

Article

Not peer-reviewed version

---

# Arabic Opinion Mining Using Machine Learning Techniques: Algerian Dialect as a Case of Study

---

[Mostefa Kara](#)\*, [Abdelkader Laouid](#), [Ahcène Bounceur](#), [Omar Aldabbas](#)

Posted Date: 25 January 2023

doi: 10.20944/preprints202301.0443.v1

Keywords: Algerian dialect ; Opinion mining ; Sentiment analysis ; Emotional detection ; Social web



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

## Article

# Arabic Opinion Mining Using Machine Learning Techniques: Algerian Dialect as a Case of Study †

Mostefa Kara <sup>1</sup>, Abdelkader Laouid <sup>1</sup>, Ahcène Bounceur <sup>2</sup> and Omar Aldabbas <sup>3</sup>

<sup>1</sup> LIAP Laboratory, University of El Oued, PO Box 789, El Oued 39000, Algeria, Algeria; karamostefa@univ-eloued.dz (M.K.); abdelkader-laouid@univ-eloued.dz (A.L.)

<sup>2</sup> Lab-STICC UMR CNRS, University of Western Brittany UBO, Brest 6285, France, France; Ahcene.Bounceur@univ-brest.fr

<sup>3</sup> Faculty of Engineering Technology, Al-Balqa' Applied University, Amman 15008, Jordan, Amman, Jordan; o.aldabbas@bau.edu.jo

† Proceedings of the 6th International Conference on Future Networks & Distributed Systems (ICFNDS 2022), December 15–16, 2022, Tashkent, Uzbekistan

**Abstract:** Social networking services such as Facebook, Twitter, and YouTube are fertile ground for analyzing texts, extracting opinions, and identifying feelings, due to the large number of texts and their diversity in all areas of life. In this manuscript, we apply four algorithms to classify tweets written in the Algerian dialect. To extract feelings, we used six features based on three polarities. In the presented work, we manually annotate a corpus of 2,891 texts and create an Algerian lexicon of idioms that contains 1328 annotated words. Our results show that there are improvements gained on the accuracy of the system, where we have achieved a better accuracy of 85.31%.

**Keywords:** Algerian dialect; opinion mining; sentiment analysis; emotional detection; social web

## 1. Introduction

Today, life is largely dependent on sensing, information extraction and analysis [1]. This information is longer available today and more precisely in digital form. With the development of Web 2.0. more and more people are communicating, sharing content, and expressing their opinions on the internet about a variety of topics, in discussion groups, blogs, forums, and other sites for product reviews [2].

At the end of 2013, Facebook opened its pages for customer recommendations, the chosen products can be rated by their fans with one of the values from 1 to 5 where 1 is very negative polarity, and 5 is very positive polarity, opinions available on the internet have a considerable impact on internet users. Surveys show that most users (80%) have already researched opinions on a product or service and that they pay twice as much for a product where its opinion is more affirmative than another [3]. Companies take these parameters into account and they know that opinion analysis is an important component of decision-making. We can see the usefulness of opinion detection in the fields of marketing, politics, psychological studies, health, road safety, tourism, etc.

As complex algorithms continue to be developed for predictive analysis of sensitive data, it is desirable that more and more computations are performed on private data as well as data stored on the blockchain [4–7]. Encryption provides a tool to ensure the privacy of private information [8–10], but it limits the functionality of operating on this data. So we are now doing our study on unencrypted data.

The objective of this work is to study public opinions which mainly concern the sentiments found in Facebook comments written in Arabic and more specifically in the Algerian dialect. These sentiments are expressed by polarities, we specify by the polarity the position of an opinion on this axis comprising different levels of positivity and negativity, therefore the polarity can be categories such as positive, negative, and neutral, or it can be defined by values describe the degree of positivity (or negativity), for example, a value between 1 and 5, where 1 denotes a very negative polarity, and 5

denotes a very positive polarity, it also takes the name "semantic orientation" [11]. So, one has a single polarity opinion where the text is associated with a single idea and an aspect-based opinion where it has to identify the polarity of each and every one of those aspects, nevertheless. For example, aspects of a washing machine are: how fast, how much electricity it uses, how much to wash, etc. While the aspects of a computer are: its weight, speed, storage space, lifespan, etc.

