Towards Unsupervised Approaches For Aspects Extraction

Marco Federici[1,2] and Mauro Dragoni[2]

[1] Universitá di Trento, Italy
[2] Fondazione Bruno Kessler, Trento, Italy
federici|dragoni@fbk.eu

Abstract. One of the most recent opinion mining research directions falls in the extraction of polarities referring to specific entities (called "aspects") contained in the analyzed texts. The detection of such aspects may be very critical especially when the domain which documents belong to is unknown. Indeed, while in some contexts it is possible to train domain-specific models for improving the effectiveness of aspects extraction algorithms, in others the most suitable solution is to apply unsupervised techniques by making the used algorithm independent from the domain. In this work, we implemented different unsupervised solutions into an aspect-based opinion mining system. Such solutions are based on the use of semantic resources for performing the extraction of aspects from texts. The algorithms have been tested on benchmarks provided by the SemEval campaign and have been compared with the results obtained by domain-adapted techniques.

1 Introduction

Opinion Mining is a natural language processing (NLP) task that aims to classify documents according to their opinion (polarity) on a given subject [1]. This task has created a considerable interest due to its wide applications in different domains like marketing, politics, and social sciences. Generally, the polarity of a document is computed by analyzing the expressions contained in the full text by leading to the issue of not distinguishing which are the subjects of each opinion. Therefore, the natural evolution of the opinion mining research field has been focused on the extraction of all subjects ("aspects") from texts in order to make systems able to compute the polarity associated to each aspect in an independent way [2].

Let us consider the following example:

Yesterday, I bought a new smartphone.
The quality of the display is very good, but the buttery lasts too little.

In the sentence above, we may identify three aspects: "smartphone", "display", and "battery". Each aspect has a different opinion associated with it, in particular:

- "display" → "very good"
- "battery" → "too little"
- "smarthphone" → no explicit opinions, therefore its polarity can be inferred by averaging the opinions associated with all other aspects.

Another important consideration related to this example is that it is easy to detect which is the domain of the analyzed text. In this case, by assuming to have a training set, it should be possible to build domain-specific models for supporting the extraction of the aspects. However, this strategy is in contrast with two considerations coming from real-world scenarios: (i) it is difficult to find annotated dataset related to all possible domains, and (ii) in the same document, it is possible to have sentences belonging to many domains by making the adoption of a domain-specific models not feasible.

To overcome these issues, we propose a set of unsupervised approaches based on natural language processing approaches that do not rely to any domain-specific information. The goal of this study is to provide techniques that are able to reach an effectiveness comparable with supervised systems.

The paper is structured as follows. In Section 2, we provide an overview of the opinion mining field with a focus on aspects extraction approaches. Section 3 presents the natural language processing layer built for supporting the approaches described in Sections 4 and 5. Section 6 discusses the performance of each algorithm; while, Section 7 concludes the paper.

2 Related Work

The topic of opinion mining has been studied extensively in the literature [3,4], where several techniques have been proposed and validated.

All the approaches presented so far operate at the document-level[5,6]; while, for improving the accuracy of the opinion classification, a more fine-grained analysis of the text, i.e., the opinion classification of every single sentence has to be performed [7,8]. In the literature, we may find approaches ranging from the use of fuzzy logic [9,10] to the use aggregation techniques [11] for computing the score aggregation of opinion words. In the case of sentence-level opinion classification, two different sub-tasks have to be addressed: (i) to determine if the sentence is subjective or objective, and (ii) in the case that the sentence is subjective, to determine if the opinion expressed in the sentence is positive, negative, or neutral. The task of classifying a sentence as subjective or objective, called "subjectivity classification", has been widely discussed in the literature [7,8] and systems implementing the capabilities of identifying opinion's holder, target, and polarity have been presented [12].

The growth of online product reviews was the perfect floor for using opinion mining techniques in marketing activities. The issue of detecting the different opinions concerning the same product expressed in the same review emerged as a challenging problem. Such a task has been faced by introducing *aspect* extraction approaches aiming to extract, from each sentence, which is the aspect the opinion refers to. In the literature, many approaches have been proposed: conditional random fields (CRF) [13,14], hidden Markov models (HMM) [15,16,17], sequential rule mining [18], dependency tree kernels [19], clustering [20], and genetic algorithms [21]. In [22,23], a method was proposed to extract both opinion words and aspects simultaneously by exploiting some syntactic relations of opinion words and aspects.

