Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 207 (2022) 2688–2697

Procedia Computer Science

www.elsevier.com/locate/procedia

26th International Conference on Knowledge-Based and Intelligent Information & Engineering Systems (KES 2022)

Evaluation of weakly-supervised methods for aspect extraction

Mohamed Ettaleb[a], Amira Barhoumi[a], Nathalie Camelin[a], Nicolas Dugué[a]

[a] *Le Mans University, LIUM, Le Mans 72000, France*

Abstract

Aspect-based sentiment analysis (ABSA) may provide more detailed information than general sentiment analysis. It aims to extract aspects from reviews and predict their polarities. In this paper, we focus on aspect extraction sub-task. We propose three weakly-supervised systems based on contextual language models and topic modeling. We evaluate and compare our systems on SemEval-2016 restaurant french benchmark. The experimental results reveal that our systems is quite competitive in aspect extraction from user reviews. We obtain 60.65% as F1 score with our best system. The latter outperforms the existing supervised ones. We deduct that weakly-supervised systems are efficient in terms of performance, time and human effort.

Keywords: Sentiment analysis; Aspect extraction; Word embeddings; CNN; LDA; Topic Modeling; Dataless Text Classification; Text Analysis

1. Introduction

Over last decades, Web has become among the most important source for customers and suppliers to evaluate and compare products and services. Therefore, Sentiment Analysis (SA) become an important research area. A classic task of SA involves determining the overall polarity expressed in a review (document level SA) or in a sentence (sentence level SA). Another more, refine-grained SA task is called Aspect-Based SA (ABSA). It deals with detecting polarities for different aspects in the review. ABSA has two sub-tasks: (i) extracting aspects from each review, and (ii) determining polarities for different extracted aspects. In this work, we focus on Aspect Extraction sub-task (AE). It consists in identifying and extracting the different aspects mentioned in a review.

Most of the existing work for AE task is based on supervised learning methods [33] that need considerable human review-annotation efforts to build training corpus. In particular, recent deep learning methods outperform existing ones [31], but need a significant volume of annotated data for training [4]. To minimize the labeling effort, semi-supervised and weakly-supervised algorithms (the latter known as data-less methods) have been proposed for text classification

E-mail address: FirstName.LastName@univ-lemans.fr

10.1016/j.procs.2022.09.327

[10]. These algorithms generally need specific resources (i.e. small labeled corpus for semi-supervised algorithms , keyword lists for weakly-supervised ones [14]. A more detailed description of different methods and approaches are explored in the survey conducted in [1].
In this work, we propose three weakly-supervised methods and evaluate them on SemEval-2016 french restaurant benchmark [24]. The choice of weakly-supervised learning is justified by the small size of the benchmark. Our proposed methods use predefined human seeds, *i.e* specific keywords that are related to the aspect. For example, seeds of the "Food" aspect could be : *menu, dish, salad, pizza, . . .* Hence, The seed list contains words which contribute towards discriminating the different aspects.
Our proposed systems do not need any labeled data but only some seeds, *i.e.* a few keywords for each aspect. All proposed systems follows this behaviour : (1) when an aspect is expressed in a review by more than one word, our systems does not extract each word as a different aspect; (2) when a review covers more than one aspect, our systems are able to extract the different aspects in the review. We evaluate our systems on SemEval-2016 french restaurant benchmark [24] for aspect extraction and demonstrate that our weakly-supervised systems are as efficient as the supervised ones.

2. Related Work

Sentiment analysis has been studied at document, sentence and aspect levels. This work focuses on the aspect level [9]. Aspect-based sentiment analysis (ABSA) task can be decomposed into two sub-tasks: aspect extraction and polarity identification. The former extracts aspects in the review, and the latter identifies the polarity for each extracted aspect. Existing work can be classified in 2 categories. The first one deals with the two sub-tasks separately [27], *i.e.* aspect extraction sub-task is followed by polarity identification. The second category jointly solves the two sub-tasks. [10] supposes that leveraging more fine-grained information by coupling two sub-tasks could enhance both.

In this work, we focus mainly on aspect extraction sub-task. Various methods have been proposed. The first methods are based on manually defined rules [35]. Frequency-based rules were helpful to identify frequent items as product features [15]. In order to enhance rules-based methods, statistics were used to eliminate false positives [15]. Grammatical dependencies were also incorporated to remove infrequent aspects [12]. Syntax-based methods, based on underlying syntactic dependencies, have been experimented to extract aspects [25]. They assume existence of a powerful syntactic parser. They also suppose that reviews are well-written and respect the language grammar.
Early supervised methods are based on Conditional Random Fields (CRF) [28]. In order to learn better features, neural networks are applied later for aspect extraction, *e.g.*, Recurrent Neural Networks (RNN) [16] and attention mechanism [8], Convolutional Neural Network (CNN) [32]. Existing work shows that supervised methods outperformed rules-based ones, but they require large fine-grained labeled corpora to train models, which may be difficult to build.

