International Journal of Information, Control and Computer Sciences
ISSN: 2517-9942
Vol:15, No:1, 2021

Opinion Mining and Sentiment Analysis on DEFT

Najiba Ouled Omar, Azza Harbaoui, Henda Ben Ghezala

***Abstract*—**Current research practices sentiment analysis with a focus on social networks, DEfi Fouille de Texte (DEFT) (Text Mining Challenge) evaluation campaign focuses on opinion mining and sentiment analysis on social networks, especially social network Twitter. It aims to confront the systems produced by several teams from public and private research laboratories. DEFT offers participants the opportunity to work on regularly renewed themes and proposes to work on opinion mining in several editions. The purpose of this article is to scrutinize and analyze the works relating to opinions mining and sentiment analysis in the Twitter social network realized by DEFT. It examines the tasks proposed by the organizers of the challenge and the methods used by the participants.

***Keywords*—**Opinion mining, sentiment analysis, emotion, polarity, annotation, OSEE, figurative language, DEFT, Twitter, Tweet.

I. Introduction

WITH the emergence of the web, especially Web 2.0, the number of documents containing information that expresses opinions, thoughts, feelings, emotions, personal judgments, and judgments of evaluation became more significant. In addition, the number of works on Opinion Mining has increased which proves the importance of Opinion on the web [2], [9]. Therefore, opinion mining and sentiment analysis are two axes of the same field of research that is globally an emerging and expanding field, several evaluation campaigns have focused on this area on a global and French scale. This is the case with DEFT (Text Mining Challenge) created in 2005 which is an annual workshop Francophone evaluation in text mining.

The aim of this article is to scrutinize and analyze the DEFT editions concerning the opinion mining and sentiment analysis in the Twitter social network. It is used in order to examine the tasks which are proposed by the organizers of the challenge and the methods that are exploited by the participants.

The importance of Opinion Mining is present in several fields, but the biggest application of Opinion Mining and Sentiment Analysis remains in the world of business and politics which can help decision-making.

II. The Importance of Opinion Mining

Mining techniques continue to invade the forms and typologies of information that envelops human life, we have gone from text mining to datamining to graph-mining and web-mining to extract meaning, analyze it and be able to predict or even help in decision-making. We are now in sentiment mining and sentiment analysis.

Najiba Ouled Omar is with the National School of Computer Science, Tunisia (e-mail: ouledomarnajiba@gmail.com).

A. Opinion Mining Process

Fig. 1 shows the steps of opinion mining.

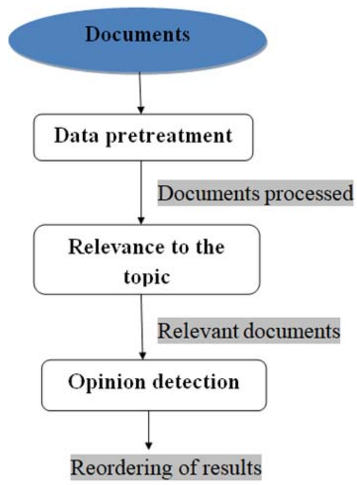

Fig. 1 Opinion Mining Process

- Data pretreatment: In this step, the texts are linguistically pretreated by removing empty words and words that do not provide any important information. Therefore, the lexical analysis removes words that have the same meaning. In this phase, grammatical labeling is carried out to determine the adverb, adjective, noun, and verb ...
- Relevance to the topic: This step allows studying the relevance of the texts to a given topic. The texts are classified, and generally, the first 1000 most relevant texts are removed and used for the next step.
- Opinion detection: Opinion detection uses several methods to reorder the relevant documents based on an opinion score.

B. Fields of Application of Opinion Mining

The importance of opinion mining and sentiment analysis runs through many areas, but the greatest application remains in business and politics:

- Marketing: Opinion mining allows any wanting company, the supplier of a product or a service to better understand what pleases and dislikes its customers by anticipating their needs and expectations in order to try to improve the product quality/service and increase profits. Thus, the customer can for his part give his opinion, to compare the products before purchasing them, not to read all the comments concerning a given product such that it is enough to see the positive percentage associated with this product and he can inspire feelings and opinions of other customers about the product which he is interested in and thus benefit from decision support.

