

Performance Evaluation of an Ontology-Based Arabic Sentiment Analysis

Salima Behdenna, Fatiha Barigou, Ghalem Belalem

Abstract—Due to the quick increase in the volume of Arabic opinions posted on various social media, Arabic sentiment analysis has become one of the most important areas of research. Compared to English, there is very little works on Arabic sentiment analysis, in particular aspect-based sentiment analysis (ABSA). In ABSA, aspect extraction is the most important task. In this paper, we propose a semantic ABSA approach for standard Arabic reviews to extract explicit aspect terms and identify the polarity of the extracted aspects. The proposed approach was evaluated using HAAD datasets. Experiments showed that the proposed approach achieved a good level of performance compared with baseline results. The F-measure was improved by 19% for the aspect term extraction tasks and 55% aspect term polarity task.

Keywords—Sentiment analysis, opinion mining, Arabic, aspect level, opinion, polarity.

I. INTRODUCTION

SENTIMENT analysis is an important research area which has aroused a great interest. It aims to analyze people's opinions, sentiments, evaluation and emotions towards entities and their attributes [1]. Sentiment analysis can be performed mainly at three levels of analysis: document level, sentence level and aspect level. However, sentiment analysis at document level or even at sentence level is often insufficient because it does not identify the opinion targets or does not attribute opinions to these targets. To obtain a finer analysis and extract finer information from the text, it is therefore necessary to move to aspect level. The goal of this level is to discover and summarize people's opinions expressed on entities and/or their aspects [2].

Due to the fast increase in the volume of Arabic opinions posted on various social media, Arabic sentiment analysis has become one of the most important area of research. As shown in Fig. 1, the first work was done on the classification of opinions at document level in English (since 2002), while for the Arabic language the first work only started in 2008 and was considerably delayed. Few studies deal with Arabic sentiment analysis (e.g. [3]), in particular ABSA. On the one hand, ABSA research is mainly based on the lexical approach; on the other hand, ABSA tasks are difficult due to the variety of semantics of opinions expressed, as well as the polarity of opinion-bearing words used to describe them. It is therefore necessary to consider the semantics of the opinions expressed on the aspects in order to improve Arabic ABSA. In this paper, we propose a semantic ABSA approach for Arabic reviews.

S. Behdenna, F. Barigou, and G. Belalem are with Department of Computer Science, Faculty of Sciences, Université Oran1 Ahmed Ben Bella, PB 1524 El

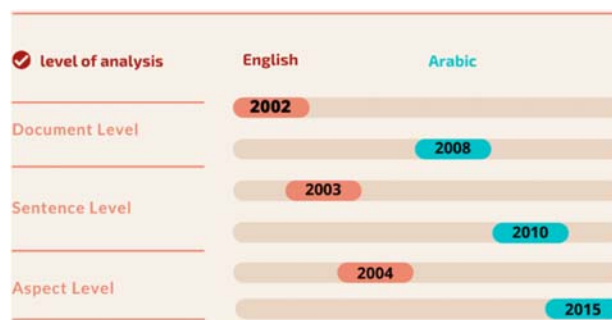


Fig. 1 Research progress in the different levels of sentiment analysis for English and Arabic

The tasks, which relate to SemEval2014-ABSA [5], include four tasks: aspect term extraction (T1), aspect-terms polarity identification (T2), aspect category identification (T3), and aspect category polarity (T4). Our work considers only the first two ABSA tasks T1 and T2. To this end, we propose to employ Description Logics (DLs) to describe the ontology. To our knowledge, it is the first attempt to use this technology to address these two tasks.

The rest of this paper is organized as following. In Section II, we present Related Work. Section II details the proposed approach. Section IV presents and discusses the results obtained with the two tasks. Finally, Section V concludes the paper.

II. RELATED WORK

Behdenna et al. [3] evaluated a method for extracting opinion targets from Arabic tweets by modeling the problem of extracting opinion targets in a machine learning classification task and by combining a number of resources available for the Arabic language.