Several unsupervised methods are proposed to automatically model different aspects. LDA-based methods [20] model each review as a mixture of aspects and output a word distribution set for each aspect . [8] uses an auto-encoder to reconstruct sentences through aspect embedding and attention mechanism to remove irrelevant words. [30] introduces a single head attention calculated by a kernel function to be the sentence summary. The unsupervised and rule-based methods is hindered by the fact that extracted aspects often do not align well with user's interested aspects, and additional human effort is needed to map topics to certain aspects.

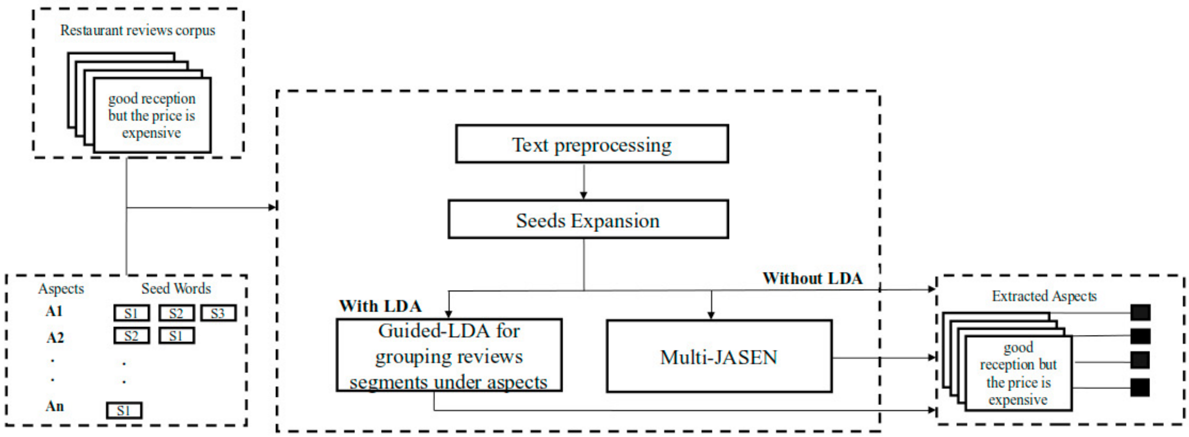
Weakly-supervised methods address aspect extraction problem by using a few keywords per aspect in order to guide and supervise the learning process. Some of these methods do not take aspect-specific polar words into consideration [18]. The semantic meaning captured by a joint (aspect, polarity) topic preserves more fine-grained information to determine the aspect in a review. Thus, it can be used to improve the performance of aspect extraction. The studies that jointly perform aspect and polarity extraction are LDA-based methods [7]. However, aspect-word distribution obtained with LDA-based methods can largely overlap among different aspects, allowing redundant aspects to appear and making them hard to classify. To solve this issue, [10] propose a tool that first learn (aspect, polarity) joint topic embeddings in the embedding space.
In this work, we propose an extension of this tool. Moreover, we propose two other methods based on aspect-word dis-

tribution through CamemBERT embeddings [17] with/without LDA. We evaluate and compare our weakly-supervised proposed methods on the french benchmark restaurant datasets (SemEval-2016) to demonstrate their effectiveness.

3. Systems description

The main objective of our systems is to perform aspect detection task, *i.e.* to classify text into a predefined set of aspects. We propose three systems: seeds expansion, guided LDA-based seeds expansion and Multi-JASEN. The fist system *seeds expansion* is used to expand predefined seed words set for each aspect. This method has been called ”query expansion approach” in the literature and has been used in several information retrieval studies [6]. With this expanded list of seeds, we search to detect the aspects in each review.
The second system *guided LDA-based seeds expansion* is based on two steps: (*1*) `Review segmentation step` : reviews are split into segments such that each segment covers a single aspect. For example, the review *”Good reception but the price is expensive”* must be segmented into the two segments : *”Good reception”* and *”but the price is expensive”*; (*2*) `Clustering step` : opinion segments referring to the same aspect of the service with different words and phrases should be grouped under the same aspect category.
The third system *Multi-JASEN* represents an improvement version of state-of-the-art tool for English aspect-based sentiment analysis task.
The general scheme of the three proposed systems is shown in Fig 1. The seed expansion, segmentation reviews, grouping and other system details are explained in the following sections. The evaluation on SemEval-2016 French benchmark allows a comparison between the three proposed systems in order to identify the best one for aspect detection sub-task.