After determining the polarity, it is necessary to determine the Analysis approaches. We used supervised learning which involves the presence of two sets of data, a training set, and a test set. The method is called supervised since the system is trained on a training subset that contains already processed models (in our case, the models are Facebook comments). On the other hand, we distinguish three opinion mining approaches, lexicon-based, corpus-based, and mixed approach. Lexicon-based approaches (which we used in this work), are also called symbolic or linguistic. So far, most studies of sentiment analysis have relied on this method. It identifies the polarity of a text using two sets of words, those which express a positive sentiment and those which express a negative sentiment. The corpus-based approach is also called statistics. Conversely to the previous method, it does not need a lexicon of positive and negative words, it needs two annotated corpora (possibly only one, if using unsupervised learning). The first corpus is intended for learning, and the second corpus is intended for testing to verify the performance of the model.

Figure 1 show an example of a part of the positive and negative collected dictionary.

Arabic	English	Arabic	English
تألّق	shine	عنف	violence
تهبّل	enchant	فاشل	failed
مزيان	nice	مهزلة	farce
روعة	magnificence	خابج	Dirty
برافو	bravo	تبليط	lie
تحياتنا	greetings	رهج	homely

**Figure 1.** Example of our dictionary part

The remainder of this paper is organized as follows: In Section 2, we will see some related work. In Section 3, we explain our proposed model. Section 4 presents an experiment result. Finally, we conclude with Section 5.

## 2. Related work

Sentiment analysis, also sometimes referred to as opinion mining, is a sub-field of computer science, it is considered a part of automatic natural language processing and aims to classify sentiments expressed in texts. Efforts such as [12] investigate the use of computers to better understand natural language. In this context, several works are carried out in all known fields of application and with different sub-objectives (corpus construction, opinion detection, comparison of features, application of methods, etc.). Pang [13] did a study that categorized the sentiments of movie reviews, he was the first to experience machine learning [14].

In [15], the authors set out to focus on the economy and more specifically on product reviews. They started with the first phase (data collection), then the preprocessing going to classification. Their corpus was collected by themselves from several web resources like reviewzat.com and jawal123.com in the form of a set of text documents. Each document is a product represented by its type, they have selected five types of products that make up this corpus, the types are Camera, Laptop PC, Cell phone, Tablet, and Television. The corpus contains 250 documents, 2,812 sentences, and 15,466 words. Two reviewers worked on the labeling of opinions, the first is an expert for product review and the second

is an Arabic language specialist and a third non-specialist was used just to validate the choices of two other reviewers to have some degree of annotation reliability.

For classification, they started with stemming, which is the process of removing all the prefixes and suffixes from a word to produce the stem or the root. The process of radicalization is difficult in Arabic because, for example, the radicalization of two words (wonderful) and (terrible) gives the word (horror). While these two words have reversed polarities. They performed their tests with three classification algorithms, which are Support Vector Machines (SVM), Naive Bayes (NB), and Nearest Neighbor (KNN).

The work of Salima mdhaffer et al. [16] was in the Tunisian dialect. The work is a survey of the resources available for Sentiment Analysis SA in the Arabic language, MSA and dialectic. They have created a training corpus available free of charge for the Tunisian dialect and the performance evaluation of the Tunisian SA dialect system in several configurations. Their corpus called TSAC consists of comments written on the official pages of Tunisian radio and television channels, namely Mosaïque FM, Jawhra FM, Shemes FM, HiwarElttounsi TV, and Nessma TV during a period from January 2015 to June 2016. TSAC contains 8215 Positive Comments and 8215 Negative Comments.

In the case of the Moroccan dialect, Abdeljalil Elouardighi et al [17] targeted Moroccan newspapers which published online comments on the Moroccan legislative elections which took place on October 07, 2016. The collection was carried out using "Facebook Graph API". During 70 days they selected 10,254 comments. After the treatment of comments (cleaning and normalization of the text such as the removal of signs, symbols, repeated letters, stop words, etc.), 6581 comments are annotated; 2908 negative and 3673 positive. They used three algorithms, Naive Bayes (NB), Random Forests (FA), and Support Vector Machines (SVM), 50% dataset for learning, 25% for validation, and 25% for testing.