- Politics: Political actors have also followed the challenge of opinion mining as before enacting a new law, politicians try to collect the opinions of Internet users on this law. In addition, it is very important to know the opinion of Internet users about a particular politician for such an election.

C. The DEFT Editions Focusing on Opinion Mining

The aim of the article is to provide an overview of the tasks of opinion mining during the different editions of the DEFT workshop.

*2014 Edition

Fraisse and Paroubek developed a guide to annotate opinions, feelings, and emotions which are of interest to French and German on statements of less than 140 characters. Their annotations are of two types, groups which consist of six types (OSEE (Expression of emotion, sentiment or opinion), source, target, modifier, negation, and recipient) and relationships which consist of five types (DIT, SUB, MOD, NEG, and RECEIVER) [6].

An expression of opinion of feeling or emotion consists of the span of text whose semantics express the OSEE; it can be carried on a verb, a noun, an adjective, an adverb or even a preposition.

The OSEE includes 18 fine semantic categories, these different fine categories are listed in three major classes and Table I shows the correspondence between fine semantic categories, type, and polarity [6].

- Intellective (opinion)
- Intellective-affective (feeling)
- Affective (emotion).

TABLE I
CORRESPONDENCE BETWEEN FINE SEMANTIC CATEGORIES, TYPE, AND POLARITY [6]

Fine semantic category	Generic class (type)	Polarity
Agreement	Opinion	+
Valorization		
Disagreement		-
Depreciation		
Satisfaction	Feeling	+
Dissatisfaction		-
Pleasure	Emotion	+
Positive surprise		
Love		
Appeasement		
Displeasure		-
Negative surprise		
Contempt		
Boredom		
Anger		
Fear		
Sadness		
Disturbance		

*2015 Edition

This edition offered participants the opportunity to identify and analyze opinion - bearing expressions, analyzing the opinions, feelings, and emotions expressed in tweets posted on the social network Twitter. It gives a corpus that is fully annotated by human annotators based on principles and rules defined in an annotation guide [6]. Thus, each tweet is annotated by means of six groups, one of which includes 18 fine categories, and five relationships. [7]

*2017 Edition

The 2017 edition focuses on opinion analysis and figurative language in tweets. Because the automatic detection of figurative language is a very active and vast subject of research in social networks and figurative language detection has gained relevance recently due to its importance for efficient sentiment analysis [3]. This detection and its role in the analysis of feelings has been the subject of several evaluation campaigns in recent years such as the SemEval campaign on tweets in English [11] and the SENTIPOLC @ Evalita campaigns on tweets in Italian in their 2014 and 2016 editions [12], [13]. The 2017 edition of the DEFT is the first evaluation campaign around these themes for French; it is interested in the influence of figurative language (especially irony, sarcasm, and humor) in the analysis of opinions in tweets in French. In fact, figurative language diverts the proper meaning to assign a meaning said figurative or imaged such as irony, sarcasm, humor, metaphor, or play on words. Hence, all tweets are annotated in both figurative/non-figurative and polarity according to the annotation guide provided by the organizers of the challenge. [3]

*2018 Edition

The 2018 edition presents IRISA's participation in DEFT, which focuses on the classification and annotation of opinion in tweets. This participation was made by IRISA's LinkMedia team who developed systems for the 2018 DEFT evaluation campaign based on opinion analysis in French tweets. Four tasks of increasing levels of complexity were proposed to participants, and the team participated in 3 [10], [4].

III. TASK TYPE

Several tasks have been proposed during the different editions in order to deal with the opinion mining in a complete way. These tasks are listed according to three types of research (simple research objective/subjective, fine level research, and figurative language research) (Table II).

IV. OPINION DETECTION APPROACHES

A. Opinion Detection and Sentiment Analysis Approaches

There are three types of approaches to opinion detection and sentiment analysis:

1- Statistical approaches: also called supervised classifications consist of grouping words into two axes of classification either in the opposition (subjective-objective), or in the distinction of subjective opinions in the opposition (positive-negative).

2- Symbolic approaches: also called unsupervised classifications (based on lexicon) use dictionaries of subjective words (the latter can be general or constructed manually).