Fig. 1. Architectures of the three proposed systems



3.1. Seeds expansion system

Traditional expansion systems can suffer from introducing irrelevant information during query expansion. To address this shortcoming, inspired by recent advances in the application of contextualized models such as CamemBERT[1] to the information retrieval task. Our system, called, Camem4AE, proposes a new seeds expansion method that leverages the strength of the CamemBERT contextualized word embeddings model to select relevant terms for expansion. The dataset in question contains restaurant domain aspect terms belonging to five aspect categories: *food, ambience, price, service* and *restaurant*. Firstly, a set of five common aspect seeds from each of the five aspect categories is manually selected to form a set of 25 seeds words. To manually select the five aspect seeds, we used CamemBERT

[1] CamemBERT is a ”version” of BERT pretrained on a French textual dataset.

embeddings to retrieve the 10 most similar terms for each aspect category, and then manually selected 5 seeds from them. Secondly, an automated procedure is used to extend the list of predefined seeds using CamemBERT embeddings retrained on restaurant training set. Specifically, each predefined seed word was automatically expanded by 5 most similar terms to build a list of 30 seeds for each aspect. Finally, we perform a search with the expanded seeds list corresponding to each aspect in each review. Each review is processed word by word, if a word exists in a list of aspect seeds, the sentence will be labeled by this aspect, and if more than one aspect is retrieved for a single word, the semantic similarity (*SM*) between the retrieved aspects and the word is computed using cosine distance of the resulting word vectors. The similarity *SM* is the value between this word and the expanded seeds list for each aspect, and is calculated using the following equation :

$$SM(w, a) = mean_{v \in a}\ Cos\_Sim(w, v) \tag{1}$$

Where w is a word in the review, v is any of the seed words chosen to define an aspect a, and Cos_Sim stands for the cosine distance between two word vectors.

3.2. Guided LDA-based seeds expansion system

In this section, CamemLDA4AE, a new system for extracting aspects from user feedback, is proposed. CamemLDA4AE is based on CamemBERT embeddings and the Guided-LDA thematic modeling algorithm. Guided-LDA was proposed by [11] and since then it has been used to discover latent topics in documents. The proposed CamemLDA4AE method is weakly supervised and requires no annotated training data.

The proposed method consists of two phases: (1) the review segmentation phase consists in splitting reviews into segments such that each segment addresses a unique aspect. For example, consider the following review sample : *"Good reception but the price is expensive"*. Here the user talks about the *"reception"* and the *"price"* aspects of the assurance. Therefore, the segmentation phase is necessary to help recover all possible aspects in the review. The second phase is grouping. Reviews segments indicating the same aspect of the product with different words and phrases should be grouped into the same aspect category.

a) Reviews Segmentation : In the collection of user reviews, aspect and seeds have high frequency. However, our observations on the dataset and on the French user reviews revealed that aspect words are mainly nouns (price, reception, service, *etc.*). In this step, all extracted frequent nouns are considered as aspects.

After discovering the list of the frequent words, the next step is to segment the reviews using the association rules between words and text similarity method. The main idea of using association rules is to determine if there is a correlation between terms, in order to distinguish between terms that represent different aspects and those that speak of the same aspect.

In the literature, an association rule [6], *i.e.*, between terms, is defined as an implication of the form $R : T_1 \Rightarrow T_2$, where T_1 and T_2 are subsets of τ, where $\tau := t_1...t_n$ is a finite set of n distinct terms in the collection and $T_1 \cap T_2 = \theta$. The termsets T_1 and T_2 are, respectively, called the *premise* and the *conclusion* of R. The rule R is said to be based on the termset T equal to $T_1 \cup T_2$. The support *Supp* of a rule $R : T_1 \Rightarrow T_2$ is then defined as:

$$Supp(R) = Supp(T_1 \cup T_2) \tag{2}$$

while its *confidence* is computed as:

$$Conf(R) = \frac{Supp(R)}{Supp(T1)} \tag{3}$$

In association rules, a high confidence value between *T1* and *T2* means a strong association between them and a low confidence value between them means a weak association.