At the same time, the social dimension of the Web opens up the opportunity to combine computer science and social sciences to better recognize, interpret, and process opinions and sentiments expressed over it. Such multi-disciplinary approach has been called *sentic computing* [24].

Above, we mentioned approaches that do not consider the domain analyzed documents belong to. The use of domain adaptation demonstrated that opinion classification is highly sensitive to the domain from which the training data is extracted. A classifier trained using opinionated documents from one domain often performs poorly when applied or tested on opinionated documents from another domain. The reason is that using the same words and even the same language constructs can carry different opinions, depending on the domain.

The classic scenario is when the same word in one domain may have positive connotations, but in another domain may have negative one; therefore, domain adaptation is needed. In the literature, different approaches related to the Multi-Domain sentiment analysis have been proposed. Briefly, two main categories may be identified: (i) the transfer of learned classifiers across different domains [25,26,27,28], and (ii) the use of propagation of labels through graph structures [29,30,9].

While on one side such approaches demonstrated their effectiveness in working in a multi-domain environment, on the other one, they suffer by the limitation in abstracting their usage within any domain different from the ones used for building the model.

3 The Underlying NLP Layer

A number of different approaches has been tested in order to accomplish aspect extraction task. Each one uses different functionalities offered by the Stanford NLP Library but every technique is characterized by a common preliminary phase.

First of all, *WordNet*[3] [31] resource is used together with Stanford's part of speech annotation to detect compound nouns. Lists of consecutive nouns and word sequences contained in Wordnet compound nouns vocabulary are merged into a single word in order to force Stanford library to consider them as a single unit during the following phases.

The entire text is then fed to the co-reference resolution module to compute pronoun references which are stored in an index-reference map.

The next operation consists in detecting which word expresses polarity within each sentence. To achieve this task *SenticNet*[4] [32], *General Inquirer dictionary* [5] [33] and *MPQA*[6] [34] sentiment lexicons have been used.

While SenticNet expresses polarity values in the continuous range from -1 to 1, the other two resources been normalized: the General Inquirer words have positive values of polarity if they belong to the "Positiv" class while negative if they belong to "Negativ" one, zero otherwise, similarly, MPQA "polarity" labels are used to infer a numerical values. Only words with a non-zero polarity value in at least one resource are considered as opinion words (e.g. word "third" is not present in MPQA and SenticNet and has a 0 value according to General Inquirer, consequently, it is not a valid opinion word; on the other hand, word "huge" has a positive 0.069 value according to SenticNet, a negative value in MPQA and 0 value according to General Inquirer, therefore, it is a possible opinion word even if lexicons express contrasting values). Every noun (single

[3] https://wordnet.princeton.edu/
[4] http://sentic.net/
[5] http://www.wjh.harvard.edu/ inquirer/spreadsheet_guide.htm
[6] http://mpqa.cs.pitt.edu/corpora/mpqa_corpus/

or complex) is considered an aspect as long as it's connected to at least one opinion and it's not in the stopword list. This list has been created starting from the "Onix" text retrieval engine stopwords list[7] and it contains words without a specific meaning (such as "thing") and special characters.

Opinions associated with pronouns are connected to the aspect they are referring to; instead, if pronouns reference can't be resolved, they are both discarded.

The main task of the system is, then, represented by connecting opinions with possible aspects. Two different approaches have been tested with a few variants. The first one relies on the syntactic tree while the second one is based on grammar dependencies.

The sentence "I enjoyed the screen resolution, it's amazing for such a cheap laptop." has been used to underline differences in connection techniques.

The preliminary phase merges words "screen" and "resolution" into a single word "Screenresolution" because they are consecutive nouns. Co-reference resolution module extracts a relation between "it" and "Screenresolution". This relation is stored so that every possible opinion that would be connected to "it" will be connected to "Screenresolution" instead. Figure 1 shows the syntax tree while Figure 2 represents the grammar relation graph generated starting from the example sentence. Both structures have been computed using Stanford NLP modules ("parse", "depparse").