Al-Dabet et al. [4] proposed a deep learning model to address the task of extracting opinion target. The proposed model is composed of bidirectional short-term memory (Bi-LSTM) as encoder, long-term memory (LSTM) as decoder with an attention mechanism and conditional random fields (CRF). This model is designed to extract opinion targets for the Arabic language. The performance of the proposed model is assessed using the SemEval-2016 annotated data set for the hotel domain. The experimental results show that the proposed model achieves an F-measure of 72.83%.

Areed et al. [5] proposed a corpus of Arabic ABSA for government mobile app reviews and proposed a combined

M'Naouer, Oran, Algeria (e-mail: behdennasalima@gmail.com, fatbarigou@gmail.com, ghalemldz@gmail.com).

approach to extract explicit and implicit aspects and classify sentiments by adopting approaches based on lexicon with rule-based models.

The Human Annotated Arabic Dataset (HAAD) developed by Al-Smadi et al. [6] was the first publicly available corpus of Arabic ABSA developed due to the limited availability of appropriate corpora. It contains 1513 book reviews selected from the LABR corpus [11] and manually annotated. The HAAD corpus was annotated based on the guidelines of SemEval 2014: Task 4.

Al-Smadi et al. [7] applied deep learning on hotel reviews to extract the opinion targets and identify the polarity of opinions. The authors used two implementations of long-term memory neural networks (LSTMs). The first is a character-level bidirectional LSTM with a conditional random field classifier (Bi-LSTM-CRF) for the extraction of opinion target expressions (OTE), and the second is an aspect-based LSTM for the aspect opinion polarity classification in which aspects-OTE are considered expressions of attention to support the identification of the polarity of opinions. The proposed approaches are assessed using a benchmark data set of Arab hotel reviews. The results show that both approaches outperform baseline research on both tasks with an improvement of 39% for the Aspect Extraction-OTE task and 6% for the Identify Opinion Polarity task.

Based on related work, it is found that most of the works for Arabic ABSA consider only the syntactic of the opinions expressed on the aspects. However, ABSA tasks are difficult due to the variety of semantics of opinions expressed. It is therefore necessary to consider the semantics of the opinions expressed on the aspects in order to improve Arabic ABSA.

III. PROPOSED APPROACH

The main goal of the proposed approach is to improve ABSA. More specifically, we are interested in the first two tasks defined in SemEval2014-ABSA; aspect term extraction (T1), aspect-terms polarity identification (T2).

ABSA requires domain knowledge to extract entities and their aspects, in order to correctly determine the polarity of each aspect. For this, we propose a semantic approach using a domain ontology to support the ABSA from reviews written in Arabic language.

This work is an extension of the published paper [8]. The old version focused exclusively on the design of the approach. The implementation is carried out at this stage. Development methodology of the proposed approach is carried out in two main stages; the first concerns the construction of the ontology, the second that will use this ontology will deal with the ABSA.

A. Ontology Development

The purpose of using ontology in ABSA is to identify entities, aspects and opinions. The case study for our proposed approach is the domain of books. We worked with this domain because we noticed that among the corpus that meet our needs and that are available, there was only that of book reviews; the

“HAAD” corpus [6]. On the other hand, the Arabic ontology of the book, to our knowledge, is not available. For this reason, we have chosen the universe of literature (books, writers) as an example of application to build a domain ontology. This can then be used to identify aspects and their polarities in book reviews written in Arabic.

The automatic and semi-automatic generation of ontology from a domain corpus is a research challenge in the Semantic Web. Many tools have been created for this task. However, to our knowledge, there are no tools for building ontologies in Arabic. For this reason, we decided to manually create this ontology.

First, we have manually collected a set of key concepts in the book domain from the HAAD corpus. Then, we classified the terms collected in classes or concepts, instances, and relations between objects. We have, for example the terms, "كتاب", "قصة" ... as classes. In addition to the hierarchical relation “is a” which links the different subclasses to their parent classes, we have defined other semantic relations like (has, expressed by...). Fig. 2 shows a snapshot of the ontology.