The elements of the review are handled one by one consecutively. Each element is checked whether it is a *frequent noun*, *seed word* in the expanded seeds list or not:

- If it's a *seed word*, then:
  - If there is a second seed word in the sentence, we check the seed word list, if both words exist in the same expanded seeds list, then both words are related with the same aspect and therefore there is no segmentation.
  - If a second seedword appears in a review, and both seedwords appear in two different expanded seeds lists, then, they present two different aspects and the segmentation is done just before the position of the second seedword in the review.
- If it's a *frequent noun* and there is a second *frequent noun* in the review, then the confidence value and the degree of similarity between the two frequent nouns are checked. The final value of the similarity is calculated as follows :
$$Sim(fn_1, fn_2) = \alpha * Conf(fn_1, fn_2) + \beta * Cos\_Sim(v_{fn_1}, v_{fn_2}) \tag{4}$$
Where fn_1 and fn_2 are two frequent nouns, $Conf(fn_1, fn_2)$ is the confidence value between fn_1 and fn_2, $Cos\_Sim(v_{fn_1}, v_{fn_2}$ is the degree of similarity between fn_1 and fn_2 using CamemBERT embedings with cosine distance, and α, β are the weighting parameters (α and β were fixed at 0.3 and 0.7 respectively).
  - If the $Sim(fn_1, fn_2)$ value is greater than the threshold value then these two words are strongly associated with each other. This means that this two words are related words of the same aspect, so no segmentation is done at this point.
  - If the $Sim(fn_1, fn_2)$ value is below the threshold value, then there is a weak association between these words. This means that this two "frequent words" are about different aspects, so segmentation is done just before the second "frequent nouns" fn_2.

When making the segmentation, word relations like possessive construction, adjective-noun, adjective-verb, adverb-verb associations are considered and the segmentation point is shifted one or two words back if necessary.

b) Guided LDA for Aspect Term Extraction : The second step of the proposed approach for aspect term extraction is a Guided-LDA model enhanced by the extended set of seeds for the considered aspects. The review corpus that we used is the first input to the guided LDA model. The expanded list of aspect seeds (see section 3.1), which is representative of the aspect categories, is the second input and will guide the model to enhance the topic-word distributions and review-topic distributions. These two inputs are passed to the Guided-LDA model, as shown in Figure 1.
The parameter K, which involves the number of topics, is mapped to the number of aspect categories with respect to aspect term extraction. The values of the guided-LDA hyperparameters α, β and *seed-confidence* which is the seed word reliability factor are set for the guided-LDA algorithm. With all the required information at hand, the guided-LDA is trained and used to predict the distributions of candidate aspect terms among the aspect categories. The probability of each token in the review segments data belonging to each of the topic/aspect categories is extracted. Finally, the tokens with higher probability of belonging to each aspect category are considered as aspect terms.

3.3. Multi-JASEN system

Jasen is a weakly supervised tool proposed by [10] for aspect-based sentiment analysis based on joint (aspect, polarity) topic embedding. Jasen learns firstly (aspect, polarity) joint topic embeddings in the word embedding space, and then uses CNN to generalize discriminative information at word level by pre-training CNN with embedding-based predictions and self-training it on unlabeled data. Jasen takes as inputs word2vec embeddings of dimension 100 and outputs the aspect in the review and its polarity. Jasen represents the state of the art tool for aspect-based sentiment analysis. It is tested on SemEval-2016 English dataset and outperforms existing methods on different domain corpora (restaurant, laptop, *etc.*) However, Jasen tool supposes that only one aspect is mentioned in the review.

In this work, we propose *Multi-JASEN*[2], an extension of Jasen. The proposed tool deals with reviews containing many aspects. Multi-JASEN is obtained by adding a review segmentation module to initial Jasen tool. This module is specific to Jasen tool and different from that one used in guided LDA-based seeds expansion system. Thus, the proposed tool is an adaptation of Jasen to deal with multi-aspect reviews.

The review segmentation algorithm is specific to sentiment analysis task. It aims to split the review in segments so that each segment contains only one aspect. This algorithm uses a seeds list for different aspects[3]. Steps of the segmentation algorithm are the following: (a) compute the number of different aspects in the review by using the keyword list, (b) if the number is greater than 1 then (i) cut firstly according to punctuation marks, (ii) cut secondly according to coordinating conjunctions, (iii) cut finally whenever there is a new aspect.
In addition to the segmentation module, an aggregation module is also added to initial Jasen in order to aggregate results obtained with different segments of each review.

