SENTIMENT ANALYSIS OF ARABIC COMMENTS USING MACHINE LEARNING AND DEEP LEARNING MODELS

Hassan Fouadi

LIIAN Laboratory Faculty of Sciences Dhar-Mahraz Fez, Morocco
hassan.fouadi@usmba.ac.ma

Hicham El Moubtahij

Modeling, Systems and Technologies of Information Team, University of Ibn Zohr, Agadir, Morocco;
h.elmoubtahij@uiz.ac.ma

Hicham Lamtougui

LIIAN Laboratory Faculty of Sciences Dhar-Mahraz Fez, Morocco
hicham.lamtougui@usmba.ac.ma

Ali Yahyaouy

LIIAN Laboratory Faculty of Sciences Dhar-Mahraz Fez, Morocco
ayahyaouy@yahoo.fr

Abstract

Sentiment analysis is a field of research that consists in analyzing the sensations, attitudes, and emotions of individuals towards entities such as products, services, and economic organizations. Likewise, the demand for Arabic Sentiment Analysis has grown rapidly due to the extensive use of the Arabic language in social media networks and has generated considerable interest from the research community. Arabic is one of the widely used languages on social networks. However, its morphological complexities, its dialect varieties, and its relatively few resources make it a challenging language for sentiment analysis. The main goal of our study is to implement and compare the performance of ASA by exploiting machine learning and deep learning models to automatically determine the sentiments by classifying them as positive or negative. For that, this study implements and evaluates a deep learning model namely the long short-term memory (LSTM) model, and three machine learning algorithms: Support vector machines (SVM), Logistic Regression (LR), K-Nearest neighbours (KNN). These classifiers are applied on the Arabic-Review (ARev) database that is manually annotated and collected from many Arabic resources. The results show that SVM and LR models are the best performing classifiers with an accuracy of 92% and 93% respectively.

***Keywords*: Arabic Sentiment Analysis; Deep Learning; KNN; LSTM; SVM**

1. Introduction

Sentiment Analysis (SA) refers to the use of natural language processing (NLP) and machine learning to identify and characterize the sentiment/polarity expressed in a piece of text [1][2]. It allows mining the immense increasing resources of shared opinions such as social networks, review sites, and blogs.

The reach of the Internet continues to expand globally. According to the International Data Corporation (IDC), the large amount of digital data generated globally exceeded 33 zettabytes in 2018, with an expected increase of 175 zettabytes by 2025. And the number of internet users worldwide has grown to 4,000 billion users. In recent years, the emergence of social media platforms as a medium of communication, and their influence on governance, development, diplomacy, and business has led to the generation of huge amounts of rich data that could be used to measure people's opinions and attitudes towards products, events, persons, news, or issues.

Many NLP researchers were attracted by this field and started exploring the applications of SA, and many research communities are working to improve this area from basic linear approaches to more complex deep neural network models [3]. However, much research has focused on English as it is the dominant language of science while other

languages have fallen behind. Recently, the increase of web content in Arabic, especially on social media, has seen an explosive growth rate in the number of internet users (130 million Arabs in January 2018) [4]. This has increased the demand for Arabic Sentiment Analysis (ASA) due to the heavy use of Arabic in social networks. As a result, there is now an evolving body of work on ASA. Arabic which is highly inflectional and derivational is an important language as far as the cultural, historical, and social aspects. Furthermore, it faces many challenges and raises important issues for SA due to its complex morphology and structure. The wide variation found in the Arabic language where there is Modern Standard Arabic (MSA), Classical Arabic (CA), and Dialects is one of the biggest challenges in building Arabic NLP resources. As a result, in some tasks, it may be necessary to create stand-alone resources for each variation where the available tools that were built for MSA cannot be adjusted for the other variations and vice-versa [5]. Indeed, the texts produced by social sites are unstructured. This is due to such factors as lack of standardization, misspellings, missing punctuation, repetitions, and non-standard words. Despite the importance of SA, and the various SA research that has been carried out over the past decade, the growth of this research has been quite slow. therefore, the analysis of Arabic sentiments still needs improvement.