3- Hybrid approaches: also called semi-supervised classifications consist in combining the strengths of the two previous approaches by taking into account all the linguistic treatment of symbolic approaches before launching the learning process as in statistical approaches.

TABLE II
TASK TYPE

	Simple research (objective/ subjective)	Fine level research	Figurative language research
2015 edition	1.determine the overall polarity of the tweet (positive, negative, mixed, neutral) 2.identify the generic classes (opinion, feeling, emotion, information) and specific (among the 18 thin categories) of these tweets	3.analyze the source, the target and the expression of opinion, feeling or emotion	
2017 edition	1.determine the overall polarity of non-figurative tweets 2.determine the overall polarity of figurative and non-figurative tweets		3. determine whether a tweet contains yes or no figurative language
2018 edition	1 .classification of tweets according to whether they relate to transport or not. 2. classification of tweets according to their polarity	3. annotation of the opinion markers and the object about which an opinion is expressed	

B. Methods Used by Participants

The majority of the methods used by the participants were based on supervised machine learning approaches, the main algorithms used are SVM, Naïve Bayes, neural network, PPMC, K nearest neighbors, and decision tree boosting; and on approaches based on lexicons of opinions, feelings and emotions such as ANEX, Affect, Lidilem, Casoar, Emotaix, Feel, Polarimots, Diko, labMT and DES (Table III) [1], [5]

For the 2015 edition, no team participated in task 3, fine annotation of opinions, feelings, and emotions indicating (judging) that is too difficult task. [7]

For the 2017 edition, most of the participants did not have recourse to the specific methods for detection of figurative language, whatever the task, the same approaches are used. [3]

For the 2018 edition, to deal with tasks 1 and 2, LinkMedia used a decision tree boosting algorithm (Bonzaiboost) and recurrent neural network (RNN). For task 3, the use of RNNs associated with CRFs (Conditional Random Field) has been experimented. Task 4 to which the team did not participate is used to determine the entity that expresses the opinion (source), the negations, the modifiers as well as the relationships between these elements. LinkMedia was the only team that participated in task 3, so it had no points of comparison with other approaches [10], [8]. Moreover, the relations between the sentiment target and the OSEE have been extracted with a simple proximity rule, in other words, a target is related to the nearest opinion markers in terms of number of words. The team defined this rule after observing examples of the corpus, but it did not perform experiments to verify its validity. [10]

TABLE III
THE DIFFERENT METHODS USED, THEIR ADVANTAGES AND DISADVANTAGES

Approaches	Advantages	Disadvantages
Lexicon-based approaches	- Domain independent - Fast time - Does not need labeled data	- Requires dictionaries that cover lot opinion words - Low accuracy - Needs strong linguistic resources
Machine learning approaches	-Unnecessity of dictionaries -High accuracy of classification -High precision and adaptability	- Dependent on the domain - Slow time - Needs human participation and labeled data
Bonzaiboost	- Relevant in the area of language processing and learning - Easy to implement - Has theoretical convergence results	- Boost misclassified examples (in the case of noisy corpora, the algorithm persists in classifying them).
BiLSTM+ Softmax	- BiLSTM (Bidirectional Long Short Term Memory): LSTM has three gates (input, output, and forget gate) -BiLSTM is more precise on the dataset using a longer sequence - With big data, BiLSTMs with higher expressiveness can lead to better results.	-BiLSTM require more memory to train -BiLSTM requires more time to train - Dropout is much more difficult to implement in BiLSTMs
BiGRU+ CRF	-BiGRU (Bidirectional Gated Recurring Units): GRU has two gates(reset and update gate) -BiGRU uses fewer training parameters and therefore uses less memory - Run faster and train faster than BiLSTM - The BiGRU unit does not need to use a memory unit to control the flow of information like the BiLSTM unit.	- When there is a larger dataset, BiLSTM work better - The BiLSTM has more parameters than BiGRU. So he learns more complex assumptions - BiGRU is not more efficient than BiLSTM. There are tasks where BiGRU outperforms BiLSTM and tasks where BiLSTM outperforms BiGRU

V. DISCUSSION AND CONCLUSION

The set of all these tasks can cover a large share of possible work in terms of analysis of the opinions, feelings, and emotions applied to short messages posted on social networks. In addition, current systems of automatic classification of the subjective or objective nature of a document are completing good results [14] while, the results on the task of polarity analysis remain inconclusive [3].