4. Experimental setups

4.1. Corpus

For evaluation, we use the french benchmark dataset in the restaurant domain in SemEval-2016 [24], where each review is annotated with aspect and polarity labels. The training set contains 3 433 different words with a total occurrences of 22 027. We apply the following preprocessing steps. Firstly, we eliminate stop-words using NLTK toolkit. However, we kept negation terms as they are relevant for sentiment analysis. Then, we apply morphological analysis (POS tagging, lemmatization and stemming). Statistics based on aspect number are reported in Table 1

Table 1. Statistics (number of aspects) on SemEval-2016 french restaurant dataset.

Aspects	**Training set**	**Test set**
Ambience	185	82
Drink	69	32
Food	658	275
Location	38	25
Service	443	162
Restaurant	563	174
Total	1525 (reviews) → 1956 (aspects)	583 (reviews) → 750 (aspects)

4.2. Embedding set

Pretrained word representations are crucial in natural language processing tasks, from non-contextual word embeddings (glove [22], word2vec [19], fasttext [2], *etc.*) to contextual ones (bert [5], elmo [23], *etc.*).
Aspect extraction is a complex task that requires domain-specific embeddings. The majority of previous work proves that using domain embeddings is important [10], even when the domain embedding corpus is small. Thus, we use both general embeddings and domain embeddings. We use contextual CamemBERT embeddings[4] of dimension 768 pretrained on wikipedia [17]. They are retrained on restaurant SemEval-2016 training set.
Before diving into the experimental results, we report here how we use the available prior knowledge (seeds list). Seeds lists are provided as input by human experts (see Table 2). They are expanded using CamemBERT embedding

[2] In *Multi-JASEN*, we kept the same hyperparameters of Jasen.
[3] The seeds list is obtained by extension of human-defined seeds of different aspects based on CamemBERT embeddings.
[4] We tested word2vec, fasttext and CamemBERT embeddings. We report, in this work, the best results obtained with CamemBERT embeddings.

retrained on our restaurant training set. We found that this model provides an accurate representation of the relationships between terms, and we used it to add new seeds that have the same context as the original available terms. Then, the new seed lists (obtained by the concatenation of the predefined and the expanded seeds) were used to detect aspects in the reviews. An example of the obtained seed lists is shown in Table 2.

Table 2. A few of the expanded seeds obtained by CamemBERT embeddings.

Aspect	**Predefined seeds**	**Expanded seeds with CamemBERT**
Ambience	ambiance, décoration, sympa, soirée, environnement /ambiance,decoration ,nice ,evening,environment /	décor, terrasse, agréable, ensoleillée ,détestable /decor, terrace ,pleasant ,sunny,detestable /
Drink	boisson,café, eau, vin, bar /drinks, coffee, water, wine, bar/	thé, citron, cocktail, alcool,bière / tea, lemon, cocktail, alcohol, beer/
Food	nourriture, pizza, frite, plat, repas /food, pizza, fries, plate, meal/	menu, viande, salade, pain, burger /menu, meat, salad, bread, burger/
Location	rue, carte, place, quartier, ville /street, map, place, district, city/	déplacement, emplacement, coin, séjour /displacement, location, corner, stay/
Service	service, compétence, serveur, réception, chef /service, competence, waiter, reception, chef/	accueil, conseille, maître, gentillesse, patron /reception, advises, master, kindness, boss/
Restaurant	restauration, propreté, vieillissant, restaurant, salle /catering, cleanliness, aging, restaurant, room/	qualité, originalité /quality, originality/

4.3. Existing systems

We compare our systems against those reported in [26], [29], [3] and [21]. They are based on supervised and unsupervised learning. We consider these existing systems as baselines.

Supervised Baselines :

- XRCE [3] : is a system based on ensemble modelling combined with rich linguistic features including lexical semantic information and syntactico-semantic dependencies to address aspect based category.
- INSIGHT [26] : is a deep learning-based approach to aspect-based sentiment analysis, which employs a convolutional neural network for aspect extraction.
- UFAL [29] : is a neural networks-based system that automatically discover linguistic patterns in data, reducing the need for language-specific tools and feature engineering.

Weakly-supervised Baselines :

- W2VLDA [21] : is a weakly-supervised system for multilingual and multidomain ABSA, that works leveraging large quantities of unlabelled textual data and an initial configuration consisting of a minimal set of seed words.

These existing systems have been evaluated on restaurant benchmark french dataset in SemEval-2016 evaluation campaign.