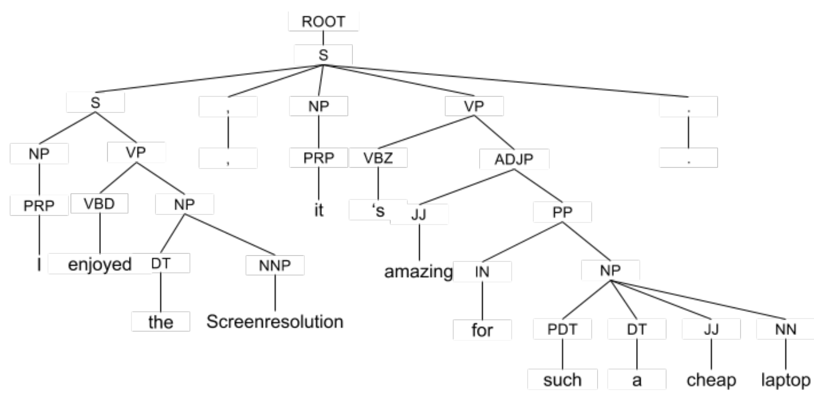

Fig. 1: Example of syntax tree.

4 Unsupervised Approaches - Syntax-Tree-Based Approach

These typologies of approaches are based on syntax tree structures created by Stanford NLP library. In order to explain how the algorithms connect opinion with aspects a few definition are needed:

- "Intermediate node": tree node which is not a leaf;
- "Sentence node": intermediate node labeled with one of the following:

[7] The used stopwords list is available at http://www.lextek.com/manuals/onix/stopwords1.html

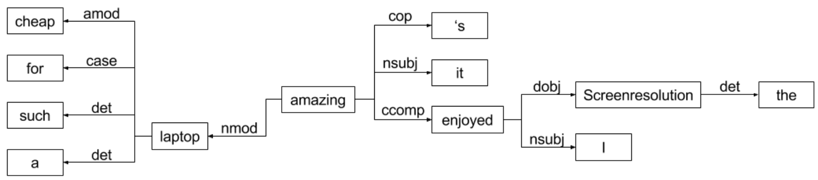

Fig. 2: Example of the grammar relations graph.

  - ROOT - Root of the tree
  - S - Sentence
  - SBAR - Clause introduced by a (possibly empty) subordinating conjunction
  - SBARQ - Direct question introduced by a wh-word or a wh phrase
  - SQ - Inverted yes/no question or main clause of a wh-question
  - SINV - Inverted declarative sentence
  - PRN - Parenthetical
  - FRAG - Fragment
- "Noun Phrase node": intermediate node labeled with NP tag

Approaches differ in rules adopted for associating intermediate nodes that define how aspects are extracted by starting from their child nodes.

Approach 1.1 Each polarized adjective is connected with each possible aspect in the same sentence.

Figure 3 shows she propagation of aspects and opinion in the tree with red lines representing propagation of aspects, blue lines for opinions and purple ones when both are propagated to the upper level.

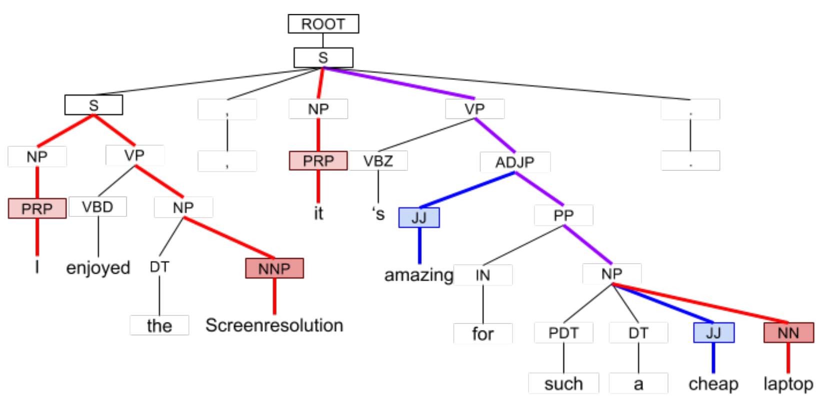

Fig. 3: Parser tree generated by the approach 1.1.

Within the sub-sentence "I enjoyed the Screenresolution" only aspects are detected, consequently, once the Sentence Level node is reached, no connection is done. On the other hand, both polarized adjectives "cheap" and "amazing" are propagated until they reach the top sentence node together with “it” and “laptop” aspects, then, they are connected with each other.

The results are shown in Figure 4.

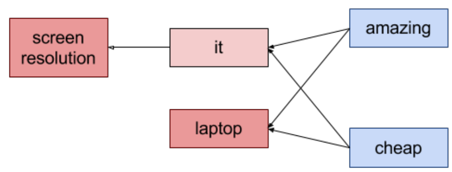

Fig. 4: Relationships generated by the approach 1.1.