Nora Al-Twairish et al. [18] have developed a hybrid sentiment analysis method for Arabic tweets in the Saudi Arabian dialect, combining a lexicon-based method and a corpus-based method. The authors made the presentation and the evaluation of a set of features for sentiment analysis using a backward selection method, also the development and comparison of three classification models for the SA Saudi tweets. They used the SVM classifier because it was reported in the majority of SA studies of tweets as the best performing classifier [19].

The authors in [20] worked on the Algerian dialect (ALGD), their approach is lexicon-based. To make their model, the authors created three lexicons, keyword lexicon, negation word lexicon, and intensity word lexicon. The authors used two other resources, a list of emoticons with polarities assigned to them and a dictionary of common expressions of the ALGD. The keyword lexicon contains 3093 words (713 positive words and 2380 negative words). The authors collected and annotated their own dataset which contains 7698 Facebook comments.

In [21], the paper classified social network texts written in Algerian dialect in Latin script retrieved from Twitter, Facebook, and YouTube. This classification was a positive and negative class approach: lexical-based, machine learning, and hybrid. The authors involved a regrouping procedure in the preprocessing phase to overcome the problems related to the Algerian dialect such as the orthographic varieties to express the same word.

### 3. Proposed model

In this section, we will present the model proposed for opinion analysis, the description of the data source on which the model is applied, and the phase of corpus annotation and lexicon construction.

#### 3.1. Data source

We exploited the dataset used in the work of [22], where the authors built their own corpus in Algerian dialect. This is done by developing a tool with the python programming language, which allows querying the Facebook API to retrieve comments. Their dataset is divided into three parts, Normal, Offensive, and Obscene. We worked on the first part because it does not contain unclean (dirty) lyrics. This part is a mixture of texts with positive, negative, and neutral sentiments. In addition,

the other two parts (offensive and obscene) are considered negative polarity texts, which will unbalance the polarity distribution in our dataset. Table 1 shows the number of comments and their sources.

**Table 1.** Collected data

Source	Normal	Offensive	Obscene	Total Comments
Facebook	2892	1497	611	5000

Like any data-gathering operation, this step requires preprocessing to get a clean and ready-to-use corpus. The preprocessing consists of filtering downloads to keep only Arabic texts and to eliminate stop words.

### 3.2. Annotating

Annotating or tagging opinions is a human task that is a bit difficult as it takes a long time to follow the comments one by one. Sometimes, it needs a lot of discussion to reach a final decision if this opinion is positive or negative! With two annotators (a third so that the conflict between the annotators will be resolved by a majority vote), we have labelled all 2891 entries by using three polarities: positive, negative, and neutral (which have the values 1, -1, 0 respectively). Table 2 divides the dataset according to the three values of polarities.

**Table 2.** Number of comments by polarity.

Polarity	Positive	Negative	Neutral	Total
Number of comments	975	525	1391	2891

In Figure 2, we have described some examples with their polarities.

Comment	Polarity
هذي جياحة نريسكي بعمرى على حاجة تستاهل ماشي علاجال بوزوج رجلين يجري مور البالون	negative
شفت المواطن مش واعي خلاص داي الزبلة حذاه يكل منا او يرمي منا	negative
في بسكرة خويا لعزیز طریق سیدی عقبه ؟	neutral
هادو روشيات تاع رخام ماشي ثلج	neutral
ربي يكثر من امثالك	positive
درس في حسن ضيفة كل إحترام وتقدير ناس المدية	positive

**Figure 2.** Example of comments annotation

The representation of Figure 2 in English is shown by Table 3.

**Table 3.** Example of comments annotation.

Comment	Polarity
It's trivial, am I risking my life for something not worth it? For someone running behind the ball	negative
I saw the citizen is never sane, makes the dumpster near him, eats here and throws here	negative
Is the road to Sidi Uqba in Biskra?	neutral
These are marble rocks, not ice	neutral
May God multiply people like you	positive
A lesson in hospitality, all the respect and appreciation of the inhabitants of Medea	positive

### 3.3. Dictionary creation

The dictionary is a collection of words that a classifier can use to assess the polarity of a text. In the literature, we have not found a specific lexicon of words in the Algerian dialect. So, we will have to establish our own dictionary, we collected 1328 words from websites. Also, we asked for help from friends in northern Algeria, south, east, and west to cover as many regions of Algeria as possible. These words are labelled as shown in Table 4.

