseille, France
patrice.bellot@lis-lab.fr

Milagros Fernández-Gavilanes
University of Vigo
Vigo, Spain
milagros.fernandez@gti.uvigo.es

Enrique Costa-Montenegro
University of Vigo
Vigo, Spain
kike@gti.uvigo.es

***Abstract*—In this paper, new schemas for feature categorization in different types of reviews, in the domain of books, are presented, so aspect extraction techniques could be later applied. We deal here with two different types of reviews: formal reviews about scholarly books, written by experts, and informal ones about fiction books, written by readers which are not necessarily highly qualified. Our final goal is to extract the most relevant aspects or features to which any opinion is expressed in these reviews, along with the sentiment associated, for later integrating it to book recommender systems, improving the quality of the recommendations. Throughout this paper, the need for different annotation schemas is proved, by developing a new review classification system, as well as making an analysis at lexical and semantic levels on both kinds of reviews, for finally concluding with the presentation of the new categorization schemas.**

***Index Terms*—book reviews, aspect extraction, category detection, recommender systems, digital humanities, sentiment analysis**

I. Introduction

A great number of scholarly documents in digital libraries can be found, as well as comments and reviews from blogs, online book shops or social networks. Manually inspecting all this information to find a book which fits in our requirements is still difficult. In this sense, book reviews and recommender systems play an important role, trying to provide personalized information to the user. In the last decade, recommender systems have started to use the content of the reviews to make better recommendations [5]. Related to this task, the goal of The Social Book Search (SBS) Lab [1], a track belonging to the CLEF (Conference and Labs of the Evaluation Forum) from 2007 to 2016, has been to research and develop techniques to support readers in complex book search tasks, providing a common evaluation dataset. However, this dataset does not contain annotations about the aspects or sentiments expressed in the reviews, which could deal with the price or the quality. Our work focuses on that point.

Aspect based sentiment analysis (ABSA) deals with extracting the specific entities mentioned in a review and the sentiment associated to them [2], so that we can know at a glance which aspects of a book are addressed in a review and the sentiment expressed towards each one. This task can help to provide more accurate results, not only for recommender systems but also for extracting, querying and organizing data in a structured way. Most of the ABSA research focuses in electronic products and restaurants domains [3], [4], however, there is not much research specifically for book reviews. Some recent works can be found [6]–[8], related to new annotated datasets for ABSA. They all work with short reviews or even tweets, whereas in this work we deal with very long reviews from expert readers as well. More formal reviews about scholarly books have not still been studied, and so they are our main target, as there are no datasets containing this kind of reviews annotated with aspects or sentiment information in books domain. Book reviews in general present a bigger challenge, due to the complexity of the texts, the length and the kind of features to extract.

Finally, we have to make reference to some works which can be found on a related task, which is identifying the genre of a text based on its style, both in textual documents [9], [10], as well as, more recently, in web documents [11]. A genre or a style is another view of a document different from a subject or a topic, and it is also a criterion to classify documents. They employ classification methods, taking into account several characteristics of the texts, such as structural, lexical, punctuations, etc.

Analyzing how users tend to express their opinions in this domain and designing new schemas, are the first steps for later developing new ABSA systems and new annotated datasets for training and testing them. Throughout this paper we perform an analysis of different kinds of book reviews extracted from two very different sources: Amazon and OpenEdition platforms, as well as a system which is able to classify any review into one of these two types, to prove the need for different ABSA models and annotation schemas. Finally, the new aspect categorization schemas are presented, which will allow new research in aspect extraction, opinion mining, recommendation and classification of digital documents.

The remaining of this article is structured as follows. Section II describes the collections used for this work. In section III a method for review classification is presented. Then, in section IV, a deeper analysis of the reviews at lexical and semantic levels is carried out. Section V presents the new annotation schemas, and finally some conclusions and future work are presented in section VI.

II. Data Description

Two different types or reviews are analyzed in this paper. On the one side, we analyze formal reviews, written by experts about scholarly books, which can be found in specialized platforms or in journals. For this study, we will extract them from the scientific digital library OpenEdition[1]. On the other side, we analyze the reviews written by any reader, usually in the majority non professional one, about fiction books, obtained from Amazon[2], where every user can express an opinion about a particular book.