Approach 1.2 Each polarized adjective is connected to each possible aspect within the same sentence or noun phrase.

Influences of this variant are underlined in Figure 5 with the same notation.

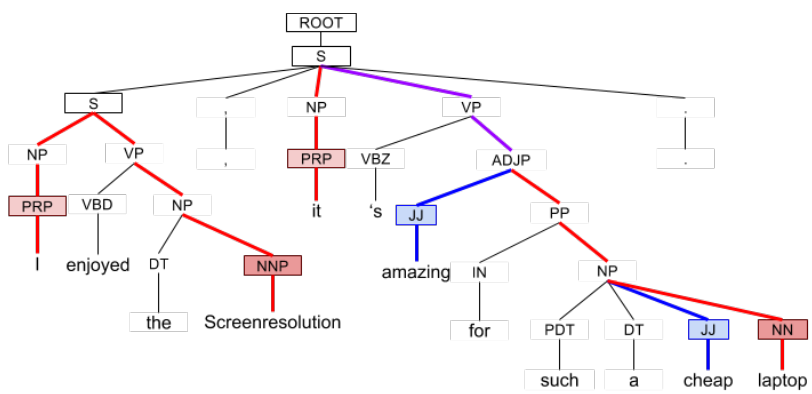

Fig. 5: Parser tree generated by the approach 1.2.

Even if extracted aspects are the same, the opinion “cheap” is associated only with the name “laptop” as shown in Figure 6.

Approach 1.3 When both aspects set and opinion words set related to a node are not empty, each opinion word is connected to the related aspect and removed from the opinion words set. Opinion words and possible aspects are removed anyway in sentence nodes.

Figure 7 shows the effects of the association rules mentioned above.

Once again, even if aspects extracted are the same, the connections are different (Figure 8).

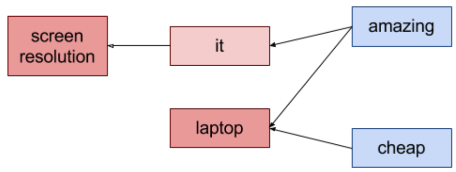

Fig. 6: Relationships generated by the approach 1.2.

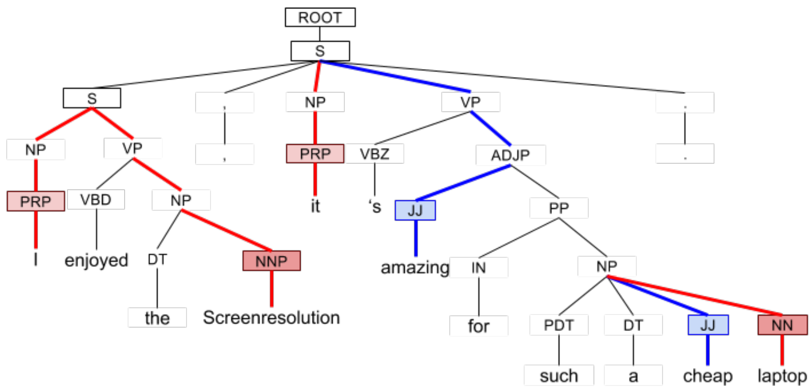

Fig. 7: Parser tree generated by the approach 1.3.

5 Unsupervised Approaches - Grammar-Dependencies-based Approach

The other set of approaches proposed in this paper exploits grammar dependencies instead of syntax tree to detect aspect-opinion associations. Grammar dependencies computed by Stanford NLP modules (which are represented by the labeled graph in picture [1.2]) can be expressed by triples: $\{Relationtype, Governor, Dependant\}$. One of the most important difference with the previous methodology is represented by the possibility of detecting opinion expressed by word that are not adjectives (such as verbs that are considered by approaches 2.2 and 2.3). Different approaches have been tested in order to detect which kind of triple can be interpreted as a connection between an opinion word and a possible aspect.

Approach 2.1 The following two rules are implemented:

Rule 1: Each adjectival modifier (amod) relation expresses a connection between an aspect and an opinion word if and only if the governor is a possible aspect and the dependant is a polarized adjective.

Rule 2: Each nominal subject (nsubj) relation expresses a connection between an aspect and an opinion word if and only if the governor is a polarized opinion and the dependant is a possible aspect.

Figure 9 underlines aspect-opinion connections mined through the process.

Resulting aspects are shown in Figure 10.