The objective of this study is to implement and compare machine learning and deep learning models to automatically determine the sentiments of comments posted in several domains (tourist, news, sports, and politics) by classifying them as positive or negative. To do so, this work implements the deep learning algorithm: LSTM, and three machine learning algorithms: SVM, LR, and KNN. These four algorithms are applied to the ARev database that is manually annotated and collected from many Arabic resources. In this respect, 10,000 comments were tagged from the most popular Algerian Facebook pages. Added to this are the datasets of 'LABR', 'Elsahar and al.', and 'Mataoui and al.'. Furthermore, the study carried out trains an Arabic word embedding model suitable for ASA.

This paper is structured as follows: Section 2 presents the most famous works that are interested in ASA, in addition to citing and discussing some recent works implemented by machine learning and deep learning models. In section 3, the ARev dataset is presented by specifying its preprocessing steps. Then, the four classification models implemented in this study are described and clarified. The mathematical formulae of the parameters used to test and compare the four classifiers are placed in section 4 where the findings of the study are discussed and interpreted. Finally, section 5 offers the conclusion of this work and ends with a future perspective on the current topic.

2. Related works

Sentimental analysis is a subfield of Natural Language Processing NLP where there is a lot of work done for the English language, but it is noted that there is little research concerning the Arabic language and its dialects. Presented in this context are the famous works interested in ASA.

2.1 Sentiment analysis for the Arabic language

In 2020, Hassan and al. published the work described in detail the field of (ASA) [6]. They presented the different difficulties and challenges related to this field and then specified the main steps needed from corpus construction to the interpretation of the results. All the techniques and approaches implemented have been well explained and presented in this work in the form of a comparative study.

In 2021, Rahab and al. presented in a chapter the main challenges and recent advances in the field of ASA [7]. They started their paper by describing the difficulty of the Arabic language and its morphology, followed by the main approaches of classification, and then in the form of a literature survey, they cited the works published since 2011 that are relevant to the ASA field by explaining all the technique used and the results found. In this work, the authors not only present the main challenges of ASA but also propose solutions and recommendations. However, many challenges remain unresolved. Some of these are cited below:

- Arabic Language Complexity
- Dialectal and Non-Arabic Content (Diglossia)
- Arabizi or Romanized Arabic
- Proper Nouns Translation

2.2 Word embedding:

Rana and al. published in 2018 a work to improve Arabic word embedding by incorporating morphological annotations into the embedding model [8]. By using linear compositionality, they managed to fit the generated word vectors to their lemma forms, and then they used sentiment analysis, subjectivity, and Arabic analogy to evaluate the effectiveness of their approach where the results obtained show improvements over those existing in the literature of ASA.

In 2020, Faris and al. exploited the word embedding technique for the topic of automatic detection of cyber hate speech written by the Arabic language [9]. The idea of this work is to propose a new deep learning approach based

on the hybrid of two different classifiers that are convolutional neural network (CNN) and long-term memory network (LSTM). The evaluation database of the approach was a collection of tweets that represent hate expressions written in the Arabic language. The word embedding and deep learning implemented in this work allow to properly classify the tweets as normal or hateful.

In this work published in 2021, the authors proposed an unsupervised approach domain adaptation for Arabic Cross-Dialect and Cross-Domain based on the Word Embedding technique [10]. During the experimental phase, they adopted the fine-grained and coarse-grained taxonomies of Arabic dialects. The results show that the proposed approach increases the performance by exactly 20.8% compared to the BERT approach (zero-shot transfer).

2.3 Deep learning and machine learning in ASA

This part begins with the work of Bolbol and al. published in 2020, which is interested in the sentiment analysis of a set of Tweets written in Arabic, they implemented and compared three machine learning classifiers which are: (1) K-Nearest Neighbors (KNN), (2) Decision Tree (DT) and (3) Logistic Regression (LR) [11]. For the experimental part, they evaluated their approach on four Arabic text datasets. Moreover, four evaluation measures are used which are successive: precision, recall, f-measure, and accuracy. In the case of large datasets, the (LR) classifier achieves a better accuracy rate (93%) than the others, but for the case of small datasets (e.g., AJGT and ASTD datasets), the KNN and LR classifiers achieved better results than the DT classifier.