A. Amazon Reviews

We randomly selected 300 reviews, associated to 40 different fiction books from the Amazon/LibraryThing corpus for English language provided by the SBS Lab [1]. We only extracted the textual content of the reviews, obtaining a total number of 2977 sentences from the 300 reviews, which are enough for our study and have already been annotated with aspects, categories and sentiment information. They are publicly available online [6].

B. OpenEdition Reviews

Our second dataset is composed of book reviews written by experts, extracted from OpenEdition, dedicated to electronic resources in the Humanities and Social Sciences. Here, we focus on the book reviews, manually extracting 50 of them in English language, containing a total of 2957 sentences. All the reviews selected are publicly available in the OpenEdition platform online.

III. SVM for Review Classification

In this section we present a method for classifying reviews according to the kind of the book and the type of evaluation: informal review about a fiction book or a formal one about a scholarly book.

For the aim of later applying aspect extraction to recommendation systems it is interesting to know in advance which reviews are more appropriate according to the kind of reader to which we want to make the recommendation. With this method we will be able to make a better selection of the reviews that should be analyzed and which we want to extract the aspects from, in order to finally obtain more personalized recommendations. By first filtering the input reviews according to one of the two types presented in this paper, we have a first idea of which kind of information could be addressed and also the kind of vocabulary we are going to find, so we can apply different aspect extraction and category detection models, and so decide which of the schemas designed, presented in the next sections, should be applied for extracting relevant information. Finally, this information extracted from the reviews selected will be added as input for the recommender system.

A binary SVM classifier with a linear kernel is constructed for the review classification, applying 5-fold cross validation. This method was chosen as it proved its performance for this kind of tasks in numerous works in the field. The inputs of the classifier are the individual sentences of the reviews, taking into account only the words appearing in them as features (bag of words), so the classifier will determine if a particular sentence belongs to a review extracted from Amazon about a fiction book (informal review) or from OpenEdition about a scholarly book (formal review). Although other binary features were also tested (lemmas, POS tags, entity recognition and bigrams), they did not show better performance. Moreover, we make sure that the sentences in training belong to different reviews, and also to different books, than those ones selected for testing. Thus we assure that the results are not biased.

The evaluation of the method is performed in terms of precision (P), recall (R) and F-score (F1) [12]. In Table I the results obtained for the classification at sentence and review level are shown for the two types of reviews, formal and informal. It can be seen that an average F-score of 84% is obtained in the classification of the individual sentences, and an average of 99% for the whole review. We determine that a review belongs to a particular type if so they do the majority of the sentences it contains.

TABLE I
Precision, recall and F-score at sentence and review levels for 2-class classification

Reviews	Sent. level			Rev. level		
	P	**R**	**F1**	**P**	**R**	**F1**
Amazon fiction	0.93	0.88	0.90	1.0	0.99	0.99
OpenEdition	0.73	0.83	0.77	0.97	1.0	0.98
Average	**0.83**	**0.86**	**0.84**	**0.99**	**1.0**	**0.99**

Finally, in order to go deeper in the review classification, another experiment is performed. In this case, a third set of reviews is obtained, also extracted from Amazon, so it means that can be written by any kind of reader, normally non-expert, but in this case about non-fiction books. As we could not find reviews for the same titles extracted from OpenEdition, we tried to find similar books to these ones, which contained user reviews. We manually extract 270 reviews, which are classified in the platform in the topics of History, Religion and Social Sciences, which are similar to the kind of topics we can find in OpenEdition platform, obtaining a total of 3048 sentences, so the number of sentences is similar in the three datasets. We perform again the classification of the reviews for the three different ones, the two datasets previously presented and this new one, using a 3-class SVM classifier. With this new experiment we want to make sure that we can classify a new review according to the kind of book (fiction or non-fiction) and to the kind of writer (expert or non-expert), and

[1] www.openedition.org
[2] www.amazon.com

that the previous experiment is not only classifying the reviews according to the platform where it was published. In Table II, the results obtained in terms of precision (P), recall (R) and F-score (F1) at sentence and review level for this new case are shown. We can observe that we are able to distinguish the three types of reviews with high precision, obtaining an F-score of 82%.