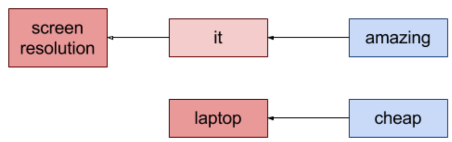

Fig. 8: Relationships generated by the approach 1.3.

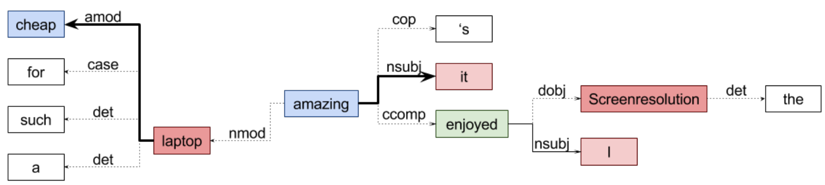

Fig. 9: Parser tree generated by the approach 2.1.

Approach 2.2 The Rules "1" and "2" are both used, in addition a third rule is introduced:

Rule 3: Each direct object (dobj) relation expresses a connection between an aspect and an opinion word if and only if the governor is a polarized word and the dependant is a possible aspect.

Figure 11 and 12 shows the results of the aspect detection process with the addition of the direct object relation.

Approach 2.3 The Rules "1" and "3" are both used, while Rule "2" is changed as follows:

Rule 2.1: Each nominal subject (nsubj) relation expresses a connection between an aspect and an opinion word if and only if the governor is a polarized word and the dependant is a possible aspect.

Figure 13 shows results of the modification of the rules. Even if the relation between "enjoyed" and "I" is detected, "I" is not considered as a valid aspect since it's has an unresolved reference in the current context.

Results are the same as the previous example (Figure 14).

6 Evaluation

Each approach has been tested on two datasets provided by the Task 12 of SemEval 2015 evaluation campaign, namely "Laptop" and "Restaurant". To evaluate results a notion of correctness has to be introduced: if the extracted aspects is equal, contained or contains the correct one, it's considered to be correct (for example if the extracted aspect is "screen", while the annotated one is "screen of the computer" or vice versa, the result of the system is considered to correct).

Tables 1 and 2 shows the number of "True Positive", "False Positive", and "False Negative" computed on mentioned datasets. The rationale behind this choice is to support the error analysis provided later and for showing strong and weak points of each aspect.

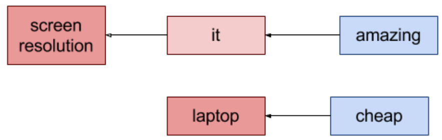

Fig. 10: Relationships generated by the approach 2.1.

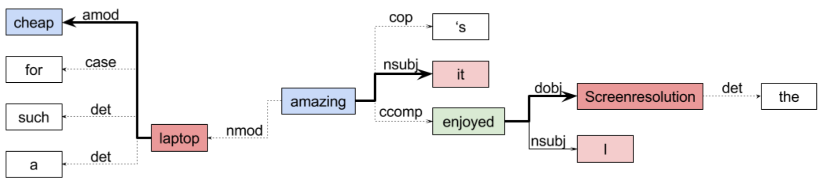

Fig. 11: Parser tree generated by the approach 2.2.

Laptop	True Positives	False Positives	False Negatives
1.1	255	459	396
1.2	211	341	440
1.3	186	322	465
2.1	155	213	496
2.2	225	318	426
2.3	257	386	394

Table 1

To evaluate performance on the "Restaurants" dataset, "null" aspect has not been considered in false negatives count.

Restaurants	True Positives	False Positives	False Negatives
1.1	316	381	197
1.2	267	246	246
1.3	259	235	254
2.1	219	161	294
2.2	238	275	275
2.3	287	330	226

Table 2

Figure 15 shows an analysis of error cases. Values have been computed according to the first 100 sentences of the "Laptop" dataset.

The majority of false negatives are given by the impossibility to detect opinions expressed by verbs. For example, in the sentence "I generally like this place" or more

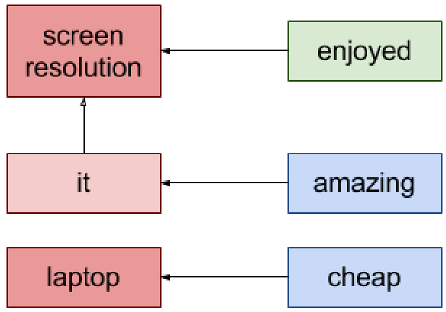

Fig. 12: Relationships generated by the approach 2.2.