5. Results and discussion

We evaluate our 3 systems for aspect extraction task with SemEval-2016 Task-5 French restaurant reviews. Moreover, we compare our weakly-supervised systems to some of the supervised baselines in order to demonstrate the efficiency of our weakly-supervised systems, in terms of performance, time and human effort.

5.1. Results

As mentioned in section 3, three systems for aspect extraction have been proposed : (*1*) Camem4AE which makes aspect detection using only the list of seeds added by CamemBERT, (*2*) CamemLDA4AE which uses associative rules for segmentation, and Guided LDA for classification, and ; (*3*) Multi-JASEN which is an adaptation of state-of-the-art Jasen tool for aspect-based sentiment analysis on English reviews in order to extract many aspects.
The proposed systems are evaluated on the Restaurant domain of SemEval-2016 french datasets and the results have been reported in tableau 3 in terms of precision, recall and F1 score. The F1 score reported by the best system *Camem4AE* is almost 60.65% and is around 7 points higher when compared to CamemLDA4AE system. This can be explained by the fact that the added seeds by CamemBERT are chosen based on the pertinent score , *i.e.,* using the strongest correlations inter-terms and leading to generate precise seeds.

Table 3. Results of the proposed methods for Aspect term extraction on SemEval-2016 french restaurant dataset.

Method	**Precision** (%)	**Recall** (%)	**F1** (%)
CamemLDA4AE	50.62	57.61	53.88
Camem4AE	**59.92**	**61.4**	**60.65**
Multi-JASEN	57.4	56.2	56.79

The evaluation results for each aspect of the product were also calculated. The extraction results of the proposed method for each aspect, including the ”general” aspect, are presented in Table 4. The level of success for each product aspect is not the same. This is related to the complexity of the statements and the characteristics of the coexistence patterns in the aspect statements. For example, the precision and recall of ”service” aspect is high because the users rarely use implicit expressions when talking about service aspect. They often directly use the word service. ”Service” and ”Food” are the most frequent words in the reviews by the way.
On the other hand, the precision in ”location” aspect is relatively low. There are two reasons for this: First, users often use very different words and statements when talking about this aspect. The second reason is that the frequency of this aspect in the comments is quite low. Finally and for the aspect ”restaurant”, its precision is low in spite of its high frequency in the train dataset. This could be explained by the following facts: restaurant is a general aspect, it implicitly brings together the other aspects (food, ambience, drink, location and service). In other words, the laters reflect ”restaurant” aspect. It could not be considered an aspect in its own right.
The sentence segmentation task and Guided-LDA both depend on the statistical distributions (co-occurrence patterns) of words. When the frequency of an aspect is relatively low, the statistical distributions for that aspect become relatively inappropriate.

Table 4. Success level of the best system for each aspect.

Aspect	**# occurrences**	**Precision** (%)	**Recall** (%)	**F1** (%)
Ambience	185	66	47	55
Drink	69	66	34	45
Food	658	69	83	75
Location	38	56	17	26
Restaurant	563	34	42	38
Service	443	72	84	78

5.2. Comparison of our best system with existing ones

We evaluate our best system *Camem4AE* in a topic classification setting using the restaurant reviews dataset from [24]. This dataset contains few thousand reviews from restaurants, classified into several categories : *ambience, food, service, etc.* We compare *Camem4AE* against the results reported in [26], [29] and [3]. A comparison of the best system's performance with the baselines considered is presented in Table 5.2. The best results are marked in bold. The best system reported the best F1 score of 60.65% compared to the weakly-supervised one [21]. The latter has been tested on only 3 aspects (*Food, Service and Ambience*) of the restaurant benchmark. And when we tested our best system on these three aspects, the reported F-measure is is approximately 71%. Thus, the best system *CamemLDA4AE* is almost at the same level as the supervised one which has the best performance.

Table 5. Comparison of our best system performance with supervised and weakly-supervised baselines

Dataset	**Methodology**	**F1**
XRCE [3]	supervised	**61.20**
INSIGHT [26]	supervised	53.59
UFAL [29]	supervised	49.93
W2VLDA [21]	weakly-supervised	58.6
Our best system	weakly-supervised	60.65

In making a comparison of the results, there is an interesting point to mention: In SemEval-2016, with the training data provided, all submissions implement supervised machine learning approaches. However, our systems are weakly-supervised. Methods based on supervised machine learning use training datasets, which gives them a great advantage. In their studies, [13] and [34] also show that supervised machine learning methods are more effective than unsupervised methods. However, supervised machine learning methods have a major drawback that they need annotated training data, which is not always possible. Keeping this fact in mind, when we compare the F1 score value of our best system with others, we believe that our weakly-supervised method performs quite well in the French SemEval-2016 dataset.