For the 2017 edition, participant submissions were evaluated using standard measures of precision and f-measurement. So, the best results for the 3 tasks in macro-f-measure clearly reveal that the use of figurative language greatly complicates the analysis of opinions. So it is necessary to develop systems and approaches allowing to analyze and annotate opinions containing figurative language. Also, it is necessary to note that in most editions of DEFT, the fine annotation task of opinions, feelings, and emotions is not achieved indicating that it is a difficult task. Only the team IRISA LinkMedia in the 2018 edition participated in this task which makes comparison with other approaches not possible. Moreover, the rule defined by this team concerning the relations between the sentiment target and the OSEE must be tested to ascertain and ensure its validity.

REFERENCES

[1] B.Azzeddine, A. Harbaoui, and BEN Ghezala H, *“Sentiment Analysis Approaches based on Granularity Levels”,* 2018.

[2] H. Ali, et al. *"Détection d'opinion: Apprenons les bons adjectifs!." INFORSID'08: INFormatique des Organisations et Systèmes d'Information et de Décision-Atelier FODOP'08*. 2008.

[3] Benamara, F., Grouin, C., Karoui, J., Moriceau, V., & Robba, I. (2017). Analyse d'opinion et langage figuratif dans des tweets: présentation et résultats du Défi Fouille de Textes DEFT2017.

[4] Graceffa, D., Ramond, A., Dusserre, E., Kalitvianski, R., Ruhlmann, M., & Padró, M. (2018). Notre tweet première fois au DEFT-2018: systèmes de détection de polarité et de transports (Systems for detecting polarity and public transport discussions in French tweets). In *Actes de la Conférence TALN. Volume 2-Démonstrations, articles des Rencontres Jeunes Chercheurs, ateliers DeFT* (pp. 287-298).

[5] Belbachir, F. (2010). Expérimentation de fonctions pour la détection d’opinions dans les blogs. *IRIT. Université Toulouse*, *3*, 95.

[6] Fraisse A. & Paroubek P. et al. (2014). Guide d’annotations d’opinions/sentiments/émotions pour le projet uComp, 2014.

[7] Hamon, T., Fraisse, A., Paroubek, P., Zweigenbaum, P., & Grouin, C. (2015, June). Analyse des émotions, sentiments et opinions exprimés dans les tweets: présentation et résultats de l'édition 2015 du défi fouille de texte (DEFT).

[8] Paroubek, P., Grouin, C., Bellot, P., Claveau, V., Eshkol-Taravella, I., Fraisse, A., ... & Torres-Moreno, J. M. (2018, May). DEFT2018: recherche d'information et analyse de sentiments dans des tweets concernant les transports en Île de France.

[9] Martineau, C., Voyatzi, S., Varga, L., Brizard, S., & Migeotte, A. (2011, October). Détection fine d'opinion et sentiments: attribution fine de la polarité et calcul incrémental de l'intensité.

[10] Minard, A. L., Raymond, C., & Claveau, V. (2018, May). Participation de l’IRISA à DeFT 2018: classification et annotation d’opinion dans des tweets.

[11] Ghosh A., LI G., Veale T., Rosso P., Shutova E., Barnden J. & REYES A. (2015). Semeval-2015 task 11 : Sentiment analysis of figurative language in twitter. In Proc of SemEval, p.470–478, Denver, CO.

[12] Basile V., Bolioli A., Nissim M., Patti V. & Rosso P. (2014). Overview of the Evalita 2014 SENTIment POLarity Classification Task. In Proc of EVALITA, p. 50–57, Pisa, Italy : Pisa University Press.

[13] Barbieri F., Basile V., Croce D., Nissim M., Novielli N. & Patti V. (2016). Overview of the Evalita 2016 SENTIment POLarity Classification Task. In Proc of Third Italian Conference on Computational Linguistics (CLiC-it 2016) and Fifth Evaluation Campaign of Natural Language Processing and Speech Tools for Italian, Napoli, Italia : CEUR Workshop Proceedings.