The approach proposed by El-Affendi and al. in 2021 is called the multilevel parallel attention neural (MPAN) model, which represents a new deep learning method [7,12]. They computed simultaneous contextualized embeddings at character, word, and sentence level by using a simple PBES: Positioning Binary Embedding Scheme. To produce better accuracies, the proposed approach computes at the output level the concatenated multilevel attention vectors. The results found are very encouraging; they obtained 94.25% for a tertiary classification collection and 95.61% for a binary classification collection. The validation of the model on a collection of IMDB film reviews (public data) gives an accuracy of 96.13%.

Finally, in 2021, the model entitled Deep ASA: Deep learning for Arabic Sentiment Analysis is exploited to predict the polarity of opinions and sentiments written in Arabic [13,14]. The proposed approach is based on two types of recurrent networks, which are LSTM and Gated Recurrent Unit (GRU). The approach shows that it is more robust to the data dependency problem. The results found in the experimental phase are very competitive. On the one hand, the accuracy found ranged from 81.11 to 94.32% and on the other hand, the classification error rate compared to the state-of-the-art models was reduced by 26%.

3. Methodology

3.1 ARev dataset

As shown in figure 1, after data consolidations, our preprocessing steps consist of three main phases: The data cleaning, the data transformation phase, and the data reduction.

3.1.1 Data collection:

The model is evaluated for the Arabic Text Sentiment Analysis task, and it is based on the ARev (Arabic Reviews) dataset [15]. ARev is built from four sources. The first is taken from 10,000 comments that were tagged as positive and negative comments, chosen from the 100,000 comments of the most popular Algerian Facebook pages. Added to this is the LABR dataset of book reviews [16]. This study also used the 'Elsahar and al.' dataset which is a multi-domain dataset [17]. The last dataset used is that of 'Mataoui and al.' which contains commentary data in Algerian Arabic [18] (Table 1).

Datasets	Type of language	Description
LABR [16]	Standard Arabic	Book reviews
Elsahar and al. [17]		Hotel reviews, restaurant reviews, product reviews, attraction reviews, movie reviews.
Mataoui and al. [18]	Algerian Dialect	Comments

Table 1. Various datasets used

3.1.2 Data pre-processing:

Most Arabic works using traditional methods have emphasized some pre-processing techniques. For this work a tokenization process is used. Through this process, the text is divided into units delimited by spaces or punctuation marks, the results are called tokens [19]. Then noise is removed to eliminate unwanted characters from the text. We start with removing non-Arabic letters, numbers, usernames, external links, and hashtags. After that, we proceed to the technique of normalization that converts a list of words to a more uniform sequence [20]. To do

that, we remove punctuation marks and Tashkeel (diacritics), then we delete repeated characters and duplicate letters. (Figure 1).

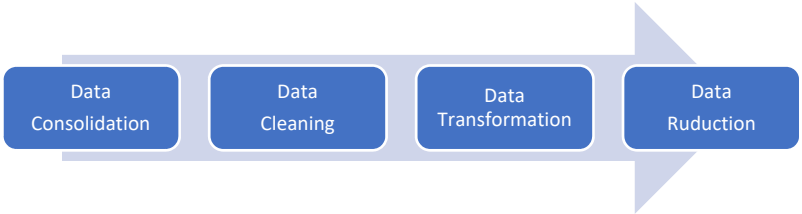

Figure 1. Preprocessing steps of ARev dataset

Table 2 present the statistics of the dataset after preprocessing and deleting the duplicate elements.

		Positive	Negative
Total	comment	24900	24900
	words	1180600	1345000
Average	words in each comment	47	53
	characters in each comment	253	294

Table 2. ARev dataset Statistics

The next step is feature selection/extraction. The purpose of this technique is to find the most pertinent features for the classification task, and the reduction of both dimensionalities of the feature space and processing time. The methods used for extracting features from the text into vectors are Bags of words (BOW) and TF-IDF.

3.2 Models of classification:

In this section, the four classifiers used in this work are presented and compared in terms of accuracy and performance. It should be mentioned that the training of these models is done on two portions of our dataset, where we reserved 15% for testing (3740 negatives, 3740 positive) and 85% for training (21192 positives, 21192 negative). The execution of our work is implemented on Google Colab (free version).