TABLE II
PRECISION, RECALL AND F-SCORE AT SENTENCE AND REVIEW LEVELS FOR 3-CLASS CLASSIFICATION

Reviews	Sent. level			Rev. level		
	P	R	F1	P	R	F1
Amazon fiction	0.81	0.61	0.69	0.91	0.77	0.83
Amazon non-fiction	0.45	0.62	0.52	0.72	0.82	0.76
OpenEdition	0.63	0.69	0.65	0.76	0.98	0.85
Average	**0.62**	**0.64**	**0.63**	**0.79**	**0.86**	**0.82**

The results obtained for the review classification confirm that there are important differences between the types of reviews, as the classifier can distinguish them with high precision, so we can think of applying different schemas when we want to extract the most important aspects addressed in them. In the next sections, we will only focus in the Amazon reviews for fiction books and the reviews from OpenEdition, about scholarly books. They will be analyzed at lexical and semantic levels in order to later propose the different annotation schemas that are more appropriate for each one.

With this system we prove that there are indeed big differences between both types of reviews so it has sense to develop different ABSA models for each one, and it also allows us to classify any new review recovered from different sources online, so we can decide which annotation schema should be applied to extract the most relevant information in each case.

IV. LEXICAL AND SEMANTIC ANALYSIS

In this section both types of reviews are analyzed according to lexical and semantic characteristics, showing the differences found between both. General properties and more specific ones related to the kind of vocabulary used in each kind of reviews are described in the next subsections.

A. General Properties

In the first place, there are some general differences between Amazon and OpenEdition reviews. One of them is the type of readers they are written for and their main objective. The former ones are usually a short description of a book, with references to the plot or the characters, if it is an enjoyable reading or if it has a slow or fast pace in the situations narrated. Otherwise, the latter are mostly written with a more educational goal, trying to assess the quality of the book for a specific field of knowledge and oriented to expert and professional readers or researchers.

Then, taking into account the length of the reviews and the sentences they contain, we find that in Amazon, the average length in the corpus extracted is 9.92 sentences per review, and 19.7 tokens per sentence, whilst for the OpenEdition reviews, it is 59.14 sentences per review, and 33.6 tokens per sentence. We can already think about the higher specificity, and the deeper analysis done about the book, in the expert-generated reviews, according to their extension.

B. Lexical/semantic Properties

When inspecting the kind of vocabulary used in both types, we can also appreciate different ways of expressing opinions about books. In Figure 1 the vocabulary size and the number of different nouns, verbs, adjectives and polar words are shown for an average review of both types. We can see that in OpenEdition reviews, a wider variety of vocabulary is used in general, represented by the vocabulary size parameter, as expected also due to the average length of the reviews. Moreover, when we focus on each type of words, the result is the same, obtaining a bigger number of different nouns, adjectives, verbs and polar words for the scholarly texts written by experts. We can think that in the OpenEdition reviews, the number of aspects evaluated from the book is much bigger, making a more detailed analysis of the book.

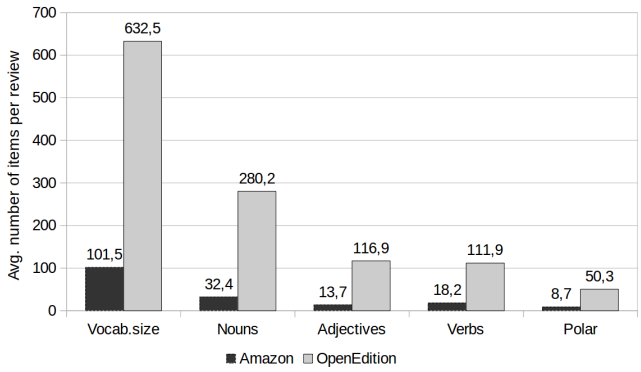

Fig. 1. Average number of different tokens per review in Amazon and OpenEdition datasets.