6. Conclusion and future work

In this work, we proposed 3 weakly-supervised systems for aspect extraction task : Camem4AE, CamemLDA4AE and Multi-JASEN. They use contextualized CamemBERT embeddings and a combination of associative rules, Guided LDA and state-of-the-art Jasen tool. They were evaluated on restaurant french benchmark in SemEval-2016 evaluation campaign.

Our best weakly-supervised system is *Camem4AE*. It does not need any annotated data and its weakly-supervised character is conveyed via a short human-predefined seeds list. Besides, *Camem4AE* performs as well as the best existing supervised ones. Thus, our best system *Camem4AE* is efficient in terms of performance, time and human effort.

Some improvements can be made to the proposed systems. In particular, the first step of the method, where the review segmentation is performed, can be improved or different algorithms can be used. The proper segmentation of the reviews is very crucial as it is decisive for the overall success of the proposed systems.

Another perspective consists on studying dataset characteristics, seeds list size and the relation between them.

To go further, another perspective consists of determining polarities for extracted aspects in order to resolve the whole aspect-based sentiment analysis task.

References

[1] Birjali, M., Kasri, M., Beni-Hssane, A., 2021. A comprehensive survey on sentiment analysis: Approaches, challenges and trends. Knowledge-Based Systems 226, 107134.

[2] Bojanowski, P., Grave, E., Joulin, A., Mikolov, T., 2017. Enriching word vectors with subword information. Transactions of the association for computational linguistics 5, 135–146.

[3] Brun, C., Perez, J., Roux, C., 2016. XRCE at semeval-2016 task 5: Feedbacked ensemble modeling on syntactico-semantic knowledge for aspect based sentiment analysis. SemEval@NAACL-HLT , 277–281.

[4] Chen, T., Xu, R., He, Y., Wang, X., 2017. Improving sentiment analysis via sentence type classification using bilstm-crf and cnn. Expert Systems with Applications 72, 221–230.

[5] Devlin, J., Chang, M., Lee, K., Toutanova, K., 2019. BERT: pre-training of deep bidirectional transformers for language understanding, in: Burstein, J., Doran, C., Solorio, T. (Eds.), NAACL-HLT, pp. 4171–4186.