**Table 4.** Statistics of our dictionary

Polarity	Positive	Negative	Total
Number of words	565	763	1328

## 4. Experiments

In our experiment, we used two PCs, the first was HP Pavillon and the second was Dell, both are multi-core I3 processors, 2.40 GHz frequency clocks, and 4 GB RAM. For application programming, we used the Python environment using the following packages: CSV, Gensim, and Scikit-learn. We will mainly implement four algorithms. These algorithms are Support Vector Machines (SVM), Decision Tree (DT), Random Forest (RF), and Naïve Bayes (NB). The general algorithm of the sentiment analysis was as follows:

---

### Algorithm 1 Analysis algorithm

---

**Require:** libraries  
 1: Read the dataset  
 2: Read the dictionary  
 3: **function** ANA  
 4:     Extract features  
 5:     Preparation  
 6:     Call to classifiers  
 7:     Show the results  
 8: **end function**

---

### 4.1. Features

As we saw previously in the work of Nora Al-Twairish et al. [18] where they found that the features that have the most important influence in the analysis of opinions are these four, the existence of positive words (and/or negative) and the number of positive (and/or negative) words in the comment. Based on their studies, we are going to do our experiment where we added on these four features two more, which are the languor of the text and the level of the sentiment. The sentiment level designates its depth, to apply this feature we have used SemEval2016 document [23]. Figure 3 presents part of this document.

Arabic	English	Sentiment level
حب	love	0.963
نجاح	success	0.925
صادق	honest	0.900
تانيب	reprimand	- 0.250
دموع	tear	- 0.287
هروب	escape	- 0.313

**Figure 3.** Example of our dictionary part with sentiment level

There is some work that uses the negation feature as a separate feature, but using the negation in this way will decrease the accuracy of the analysis [18]. In our work, we have integrated the negation



in the features: the number of positive words and the number of negative words, where the model checks whether this negation is followed by a positive (negative) word or not! If so, the model adds 'one' to the number of negatives (positives) words, if the negation is followed neither by a positive word nor by a negative word, the model does not mean anything. Table 5 shows the used features in our implementation.

Table 5. The used features

Feature	Abbreviation	Meaning
Has Positive Word	HPW	0 or 1
Has Negative Word	HNW	0 or 1
Positive Word Count	PWC	$\geq 0$
Negative Word Count	NWC	$\geq 0$
CommentLength	CL	Digital $> 0$
SentimentLevel	SL	$-1 \leq V \leq 1$

The following algorithm presents the source code which allows the counting of positive words:

Algorithm 2 Count Pos algorithm

**Require:** corpus.text, pos.text

```
1: function COUNTPOS
2:   cp = []
3:   count = 0
4:   for t in corpus.text do
5:     for w in t.split do
6:       for w1 in pos.text do
7:         if w = w1
8:           count ← count + 1
9:         end for
10:      end for
11:      cp.append(count)
12:      count ← 0
13:    end for
14:    corpus[pos - count] ← cp
15: end function
```

To test for the existence of a positive word, we used the following algorithm:

Algorithm 3 Existence Pos algorithm

**Require:** corpus.text, pos.text

```
1: function EXISTPOS
2:   op = []
3:   for t in corpus.text do
4:     IF any (w in t.split() for w in pos.text) '
5:       op.append(1)
6:     Else
7:       op.append(0)
8:     end for
9:     corpus[one - pos] ← op
10: end function
```

4.2. Results

Because we use the supervised learning method and lexicon-based approach, we have divided the corpus into two parts, 80% for training and 20% for testing. We carried out several tests, the Accuracy results are presented in Table 6.

**Table 6.** Classification results

Test	Feature	SVM	DT	RF	NB
1	All features	85.14	83.07	85.31	84.28
2	HPW, HNW, PWC, and NWC	84.11	84.45	84.28	82.38
3	PWC, NWC, CL, and SL	84.62	83.24	85.31	83.76
4	PWC and NWC	84.11	84.28	84.45	67.87
5	HPW, HNW, CL, and SL	84.11	84.11	84.11	48.35
6	HPW and HNW	84.11	84.11	84.11	72.30
7	CL and SL	47.49	47.66	48.18	48.18