With respect to the nouns used in both kinds of reviews, we can see in Table III those ones which are the most common in each corpus, along with the number of sentences which contain each noun in each dataset. In this way we can make a fair comparison between the two types, as they both contain a similar number of sentences. The most common nouns used in Amazon reviews are: *book* (appearing in 703 sentences), *story* or *character*; while in OpenEdition are: *book* (appearing only in 209 sentences), *chapter* or *author*. Moreover, there are some other nouns, like *mystery*, *adventure*, *plot* or *thriller* which are common in Amazon reviews, but very rare in OpenEdition ones. On the contrary, we find that *analysis*, *study*, *theory* or *essay* are much more common in OpenEdition texts.

When focusing on the adjectives, we see in Table IV the most frequent adjectives in each dataset. We find that adjectives like *good*, *great*, *bad* and others like *mysterious*, *funny* or *amazing* are very common in Amazon reviews, whilst in OpenEdition we frequently find others like *social*, *political*, *literary*, *historical*, *critical* or *economic*, which are very specific to this kind of reviews.

TABLE III
MOST FREQUENT NOUNS IN AMAZON AND OPENEDITION REVIEWS

Word	Amazon	OE	Word	OE	Amazon
book	703	209	book	209	703
story	176	36	chapter	195	26
character	136	22	author	158	101
series	126	16	history	131	19
author	101	158	study	120	3
novel	67	21	work	113	35
reader	60	101	reader	101	60
plot	53	5	literature	91	8
fantasy	49	9	society	84	11

TABLE IV
MOST FREQUENT ADJECTIVES IN AMAZON AND OPENEDITION REVIEWS

Word	Amazon	OE	Word	OE	Amazon
good	139	32	social	116	6
great	90	24	political	115	4
new	77	113	new	113	77
young	51	22	literary	85	7
strange	42	1	historical	83	9
interesting	42	34	russian	81	0
different	39	80	cultural	81	2
bad	34	6	different	80	39
real	33	29	military	66	0

Similarly, in Table V it is shown that the verbs used in Amazon reviews are very general, like *get*, *find*, *give* or *come*. However, in formal reviews published in OpenEdition the verbs used belong to a more formal language, such as *include*, *argue* or *provide*.

TABLE V
MOST FREQUENT VERBS IN AMAZON AND OPENEDITION REVIEWS

Word	Amazon	OE	Word	OE	Amazon
read	266	64	write	122	86
get	155	20	take	108	76
find	135	63	include	83	18
go	132	59	become	74	49
think	108	21	argue	74	2
give	88	99	provide	70	10
come	88	70	show	65	21
write	86	122	develop	55	11
know	82	36	consider	52	2

Then, sentiment-bearing words are analyzed for both datasets, so it can be seen how the positive or negative opinions are expressed in both. For this task the SOCAL sentiment lexicon [13] is applied, which is composed of a list of words with a polarity associated, expressed on a scale between -5 (most negative) and +5 (most positive). Then, we extract all the words with a polarity higher than 1.5 or lower than -1.5. The most frequent polar words in both datasets are shown in Table VI, along with the number of sentences where they appear.

We can observe that, while in Amazon reviews the most frequent words for expressing sentiment are *good*, *great* or *love*, in OpenEdition we find more often others like *war, critical* or *moral*, which belong again to a more specific and formal vocabulary than the first ones.

TABLE VI
MOST FREQUENT POLAR WORDS IN AMAZON AND OPENEDITION EXTRACTED FROM SOCAL

Word	Amazon	OE	Word	OE	Amazon
good	139	32	problem	58	31
great	90	24	modern	55	13
like	77	9	war	54	8
love	70	7	critical	49	1
recommend	58	7	important	47	11
enjoy	48	6	moral	44	3
bad	35	6	criticism	29	3

C. Z-Score Experiments

In this section, the Z-Score measure is used in order to identify the most salient words belonging to the specific classes (Amazon and OpenEdition). Other authors have used the Z-Score as a measurement of the importance of different terms in a dataset [14]. A high Z-Score for a word in a particular dataset, compared to the other, means that it clearly belongs to that specific context.

We compute the Z-Score for each term t_i in a dataset C_j (t_ij) as:

$$Z_{score(t_{ij})} = \frac{tfr_{ij} + n_j \cdot P(t_i)}{\sqrt{n_j \cdot P(t_i) \cdot (1 - P(t_i))}} \quad (1)$$

where tfr_{ij} is the relative frequency of the term t_i in a particular dataset C_j; n_j is the total number of terms in the dataset C_j; and $P(t_i)$, the term probability over the whole corpus (both datasets together).