[6] Ettaleb, M., Latiri, C., Bellot, P., 2018. A combination of reduction and expansion approaches to deal with long natural language queries, in: Knowledge-Based and Intelligent Information & Engineering Systems: Proceedings of the 22nd International Conference KES-2018, Elsevier. pp. 768–777.
[7] García-Pablos, A., Cuadros, M., Rigau, G., 2018. W2vlda: almost unsupervised system for aspect based sentiment analysis. Expert Systems with Applications 91, 127–137.
[8] He, R., Lee, W.S., Ng, H.T., Dahlmeier, D., 2017. An unsupervised neural attention model for aspect extraction, in: Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), pp. 388–397.
[9] Hu, M., Liu, B., 2004. Mining and summarizing customer reviews, in: SIGKDD, pp. 168–177.
[10] Huang, J., Meng, Y., Guo, F., Ji, H., Han, J., 2020. Weakly-supervised aspect-based sentiment analysis via joint aspect-sentiment topic embedding. arXiv preprint arXiv:2010.06705 .
[11] Jagarlamudi, J., III, H.D., Udupa, R., 2012. Incorporating lexical priors into topic models, in: EACL, pp. 204–213.
[12] Karaoğlan, K.M., Fındık, O., 2022. Extended rule-based opinion target extraction with a novel text pre-processing method and ensemble learning. Applied Soft Computing 118, 108524.
[13] Kennedy, A., Inkpen, D., 2006. Sentiment classification of movie reviews using contextual valence shifters. Comput. Intell. 22, 110–125.
[14] Kim, S., Zhang, J., Chen, Z., Oh, A.H., Liu, S., 2013. A hierarchical aspect-sentiment model for online reviews, in: desJardins, M., Littman, M.L. (Eds.), Proceedings of the Twenty-Seventh AAAI Conference on Artificial Intelligence, July 14-18, 2013, Bellevue, Washington, USA, AAAI Press.
[15] Li, Z., Zhang, M., Ma, S., Zhou, B., Sun, Y., 2009. Automatic extraction for product feature words from comments on the web, in: Asia Information Retrieval Symposium AIRS, pp. 112–123.
[16] Liu, P., Joty, S., Meng, H., 2015. Fine-grained opinion mining with recurrent neural networks and word embeddings, in: Proceedings of the 2015 conference on empirical methods in natural language processing, pp. 1433–1443.
[17] Martin, L., Muller, B., Ortiz Suárez, P.J., Dupont, Y., Romary, L., de la Clergerie, É., Seddah, D., Sagot, B., 2020. CamemBERT: a tasty French language model, in: ACL, pp. 7203–7219.
[18] Meng, Y., Zhang, Y., Huang, J., Xiong, C., Ji, H., Zhang, C., Han, J., 2020. Text classification using label names only: A language model self-training approach. arXiv preprint arXiv:2010.07245 .
[19] Mikolov, T., Sutskever, I., Chen, K., Corrado, G.S., Dean, J., 2013. Distributed representations of words and phrases and their compositionality. Advances in neural information processing systems 26.
[20] Ozyurt, B., Akcayol, M.A., 2021. A new topic modeling based approach for aspect extraction in aspect based sentiment analysis: Ss-lda. Expert Systems with Applications 168, 114231.
[21] Pablos, A.G., Cuadros, M., Rigau, G., 2018. W2VLDA: almost unsupervised system for aspect based sentiment analysis. Expert Syst. Appl. 91, 127–137.
[22] Pennington, J., Socher, R., Manning, C.D., 2014. Glove: Global vectors for word representation, in: Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP), pp. 1532–1543.
[23] Peters, M.E., Neumann, M., Iyyer, M., Gardner, M., Clark, C., Lee, K., Zettlemoyer, L., 2018. Deep contextualized word representations, in: NAACL, pp. 2227–2237.
[24] Pontiki, M., Galanis, D., Papageorgiou, H., Androutsopoulos, I., Manandhar, S., AL-Smadi, M., Al-Ayyoub, M., Zhao, Y., Qin, B., De Clercq, O., Hoste, V., Apidianaki, M., Tannier, X., Loukachevitch, N., Kotelnikov, E., Bel, N., Jiménez-Zafra, S.M., Eryiğit, G., 2016. SemEval-2016 task 5: Aspect based sentiment analysis, in: Proceedings of the 10th International Workshop on Semantic Evaluation (SemEval-2016).
[25] Rana, T.A., Cheah, Y., 2017. A two-fold rule-based model for aspect extraction. Expert Syst. Appl. 89, 273–285.
[26] Ruder, S., Ghaffari, P., Breslin, J.G., 2016. INSIGHT-1 at semeval-2016 task 5: Deep learning for multilingual aspect-based sentiment analysis, in: SemEval@NAACL-HLT, pp. 330–336.
[27] Ruidan, H., 2020. Exploiting document knowledge for aspect-level sentiment classification. US Patent 10,726,207.
[28] Shu, L., Xu, H., Liu, B., 2017. Lifelong learning crf for supervised aspect extraction. arXiv preprint arXiv:1705.00251 .
[29] Tamchyna, A., Veselovská, K., 2016. UFAL at semeval-2016 task 5: Recurrent neural networks for sentence classification, in: SemEval@NAACL-HLT 2016, pp. 367–371.
[30] Tulkens, S., van Cranenburgh, A., 2020. Embarrassingly simple unsupervised aspect extraction. arXiv preprint arXiv:2004.13580 .
[31] Wang, J., Xu, B., Zu, Y., 2021. Deep learning for aspect-based sentiment analysis, in: 2021 International Conference on Machine Learning and Intelligent Systems Engineering (MLISE), IEEE. pp. 267–271.
[32] Xu, H., Liu, B., Shu, L., Yu, P.S., 2018. Double embeddings and cnn-based sequence labeling for aspect extraction. arXiv preprint arXiv:1805.04601 .
[33] Zad, S., Heidari, M., Jones, J.H., Uzuner, O., 2021. A survey on concept-level sentiment analysis techniques of textual data, in: 2021 IEEE World AI IoT Congress (AIIoT), IEEE. pp. 0285–0291.
[34] Zhang, H., Gan, W., Jiang, B., 2014. Machine learning and lexicon based methods for sentiment classification: A survey, in: 2014 11th Web Information System and Application Conference, pp. 262–265.
[35] Zhuang, L., Jing, F., Zhu, X.Y., 2006. Movie review mining and summarization, in: Proceedings of the 15th ACM international conference on Information and knowledge management, pp. 43–50.