For the formal representation of this ontology, we have chosen to use description logics formalism (DLs). The motivation for this choice is twofold: first, the domain entities and their aspects are described by concepts and roles [9]. Second, DL is based on formal and logical semantics [10].

- TBox allows introducing the terminology, i.e., a set of concept axioms and role axioms. There are 9 top level concepts which are associated to book domain: (كتاب) (Book), (ميزة) (Feature), (قصة) (story), (رواية) (novel), (شخص) (Person), (كاتب) (writer), (كاتب-ميزة) (Feature Writer), (قارئ) (reader), (رأي مصطلح) (Opinion word).
- ABox contains assertions about individuals, specifying their class.

To validate and check consistency of the constructed ontology we use Pellet Reasoner.

B. Aspect-Based Sentiment Analysis

1) Preprocessing

Several treatments must be taken and multiple NLP techniques must be utilized, including sentence splitting, tokenization to tokenize every sentence into words, normalization, Part of Speech (POS) tagging; the words are tagged using Stanford Arabic part of speech tagger and removing the stop words.

2) Explicit Aspect-Term Extraction (T1)

The proposed approach will be applied only to aspect terms that are explicitly expressed. First, only words whose grammatical category matches nouns (e.g., NN: singular noun, NNP: proper singular noun) are considered candidates to be aspect terms. Then, to determine this, they are mapped to the concepts included in the T-Box of our book ontology. An explicit aspect term in the review corresponds to a concept present in the T-Box.