In Table VII we can see the result of computing the Z-Score for every term in each dataset. The terms which achieve the highest Z-Score in each dataset are displayed, along with the correspoding Z-Score achieved by the same term in the other dataset. It can be observed that those words with a high Z-Score in Amazon reviews have a low Z-Score in OpenEdition ones and vice versa. Therefore, we can state that the kind of vocabulary used in both is very different.

TABLE VII
LIST OF TERMS WITH THE HIGHEST Z-SCORE IN AMAZON AND OPENEDITION REVIEWS, ALONG WITH THEIR Z-SCORE VALUE

Word	Amazon	OE	Word	OE	Amazon
book	21.84	-16.64	study	5.25	-6.89
read	15.06	-11.48	political	4.83	-6.33
story	11.42	-8.71	chapter	4.73	-6.21
series	10.73	-8.18	social	4.49	-5.89
character	10.35	-7.89	military	4.43	-5.81
good	9.85	-7.51	cultural	4.37	-5.74
like	8.40	-6.40	law	4.12	-5.42
love	7.93	-6.04	volume	4.07	-5.34
novel	7.81	-5.96	literature	4.04	-5.31
great	7.54	-5.75	argue	3.84	-4.91

V. NEW ANNOTATION SCHEMAS

With the previous sections it is clear that there are differences in the vocabulary and the way of expressing opinions between the reviews written by experts about scholarly books and the ones written by any user about fiction books or novels.

For this reason, when trying to extract the most relevant aspects from these texts, not the same models can be applied. In this section we present a different set of categories for each type of reviews, representing the aspects most frequently addressed, by looking at the most common words appearing in each dataset, shown before. As far as we know, there is not much research on book reviews in general, and even less on expert book reviews. Therefore, these schemas are a first step also for developing new annotated datasets for ABSA, so they can be used for training supervised systems, as well as for testing new proposals in the context of books.

A. Informal Reviews

The categories defined for these reviews try to cover the greater part of the aspects mentioned by the readers. The difference between aspects (also called targets) and categories is that the first ones are the specific entities mentioned explicitly in the text, whilst the second ones are a more coarse-grained classification of those entities. We can see the difference between both in the example of the Figure 2.

A distinction was made between the categories related to the book itself (starting with *"B#"*) and the ones related to the content of the book (starting with *"C#"*) and are the following: 1) B#GENERAL, when the reviewer makes reference to a general aspect about the book; 2) B#QUALITY for the reviews talking about the quality of the book, such as the way of writing; 3) B#STRUCTURE refers to aspects about the chapters, index, summary or other features related to the structure of the book; 4) B#AUTHOR, when mentioning information about the author; 5) B#PERIOD, which is related to the publishing date; 6) B#TITLE for features related to the title of the book; 7) B#AUDIENCE refers to the kind of readers which the book was written for; 8) B#PRICE, when talking about the price of the book; 9) B#LENGTH when talking about the number of pages or the size; 10) C#PLOT is related to the storyline; 11) C#CHARACTERS for reviews which talk about any aspect of the characters in the storyline; 12) C#PERIOD referring to the period when the plot passes and 13) C#GENRE, related to the literary genre of the book.

This set of categories was already used to create a new dataset of Amazon reviews annotated with the aspects extracted, along with the categories to which they belong and the sentiment associated, and it is publicly available online [3]. More information about this dataset, as well as the annotation process, can be found in [6] and in Figure 2, an example of an annotated sentence from this dataset is presented.

When observing the aspects annotated in this dataset, we realise that most of them are related to the *plot* (35.06% of the aspects annotated) or *characters* (27.84%). However, we rarely find this kind of information in OpenEdition reviews. Moreover, by inspecting the specific aspects manually extracted from these reviews, it can be observed that they appear very infrequently in OpenEdition ones.