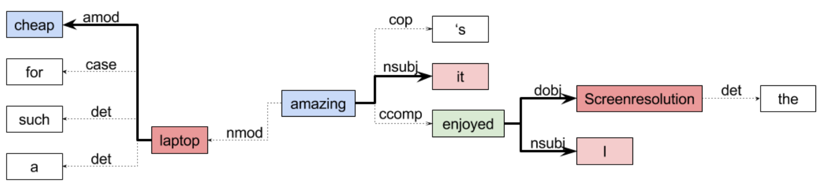

Fig. 13: Parser tree generated by the approach 2.3.

complex expressions "tech support would not fix the problem unless I bought your plan for $150 plus".

Other issues are correlated to the association algorithm. Figures 16 and 17 show error categories in approaches 1.3 and 2.1 respectively, always computed on the same 100 sentences of the "Laptop" dataset.

Even if the syntax-tree-based approach tends to produce a significant number of true positives, relationships are often imprecise. A relevant example is represented by the sentence "I was extremely happy with the OS itself." in the "Laptop" dataset. Approach 1.3 connects the opinion adjective "happy" with the potential aspect "OS", correctly recognized as an aspect in the sentence, while approach 2.1 does not detect such a relation because the word "happy" is connected to "I" which is not a potential aspect.

A relevant part of false positives are generated by approaches that are not able to discriminate aspects from the entity itself. In facts, almost half of them consists in associations between opinion words and the entity reviewed that are correct. However, they must not be considered during the aspect extraction task (for example the aspect "laptop" in the example sentence should not be considered according to the definition of aspect).

7 Conclusions

In this paper, we presented a set of unsupervised approaches for aspect-based sentiment analysis. Such approaches have been tested on two SemEval benchmarks: the "Laptop" and "Restaurant" datasets used in the Task 12 of SemEval 2015 evaluation campaign. Results demonstrated how without using learning techniques the results can be compa-

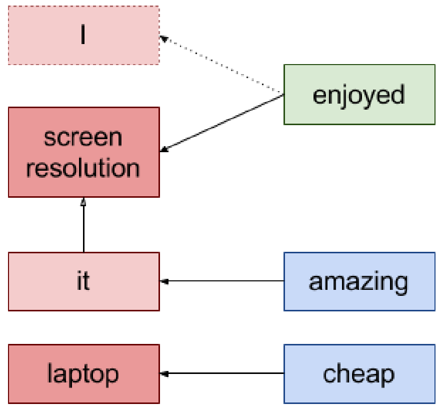

Fig. 14: Relationships generated by the approach 2.3.

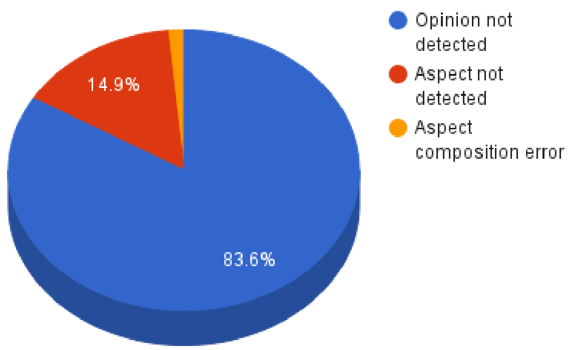

Fig. 15: Overall error analysis.

rable with the ones obtained by trained systems. Future work includes refinement of the proposed approaches in order to make them suitable for real-world implementation.

References

1. Pang, B., Lee, L., Vaithyanathan, S.: Thumbs up? sentiment classification using machine learning techniques. In: Proceedings of EMNLP, Philadelphia, Association for Computational Linguistics (July 2002) 79–86

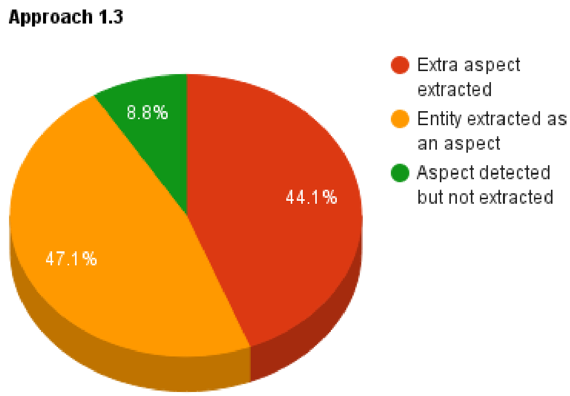

Fig. 16: Error analysis of Approach 1.3.