The best result in all the tests is 85.31% got by the RF with the use of all the features. The best results of SVM and NB (with the first test) are 85.14%, 84.28% respectively, on the other hand, DT reached its maximum measurement (84.45%) in test (2). From the tests (2), (4), and (6), we noticed that the two couples (PWC, NWC) and (HPW, HNW) had almost the same influence weight, this is logical because the number of mixed comments (which contain positive words and at the same time negative words) is usually small. This is shown when we exploited the number of polarity words, the measure was a little larger (from 84.11% to 84.45%); except NB, which has had a considerable variation. According to tests (3) and (4), we found that adding features (CL, SL) to features (PWC, NWC) improved the results of classifiers; except for DT, the result was lowered. In tests (5) and (6), we found that adding features (CL, SL) to features (HPW, HNW) did nothing except that the measure of NB was decreased. From the last test, we found that the features (CL, SL) cannot be alone, since if they do, they will give the poorest result (less than 49 %). These results show that SVM is generally regarded as a better classifier. Table 7 summarizes the comparison with other works.

**Table 7.** Results comparison

Works	Classifier	Accuracy	F-measure
ours	RF	85.31	84.93
[21]	SVM	81	81
[20]	/	79.13	/
[16]	MLP	NA	78
[17]	SVM	/	78
[18]	SVM	NA	69
[24]	SVM	82.3	83.9
[24]	DL	85.3	86.8
[25]	DL	71.8	66.2

## 5. Conclusion and future work

In this paper, we have done sentiment analysis on a manually annotated corpus that contains 2891 texts in the Algerian dialect labelled as the following: 975 positive texts, 525 negative texts, and 1391 neutral. The objective of which is the detection of the polarities of these tweets in three ways, a positive tweet, a negative tweet, and a neutral tweet. The most difficult phase is the creation of a lexicon of Algerian idioms because of the great difficulty of the dialect. We used six features and leveraged four machine learning classifiers which are Support Vector Machine (SVM), Decision Tree (DT), Decision Tree Forest (RF), and Bayesian Naive (NB). The evaluation of these classifiers is done by 20% of the corpus. We did seven different tests, the first was done using all the features and the other tests were done by substituting these features. We have found that the correct Accuracy (85.31%) is achieved by the Random Forest (RF) classifier. Finally, we cited some examples of analysis errors in our model and explained how the model made them. In future work, we intend to enrich our dictionary with more Algerian dialect words by covering other areas more broadly. Also, using other configurations such as bigram, trigram, and mixed.