[3]http://www.gti.uvigo.es/index.php/en/book-reviews-annotated-dataset-for-aspect-based-sentiment-analysis

B. Formal Reviews

For the professional reviews from OpenEdition, some of the categories proposed were taken from the previous list, as some common features can be found in both: 1) B#AUTHOR 2) B#PERIOD 3) B#TITLE 4) B#AUDIENCE 5) B#PRICE and 6) C#PERIOD.

The rest of the categories were defined by manually inspecting several reviews from OpenEdition and different studies published about good practices for writing scholarly book reviews. In [15] the authors expose the objectives to achieve when writing reviews or scholarly articles, as well as some standards and guidelines, distinguishing also different types of scholarly reviews. Some of these standards mentioned are scientific originality, research methods, clarity of the results, etc. In [16] they explore how academics from different disciplines perform the task of reading and writing book reviews for journals, providing some results from several questionnaires that were filled in by different professionals. Then in [17] the process of writing a review is addressed and some desirable and undesirable characteristics that a scholarly book review should have are also presented.

After analyzing different studies about professional book reviews, we decided to add the follwing categories: 1) B#JUDGEMENT makes reference to the opinion the reviewer presents about the book; 2) B#ORGANIZATION refers to the development of the different sections or chapters in the book, the structure and the format; 3) B#WRITING STYLE, which is about the way of writing of the author and the rhetorical devices utilized; 4) C#TECHNICAL FEATURES, when the review talks about the table of contents, illustrations, figures, etc.; 5) C#PRESENTATION includes those sentences mentioning a summary of the contents of the book or the subject addressed; 6) C#SCIENTIFIC CONTEXT refers to the state of the art or the situation of the book amongst other works about the same subject; 7) C#ARGUMENTATION, when discussing the principles presented by the author for reaching the conclusion and 8) C#METHODOLOGY will be tagged if the review talks about the scientific methods applied, if it is the case, to reach the results expounded in the book.

This dataset has not been manually annotated yet for this work. However, some examples were prepared and annotated in order to understand the new categories designed:

- The author's discussion of these problems drawn from real life cases do an excellent job of explaining how the process works. - B#JUDGEMENT
- I agree with his general point about the specialized studies of civil-military relations but fear that the SMI model he proposes is impractical because it casts too broad a net. - C#ARGUMENTATION
- The book's introduction also contains a good discussion of our field's evolution over the past several decades. - B#ORGANIZATION, C#PRESENTATION
- As an encoded narrative, SF relies on a style and on a specific language which transforms the reading of an SF story into "an active process of translation". -

```
−<sentence id="000_0006753000_31:12">
 −<text>
    I enjoyed the story, but I still don't think its right for pre-teens
  </text>
 −<Opinions>
    <Opinion category="B#GENERAL" occurrence="1" polarity="positive" target="story"/>
    <Opinion category="B#AUDIENCE" occurrence="1" polarity="negative" target="pre-teens"/>
  </Opinions>
</sentence>
```

Fig. 2. Example of an annotated sentence extracted from an Amazon review.

B#WRITING STYLE

- From Edgar Allan Poe to H. G. Wells, Jules Verne and the Vernians on both sides of the Atlantic, Stableford maps out the literary production of the period to describe the overlapping of scientific romance with both utopian and dystopian categories and its hybridization with other literary forms. - C#SCIENTIFIC CONTEXT
- This initial reserve put aside, the volume is an ideal up-to-date companion for advanced undergraduate and post-graduate students as well as a reliable teaching resource with a chronology, a bibliography (from which websites are absent) and a useful index complementing sophisticated essays. - B#AUDIENCE, B#ORGANIZATION
- The first striking characteristic of this monograph is the beautiful full-page illustration on the cover. - C#TECHNICAL FEATURES

VI. Conclusions

In this paper an analysis and classification of two different document collections (formal reviews from OpenEdition about scholarly books and more informal ones about fiction books from Amazon) are performed in order prove the need for different models and annotation schemas for ABSA. Then new schemas are then proposed, which will be useful for later applying the aspect extraction task in the domain of books, and the sentiment detection associated to each one. There is not much research in ABSA dealing with book reviews in general and even less dealing with formal scholarly book reviews, and new datasets in this context are needed to develop new approaches. The analysis and schemas presented in this paper are the first steps to achieve this. Moreover, aspect extraction would be then useful for recommendation systems and other tasks dealing with structuring, classifying and organizing data, for example in digital libraries.