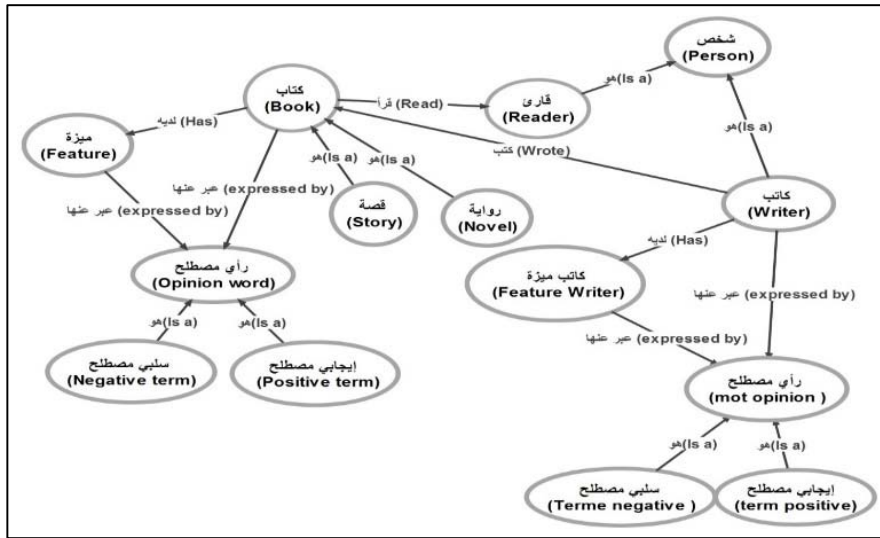


Fig. 2 Modeling of book domain knowledge

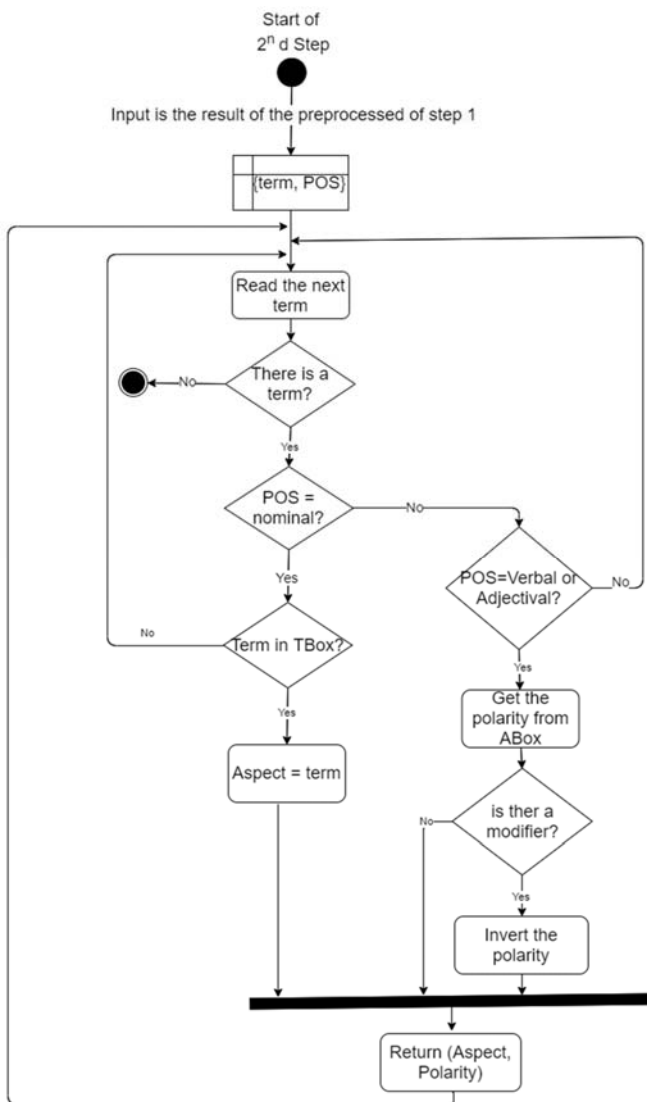


Fig. 3 The whole task for explicit aspect-term extraction and aspect-term polarity identification process

3) Aspect-Term Polarity Identification (T2)

To identify the polarity of the aspect term, the polarity modifiers combined with ABox are used to determine the polarity of the opinion words. When one of the polarity modifiers appeared before the opinion word, it would change the polarity of the opinion terms from positive to negative, and vice versa. For example, in the sentence "هذه الرواية لم تعجبني" "This novel did not interest me". The denier "لم" ("did not") reverses the polarity of the opinion word "تعجبني" ("interest me").

Using ABox, the proposed approach checks whether each adjective or verb is an opinion word, then identifies the existence of the polarity modifiers before retrieving and identifying its polarity. If the polarity modifiers exist, the proposed approach reverses the polarity of the current opinion word. If not, our approach derives its polarity from ABox. These two steps of ABSA are presented in Fig. 3.

IV. EXPERIMENT

To evaluate the proposed approach, we used HAAD, the first ABSA dataset for Arabic language [6]. It consists of 1513 books review in Arabic selected from the LABR dataset and annotated by humans with aspect terms and their polarities [6]. It is provided in three files (training file, test-gold file, and the test file), and it is equipped with a tool for common evaluation technique. This tool yields us the possibility to compare our results with those of the Test-gold file by computing the same measures used in the baseline approach. The details of this baseline approach can be found in [6]. We use this baseline approach in the comparative study.

A. Evaluation Metrics

In order to evaluate the extraction of the aspect terms (T1), the F-measure (F1) is used. The F-measure is defined as the harmonic mean of the recall (R) and the precision (P). Precision and recall are two standard evaluation measures commonly used to evaluate and measure the performance of approaches.

They are defined as:

$$F1 = \frac{2 \cdot P \cdot R}{P + R} \quad (1)$$

$$P = \frac{TP}{TP + FP} \quad (2)$$

$$R = \frac{TP}{TP + FN} \quad (3)$$

where TP: True Positive, TN: True Negative, FP: False Positive, FN: False Negative.

In order to evaluate aspect term polarity (T2), the accuracy (Acc) of the proposed approach is computed and compared with the baseline approach:

$$Acc = \frac{TN + TP}{TN + TP + FP + FN} \quad (4)$$

V. RESULTS AND DISCUSSION

A. Results

In this section, we present the experimental results obtained. The results are very promising and show the effectiveness of the proposed approach compared to the basic approach.