## References

1. Hammoudeh, M.; Newman, R.; Dennett, C.; Mount, S.; Aldabbas, O. Map as a service: a framework for visualising and maximising information return from multi-modal wireless sensor networks. *Sensors* **2015**, *15*, 22970–23003.
2. Ahsan, M.S.A.; Arefin, M.S.; Kayes, A.; Hammoudeh, M.; Aldabbas, O. A Framework for Identifying Influential People by Analyzing Social Media Data. *Applied Sciences* **2020**, *10*, 8773.
3. Gillot, S. Fouille d'opinions. In *Traitement du texte et du document*; 2010.
4. Kara, M.; Laouid, A.; Hammoudeh, M.; Alshaikh, M.; Bounceur, A. Proof of Chance: A Lightweight Consensus Algorithm for the Internet of Things. *IEEE Transactions on Industrial Informatics* **2022**.
5. Kara, M.; Laouid, A.; Bounceur, A.; Lalem, F.; AlShaikh, M.; Kebache, R.; Sayah, Z. A Novel Delegated Proof of Work Consensus Protocol. 2021 International Conference on Artificial Intelligence for Cyber Security Systems and Privacy (AI-CSP). IEEE, 2021, pp. 1–7.
6. Kara, M.; Laouid, A.; Bounceur, A.; Hammoudeh, M.; Alshaikh, M.; Kebache, R. Semi-Decentralized Model for Drone Collaboration on Secure Measurement of Positions. The 5th International Conference on Future Networks & Distributed Systems, 2021, pp. 64–69.
7. Habib, A.; Laouid, A.; Kara, M. Secure Consensus Clock Synchronization in Wireless Sensor Networks. 2021 International Conference on Artificial Intelligence for Cyber Security Systems and Privacy (AI-CSP). IEEE, 2021, pp. 1–6.
8. Chait, K.; Laouid, A.; Laouamer, L.; Kara, M. A Multi-Key Based Lightweight Additive Homomorphic Encryption Scheme. 2021 International Conference on Artificial Intelligence for Cyber Security Systems and Privacy (AI-CSP). IEEE, 2021, pp. 1–6.
9. Kahla, M.E.; Beggas, M.; Laouid, A.; Kara, M.; AlShaikh, M. Asymmetric Image Encryption Based on Twin Message Fusion. 2021 International Conference on Artificial Intelligence for Cyber Security Systems and Privacy (AI-CSP). IEEE, 2021, pp. 1–5.
10. KARA, M.; LAOUID, A.; BOUNCEUR, A.; HAMMOUDEH, M.; ALSHAIKH, M. Perfect Confidentiality through Unconditionally Secure Homomorphic Encryption Using OTP With a Single Pre-Shared Key. *Journal of Information Science and Engineering* **2023**, *39*, 183–195.
11. Thonet, T. Modèles thématiques pour la découverte non supervisée de points de vue sur le Web. PhD thesis, Université de Toulouse, Université Toulouse III-Paul Sabatier, 2017.
12. Adel, B.; Meftah, M.C.E.; Laouid, A.; Hammoudeh, M. Machine Learning to Classify Religious Communities and Detect Extremism on Social Networks: ML to CRCs and DE Through Text Tweets on SNs. *International Journal of Organizational and Collective Intelligence (IJOICI)* **2022**, *12*, 1–19.
13. Pang, B.; Lee, L.; Vaithyanathan, S. Thumbs up? Sentiment classification using machine learning techniques. *arXiv preprint cs/0205070* **2002**.
14. Hermann, C. Entre Web 2.0 et 3.0: opinion mining. PhD thesis, Haute Ecole de Gestion & Tourisme, 2010.
15. Sghaier, M.A.; Abdellaoui, H.; Ayadi, R.; Zrigui, M. Analyse de sentiments et extraction des opinions pour les sites e-commerce: application sur la langue arabe, 2014.
16. Mdhaftar, S.; Bougares, F.; Esteve, Y.; Hadrich-Belguith, L. Sentiment analysis of tunisian dialects: Linguistic ressources and experiments. Third Arabic Natural Language Processing Workshop (WANLP), 2017, pp. 55–61.
17. Elouardighi, A.; Maghfour, M.; Hammia, H.; Aazi, F.Z. Analyse des sentiments à partir des commentaires Facebook publiés en Arabe standard ou dialectal marocain par une approche d'apprentissage automatique. *EGC*, 2018, pp. 329–334.
18. Al-Twairish, N.; Al-Khalifa, H.; Alsalman, A.; Al-Ohali, Y. Sentiment analysis of arabic tweets: Feature engineering and a hybrid approach. *arXiv preprint arXiv:1805.08533* **2018**.
19. Abdul-Mageed, M.; Diab, M.; Kübler, S. SAMAR: Subjectivity and sentiment analysis for Arabic social media. *Computer Speech & Language* **2014**, *28*, 20–37.
20. Mataoui, M.; Zelmatti, O.; Boumechache, M. A proposed lexicon-based sentiment analysis approach for the vernacular Algerian Arabic. *Research in Computing Science* **2016**, *110*, 55–70.
21. Bettiche, M.; Mouffok, M.Z.; Zakaria, C. Opinion mining in social networks for Algerian dialect. International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems. Springer, 2018, pp. 629–641.

22. Mansour, A.E. La détection automatique du discours abusif, offensant et obscène dans le dialecte algérien **2018**.
23. Mohammad, S.; Kiritchenko, S.; Sobhani, P.; Zhu, X.; Cherry, C. Semeval-2016 task 6: Detecting stance in tweets. Proceedings of the 10th international workshop on semantic evaluation (SemEval-2016), 2016, pp. 31–41.
24. Abo, M.E.M.; Idris, N.; Mahmud, R.; Qazi, A.; Hashem, I.A.T.; Maitama, J.Z.; Naseem, U.; Khan, S.K.; Yang, S. A multi-criteria approach for arabic dialect sentiment analysis for online reviews: Exploiting optimal machine learning algorithm selection. *Sustainability* **2021**, *13*, 10018.
25. Mahdaouy, A.E.; Mekki, A.E.; Essefar, K.; Mamoun, N.E.; Berrada, I.; Khoumsi, A. Deep multi-task model for sarcasm detection and sentiment analysis in Arabic language. *arXiv preprint arXiv:2106.12488* **2021**.

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.