3.2.1 SVM model:

Support vector machines (SVMs), which are based on sound mathematical theory, fall within the human research development area of learning techniques [21]. The simplest case of SVMs is when the training data comes only from two different classes (+1 or -1), which is called binary classification. The idea of SVMs is to search for a hyperplane (straight line in the case of two dimensions) that best separates these two classes (figure 2).

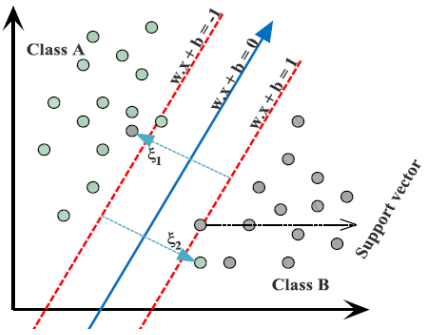

Figure 2. SVM binary

The separating hyperplane is represented by the following equation (1):

$$F1-score(11) = 2*\frac{(Precision*Recall)}{(Precision+Recall)} \quad H(x) = wT\ x + b \ (1)$$

Where w is a vector of m dimensions and b is a term.

The decision function, for an example x, can be expressed as follows:

Since the two classes are linearly separable, there are no examples that lie on the hyperplane, that satisfy H(x)=0. The following decision function should then be used:

$$\begin{cases} Classe = 1\ Si\ H(x) > 0\ (2) \\ Classe = -1\ Si\ H(x) < 0 \end{cases}$$

$$\begin{cases} Classe = 1\ Si\ H(x) > 1\ (3) \\ Classe = -1\ Si\ H(x) < -1 \end{cases}$$

The power of the SVM algorithm is in the kernel [22], the linear function is:

$$w^T x + b = b + \sum_{i=1}^{m} \propto_i x^T x^{(i)} (4)$$

where ∝ and x(i) are a vector of coefficients and a training example. The following two elements are replaced: x by the output of a given characteristic function g(x).

The scalar product by a function $k(x,\ x(i)) = g(x)$.

g(x(i)) called a kernel where [22]:

$$g(x(i)) = g(x)T(5)$$

3.2.2 Logistic Regression model:

Logistic regression is a mathematical model that combines a set of predictor variables (X) with a binomial random variable (Y) [23]. It is commonly used in the field of artificial intelligence (AI) and machine learning. Regression is a linear modelling technique, which studies the relationship between the main variable and explanatory variables. It is called logistic when there is a logistic function relationship between the variable of interest and the other variables. Logistic regression is also a multivariate analysis model, which can take different forms: binary or multinomial, logistic, or linear plus [24].

In the case of binary logistic regression, we get a simple answer with a choice between two solutions, that is, a binary type results like yes or no, true, or false, success or failure [24]. Mathematically, we try to find all the limits which separate the classes by applying the following formula [25]:

$$f(u) = \sum_{n=1}^{L} y^{(n)} \log\left(1 + e^{-u^T s}\right) + (1 - y^{(n)}) \log(1 + e^{u^T s})(6)$$

Such as:

- L represents the number of data – points.
- y (n) is the label of data – point n.
- s is the matrix of instances.
- u is the weight vector.

3.2.3 KNN model:

In machine learning, the k nearest neighbors' method KNN is a supervised learning model, it does not require any training step. The classification by this algorithm is done by calculating the distance between the training data and the test data to finally obtain the nearest neighbors [26, 27]. Two different strategies are applied to find the short distance between the points, the first one is based on the calculation of the Euclidean distance, and the second one is done by calculating the cosine similarity between the test and training data. For the case of this work, the first strategy was used [28].

In a Euclidean space, the documents are represented as points. With this algorithm we can calculate the Euclidean distance between two points (M = (x, y) and N = (a, b)) by applying the following equation [29]:

$$d(M, N) = f\overline{(x-a)^2 + (y-b)^2}(7)$$

The choice of the K-value to be used to predict with K-NN varies according to the dataset. Generally, the fewer neighbors (a small number of K) are used, the more prone to underfitting. On the other hand, the more neighbors (a larger number of K) are used, the more reliable the prediction will be.