1) Results of the First Task: T1

Fig. 4 shows the results of the first task: extraction of the aspect term in terms of precision, recall and F-measure.

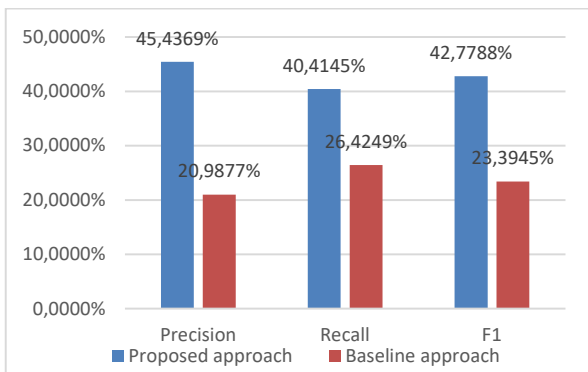


Fig. 4 Aspect term extraction performance

We studied the performance of extracting the explicit aspect term. As shown in Fig. 4 the results indicate that:

- The best results were obtained with the proposed approach.
- Significant performance improvement is observed in terms of 19% F-measure, 24% accuracy and 14% recall.
- Our approach failed to extract all appearance terms, due to annotation errors made by annotators regarding appearance terms in the HAAD test dataset as in the review: "لم يعجبني أبدا" (I never liked it) the word "يعجبني" is incorrectly annotated as an aspect term when it is an opinion word. While the term aspect is expressed implicitly.
- Further, the major limitation of our approach is due to the implicit aspect terms present in objective sentences and metaphorical sentences.

2) Results of the Second Task: T2

Fig. 5 shows the results obtained by our semantic approach for the second task: identification of the polarity of the aspect term, in terms of precision, recall and F-measure, for the three polarity classes: positive, negative and neutral. By looking at these results, we can see that the proposed approach gives us a precision of 98%, a recall of 84% and an F-measure of 90% for the negative polarity class. For the neutral polarity class, we arrive at a precision of 100%, a recall of 66% and an F-measure of 80%. While for the positive polarity class, we have a precision of 57%, a recall of 93% and an F-measure of 70%.

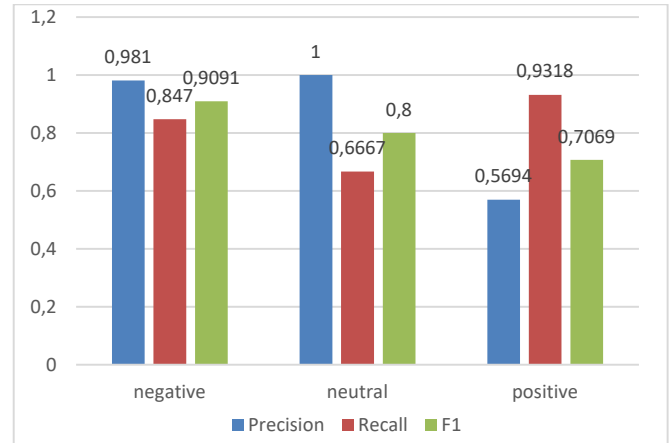


Fig. 5 Experimental results of identifying the polarity of the aspect term

The analysis of the results allowed us to better understand the reasons for the decline in the performance of some classes compared to others, in particular the precision of the positive class. This low accuracy rate is mainly due to the ambiguity of the polarity of certain opinion words. The system does not recognize the polarity of contextual words i.e., words that change their polarities from one domain to another because it does not have sufficient linguistic rules to recognize them; as in the review: "قصة قصيرة" "A short story". Overall, the system makes a good identification of the polarity of the aspect term.