As future work, we plan to keep working in the aspect extraction task for books domain, mostly for scholarly reviews, as we think it is a bigger challenge, still not exploited in the state of the art. Then we will design a strategy to add these aspects as input for a recommendation system and so improve the quality of the recommendations.

Acknowledgment

This research was supported by the French ANR program “Investissements d’Avenir” EquipEx DILOH (ANR-11-EQPX-0013).

References

[1] M. Koolen, T. Bogers, M. Gäde, M. Hall, I. Hendrickx, H. Huurdeman, J. Kamps, M. Skov, S. Verberne, and D. Walsh, “Overview of the CLEF 2016 social book search lab,” International Conference of the Cross-Language Evaluation Forum for European Languages, pp. 351–370, 2016.

[2] K. Schouten, F. Frasincar, “Survey on Aspect-Level Sentiment Analysis,” IEEE Transactions on Knowledge and Data Engineering, pp. 813–830, 2016.

[3] G. Ganu, N. Elhadad, and A. Marian, “Beyond the stars: improving rating predictions using review text content,” WebDB 2009, pp. 1–6.

[4] M. Hu, and B. Liu, “Mining and summarizing customer reviews,” Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining, 2004, pp. 168–177.

[5] L. Chen, G. Chen, and F. Wang, “Recommender systems based on user reviews: the state of the art,” User Modeling and User-Adapted Interaction, vol. 25, pp. 99–154, 2015.

[6] T. Álvarez-López, M. Fernández-Gavilanes, E. Costa-Montenegro, J. Juncal-Martínez, S. García-Méndez, and P. Bellot, “A book reviews dataset for aspect based sentiment analysis,” Proceedings of 8th Language & Technology Conference, pp. 49–53, 2017.

[7] M. Al-Smadi, O. Qawasmeh, B. Talafha, and M. Quwaider, “Human annotated arabic dataset of book reviews for aspect based sentiment analysis,” Future Internet of Things and Cloud (FiCloud), 2015 3rd International Conference on, IEEE, pp. 726–730, 2015.

[8] C. Freitas, E. Motta, R.L. Milidiú, and J. César, “Sparkling vampire... lol! annotating opinions in a book review corpus,” New Language Technologies and Linguistic Research: A Two-Way Road. Cambridge Scholars Publishing, pp. 128–146, 2014.

[9] B. Kessler, G. Numberg, and H. Schtze, “Automatic detection of text genre,” Proceedings of 8th Conference on European chapter of the Association for Computational Linguistics, pp. 32-38, 1997.

[10] E. Stamatatos, N. Fakotakis, and G. Kokkinakis, “Automatic text categorization in terms of genre and author,” Computational linguistics, 26(4), pp. 471-495, 2000.

[11] C. S. Lim, K. J. Lee, and G. C. Kim, “Multiple sets of features for automatic genre classification of web documents,” Information processing & management, 41(5), pp. 1263-1276, 2005.

[12] M. Sokolova, and G. Lapalme, “A systematic analysis of performance measures for classification tasks,” Information Processing & Management, pp. 427–437, 2009.

[13] M. Taboada, J. Brooke, M. Tofiloski, K. Voll, and M. Stede, “Lexicon-based methods for sentiment analysis,” Computational Linguistics. 37, 267-307, 2011.

[14] H. Hamdam, P. Bellot, and F. Béchet, “The impact of z_score on Twitter sentiment analysis,” 8th International Workshop on Semantic Evaluation (SemEval), pp. 636-641, 2014.

[15] F. Dochy, “A guide for writing scholarly articles or reviews for the Educational Research Review,” Educational Research Review, pp. 1–6, 2006.

[16] J. Hartley, “Reading and writing book reviews across the disciplines,” Journal of the Association for Information Science and Technology, pp. 1194–1207, 2006.

[17] A. Lee, B. Green, C. Johnson, and J. Nyquist, “How to write a scholarly book review for publication in a peer-reviewed journal: a review of the literature,” Journal of Chiropractic Education, pp. 57–69, 2010.