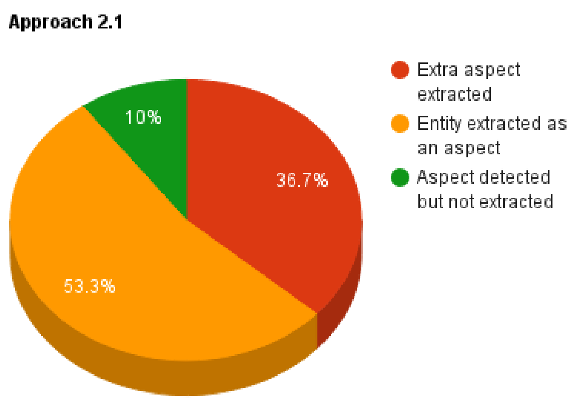

Fig. 17: Error analysis of Approach 1.3.

2. Hu, M., Liu, B.: Mining and summarizing customer reviews. In: Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining, ACM (2004) 168–177
3. Pang, B., Lee, L.: Opinion mining and sentiment analysis. Foundations and Trends in Information Retrieval **2**(1-2) (2008) 1–135

4. Liu, B., Zhang, L.: A survey of opinion mining and sentiment analysis. In Aggarwal, C.C., Zhai, C.X., eds.: Mining Text Data. Springer (2012) 415–463
5. Dragoni, M.: Shellfbk: An information retrieval-based system for multi-domain sentiment analysis. In: Proceedings of the 9th International Workshop on Semantic Evaluation. SemEval '2015, Denver, Colorado, Association for Computational Linguistics (June 2015) 502–509
6. Petrucci, G., Dragoni, M.: An information retrieval-based system for multi-domain sentiment analysis. In Gandon, F., Cabrio, E., Stankovic, M., Zimmermann, A., eds.: Semantic Web Evaluation Challenges - Second SemWebEval Challenge at ESWC 2015, Portorož, Slovenia, May 31 - June 4, 2015, Revised Selected Papers. Volume 548 of Communications in Computer and Information Science., Springer (2015) 234–243
7. Riloff, E., Patwardhan, S., Wiebe, J.: Feature subsumption for opinion analysis. In: EMNLP. (2006) 440–448
8. Wilson, T., Wiebe, J., Hwa, R.: Recognizing strong and weak opinion clauses. Computational Intelligence **22**(2) (2006) 73–99
9. Dragoni, M., Tettamanzi, A.G., da Costa Pereira, C.: Propagating and aggregating fuzzy polarities for concept-level sentiment analysis. Cognitive Computation **7**(2) (2015) 186–197
10. Dragoni, M., Tettamanzi, A.G.B., da Costa Pereira, C.: A fuzzy system for concept-level sentiment analysis. In Presutti, V., Stankovic, M., Cambria, E., Cantador, I., Iorio, A.D., Noia, T.D., Lange, C., Recupero, D.R., Tordai, A., eds.: Semantic Web Evaluation Challenge - SemWebEval 2014 at ESWC 2014, Anissaras, Crete, Greece, May 25-29, 2014, Revised Selected Papers. Volume 475 of Communications in Computer and Information Science., Springer (2014) 21–27
11. da Costa Pereira, C., Dragoni, M., Pasi, G.: A prioritized "and" aggregation operator for multidimensional relevance assessment. In Serra, R., Cucchiara, R., eds.: AI*IA 2009: Emergent Perspectives in Artificial Intelligence, XIth International Conference of the Italian Association for Artificial Intelligence, Reggio Emilia, Italy, December 9-12, 2009, Proceedings. Volume 5883 of Lecture Notes in Computer Science., Springer (2009) 72–81
12. Aprosio, A.P., Corcoglioniti, F., Dragoni, M., Rospocher, M.: Supervised opinion frames detection with RAID. In Gandon, F., Cabrio, E., Stankovic, M., Zimmermann, A., eds.: Semantic Web Evaluation Challenges - Second SemWebEval Challenge at ESWC 2015, Portorož, Slovenia, May 31 - June 4, 2015, Revised Selected Papers. Volume 548 of Communications in Computer and Information Science., Springer (2015) 251–263
13. Jakob, N., Gurevych, I.: Extracting opinion targets in a single and cross-domain setting with conditional random fields. In: EMNLP. (2010) 1035–1045
14. Lafferty, J.D., McCallum, A., Pereira, F.C.N.: Conditional random fields: Probabilistic models for segmenting and labeling sequence data. In: ICML. (2001) 282–289
15. Freitag, D., McCallum, A.: Information extraction with hmm structures learned by stochastic optimization. In: AAAI/IAAI. (2000) 584–589
16. Jin, W., Ho, H.H.: A novel lexicalized HMM-based learning framework for web opinion mining. In: Proceedings of the 26th Annual International Conference on Machine Learning. ICML '09, New York, NY, USA, ACM (2009) 465–472
17. Jin, W., Ho, H.H., Srihari, R.K.: Opinionminer: a novel machine learning system for web opinion mining and extraction. In: KDD. (2009) 1195–1204
18. Liu, B., Hu, M., Cheng, J.: Opinion observer: analyzing and comparing opinions on the web. In: WWW. (2005) 342–351
19. Wu, Y., Zhang, Q., Huang, X., Wu, L.: Phrase dependency parsing for opinion mining. In: EMNLP. (2009) 1533–1541
20. Su, Q., Xu, X., Guo, H., Guo, Z., Wu, X., Zhang, X., Swen, B., Su, Z.: Hidden sentiment association in chinese web opinion mining. In: WWW. (2008) 959–968
21. Dragoni, M., Azzini, A., Tettamanzi, A.: A novel similarity-based crossover for artificial neural network evolution. In Schaefer, R., Cotta, C., Kolodziej, J., Rudolph, G., eds.: Parallel