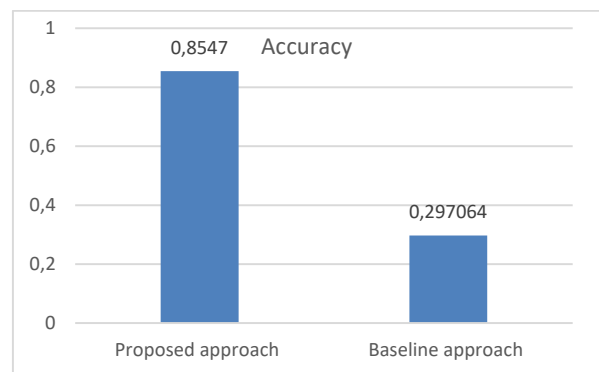


Fig. 6 Accuracy of the identification of the polarity of the aspect term

To better see the impact of using ABox, we compared the performance of the second task of our approach with that of the

baseline approach. Fig. 6 shows the accuracy of the second task for the two approaches.

We have studied the performance of identifying the polarity of the aspect term. As shown in Fig. 6, the results indicate that:

- In terms of accuracy, the results are very interesting, an improvement of 55% is observed.
- Our approach has failed to identify the ambiguous polarity of opinion words.
- In addition, the major limitation of our approach is due to the opinions expressed in the metaphorical sentences; as in the review: "كتاب ملحد" "An atheist book" is difficult to identify the polarity of opinion.
- These results tend to confirm our hypothesis to use the ABox combined with the polarity modifiers to improve the identification of the polarity of opinion words.

B. Discussion

For the two tasks: Extraction of the aspect term and identification of the polarity of the aspect terms, the best results were obtained with the proposed approach. The results are very promising and show the effectiveness of the proposed approach compared to the baseline approach.

Fig. 7 shows the precision and F-measure obtained by each approach and for each task.

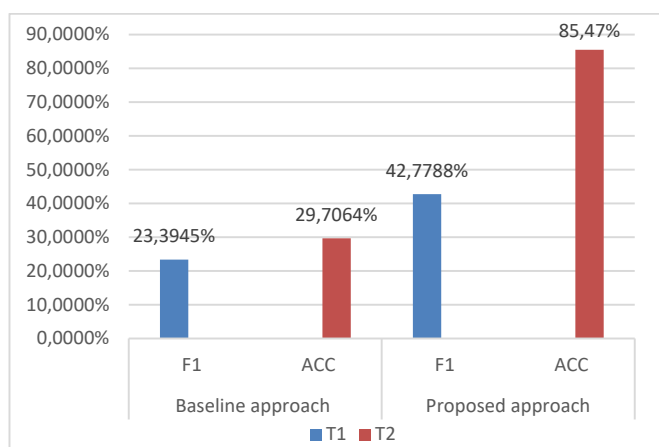


Fig. 7 Comparison with the baseline approach

The experimental study allowed us to make several observations:

- The proposed approach achieves the best performance for both tasks.
- The proposed approach surpasses the baseline approach.
- Regarding task T1, a significant improvement in performance is observed in terms of F-measure of 19%, precision of 24% and recall of 14%.
- Concerning the T2 task, the precision is considerably improved by 55%.

We also see two results:

- The experimental results are very promising and show the efficiency of the proposed semantic approach compared to the baseline approach.
- The major limitations of our approach are due to the implicit aspect terms present in the objective sentences and

the metaphorical sentences and to the opinions expressed by the metaphorical sentences or by ambiguous opinion words.

VI. CONCLUSION

This work focuses on ABSA for the Arabic language, considering, particularly two ABSA tasks: Aspect Term Extraction and Aspect Term Polarity. We proposed a semantic ABSA approach for Arabic reviews based on Description logics (DLs) for the identification of opinion aspects and their polarity. The proposed approach was evaluated using HAAD datasets. Experiments showed that the proposed approach outperforms the baseline approach proposed by [6] with an overall enhancement around 19% for T1 in terms of F-measure and 55% for T2 in terms of accuracy.

For future work, we will improve the proposed approach by studying the implicit aspect term extraction, and detecting the polarity of ambiguous words. For this, we plan to use the benefits of reasoning services of DLs combined with the linguistic rules to deal with this kind of challenging expressions.

REFERENCES

- [1] B. Liu, "Sentiment analysis and opinion mining," *Synthesis lectures on human language technologies*, vol. 5, no. 1, pp. 1-167, 2012.
- [2] S. Behdenna, F. Barigou, and G. Belalem, "Sentiment Analysis at Document Level," in *Conf. International Conference on Smart Trends for Information Technology and Computer Communications*. Springer, Singapore, 2016, pp. 159-168.
- [3] S. Behdenna, F. Barigou, and G. Belalem, "Sentiment Analysis of Arabic Tweets: Opinion Target Extraction," *Digital Information Management J.*, vol. 16, no. 6, pp. 325, 2018.
- [4] S. Al-Dabet, S. Tedmori, and M. Al-Smadi, "Extracting Opinion Targets Using Attention-Based Neural Model," *SN Computer Science J.*, vol. 1, no 5, pp. 1-10, 2020.
- [5] S. Areed, O. Alqaryouti, B. Siyam, and K. Shaalan, "Aspect-based sentiment analysis for Arabic government reviews" In *Recent Advances in NLP: The Case of Arabic Language*, 2020, pp. 143-162.
- [6] M. Al-Smadi, O. Qawasmeh, B. Talafha, and M. Quwaider, "Human annotated arabic dataset of book reviews for aspect based sentiment analysis," in *Conf. 2015 3rd International Conference on Future Internet of Things and Cloud*. IEEE, pp. 726-730, 2015.
- [7] M. Al-Smadi, B. Talafha, M. Al-Ayyoub, and Y. Jararweh, "Using long short-term memory deep neural networks for aspect-based sentiment analysis of Arabic reviews," *International Journal of Machine Learning and Cybernetics.*, vol. 10, no 8, pp. 2163-2175, 2019.
- [8] S. Behdenna, F. Barigou, and G. Belalem, "Towards Semantic Aspect Based Sentiment Analysis for Arabic Reviews," *International Journal of Information Systems in the Service Sector (IJISSS) J.*, vol. 12, no. 4, pp. 1-13, 2020.
- [9] F. Baader, I. Horrocks, and U. Sattler, "Description logics as ontology languages for the semantic web," in *Conf. Mechanizing mathematical reasoning*, Springer, Berlin, 2005, pp. 228-248.
- [10] G. Shu, O. F. Rana, N. J. Avis, and C. Dingfang, "Ontology-based semantic matchmaking approach," *Advances in engineering software J.*, vol. 38, no. 1, pp.59-67, 2007.
- [11] M.Aly, A. Atiya, "Labr: A large scale arabic book reviews dataset. In Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics, vol. 2, pp. 494-498, 2013.

S. Behdenna is a PhD candidate in the Department of Computer Science in the Faculty of Sciences at the University of Oran 1, Ahmed Ben Bella in Algeria. She received her Magister degree in Computer Science in 2013 from the University of Sciences at the University of Oran 2, Mohamed Boudief. She supervised 2 Master in computer science. Her research interests are: Sentiment Analysis and Opinion mining, Text Mining, Data Mining, Machine Learning.

Dr. F. Barigou is a lecturer at Computer Science Department of Université Oran 1, Oran, Algeria. She does research in Text Data Mining, Big data and Artificial Intelligence. Her current projects are Arabic Sentiment Analysis, Cloud Computing and IOT in healthcare.

Prof. G. Belalem Graduated from Department of computer science, Faculty of exact and applied sciences, University of Oran1 Ahmed Ben Bella, Algeria, where he received PhD degree in computer science in 2007. His current research interests are distributed system; grid computing, cloud computing, replication, consistency, fault tolerance, resource management, economic models, energy consumption, Big data, IoT, mobile environment, images processing, Supply chain optimization, Decision support systems, High Performance Computing.