Problem Solving from Nature - PPSN XI, 11th International Conference, Kraków, Poland, September 11-15, 2010, Proceedings, Part I. Volume 6238 of Lecture Notes in Computer Science., Springer (2010) 344–353
22. Qiu, G., Liu, B., Bu, J., Chen, C.: Expanding domain sentiment lexicon through double propagation. In: IJCAI. (2009) 1199–1204
23. Qiu, G., Liu, B., Bu, J., Chen, C.: Opinion word expansion and target extraction through double propagation. Computational Linguistics **37**(1) (2011) 9–27
24. Cambria, E., Hussain, A.: Sentic Computing: Techniques, Tools, and Applications. Volume 2 of SpringerBriefs in Cognitive Computation. Springer, Dordrecht, Netherlands (2012)
25. Blitzer, J., Dredze, M., Pereira, F.: Biographies, bollywood, boom-boxes and blenders: Domain adaptation for sentiment classification. In: ACL. (2007) 187–205
26. Pan, S.J., Ni, X., Sun, J.T., Yang, Q., Chen, Z.: Cross-domain sentiment classification via spectral feature alignment. In: WWW. (2010) 751–760
27. Bollegala, D., Weir, D.J., Carroll, J.A.: Cross-domain sentiment classification using a sentiment sensitive thesaurus. IEEE Trans. Knowl. Data Eng. **25**(8) (2013) 1719–1731
28. Yoshida, Y., Hirao, T., Iwata, T., Nagata, M., Matsumoto, Y.: Transfer learning for multiple-domain sentiment analysis—identifying domain dependent/independent word polarity. In: AAAI. (2011) 1286–1291
29. Ponomareva, N., Thelwall, M.: Semi-supervised vs. cross-domain graphs for sentiment analysis. In: RANLP. (2013) 571–578
30. Huang, S., Niu, Z., Shi, C.: Automatic construction of domain-specific sentiment lexicon based on constrained label propagation. Knowl.-Based Syst. **56** (2014) 191–200
31. Fellbaum, C.: WordNet: An Electronic Lexical Database. MIT Press, Cambridge, MA (1998)
32. Cambria, E., Speer, R., Havasi, C., Hussain, A.: Senticnet: A publicly available semantic resource for opinion mining. In: AAAI Fall Symposium: Commonsense Knowledge. (2010)
33. P.J, S., Dunphy, D., Marshall, S.: The General Inquirer: A Computer Approach to Content Analysis. Oxford, England: M.I.T. Press (1966)
34. Deng, L., Wiebe, J.: MPQA 3.0: An entity/event-level sentiment corpus. In Mihalcea, R., Chai, J.Y., Sarkar, A., eds.: NAACL HLT 2015, The 2015 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Denver, Colorado, USA, May 31 - June 5, 2015, The Association for Computational Linguistics (2015) 1323–1328