The image above left (figure 3) represents points in a 2D plane with three possible labeling types (red, green, blue). For the 5-NN classifier, the boundaries between each region are quite smooth and regular. As for the NN classifier, we notice that the boundaries are "chaotic" and irregular. The latter comes from the fact that the algorithm tries to fit all the blue points into the blue regions, the reds with the reds, etc. It is a case of overfitting.

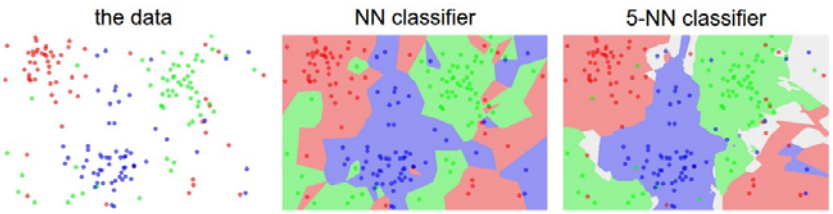

Figure 3. Effect of k in KNN

3.2.4 LSTM model:

Long Term Memory (LSTM) is a Recurrent Neural Network (RNN) architecture used in the field of deep learning. It has feedback connections, unlike direct propagation neural networks. In our work, we implemented this model using the TensorFlow library. To exploit word embedding, CBOW model are used, it based on Gensim which is an open-source toolbox library in Python. For training the model, the following parameters were used:

- The number of LSTM units = 64.
- The batch size = 20.
- Iterations = 100k.

4. Experiment and Result

4.1 Experiments setup:

The experiments were applied to the ARev dataset by using the Google Colab tool where we can take advantage of the performance of TensorFlow for the LSTM model. Figure 4 below shows the process overview of the classification followed in our experiments.

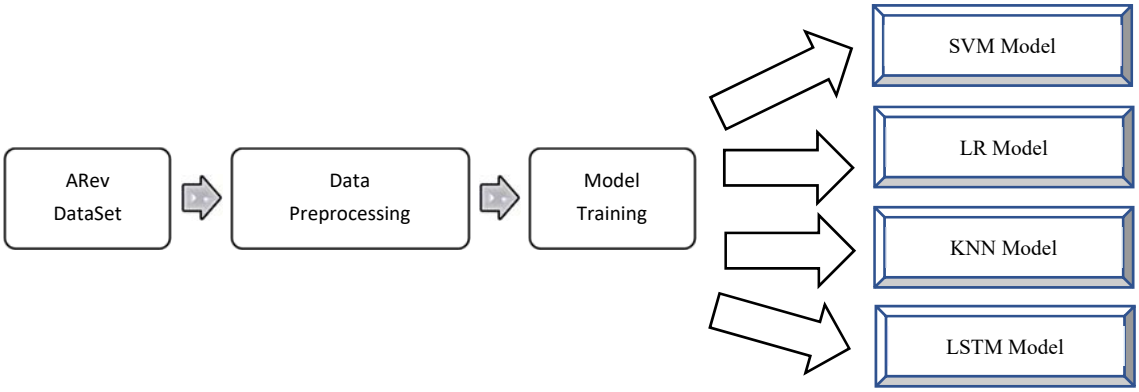

Figure 4. The methodology of our approach

We evaluated the performance of the classifiers used in this work by the following four metrics [30]:

$$Precision(8) = \frac{TP}{(TP + FP)}$$

$$Recall(9) = \frac{TP}{(TP + FN)}$$

$$Accuracy(10) = \frac{(TP + TN)}{(TP + TN + FP + FN)}$$

$$F1 - score(11) = 2 * \frac{(Precision * Recall)}{(Precision + Recall)}$$

Where:

- TP: True Positive
- TN: True Negative
- FP: False Positive
- FN: False Negative

4.2 Results and discussion:

Below are presented in detail the results found for this study's tested with the ARev dataset:

	F1-score	Recall	Precision	Accuracy
Negative	92%	91%	93%	92%
Positive	92%	94%	91%	

Table 3. Results of SVM model

	F1-score	Recall	Precision	Accuracy
Negative	47%	42%	54%	51%
Positive	54%	60%	49%	

Table 4. Results of KNN model

	F1-score	Recall	Precision	Accuracy
Negative	92%	91%	94%	93%
Positive	93%	94%	92%	

Table 5. Results of LR model

For the LSTM model, it gave an accuracy value of 81%. However, we obtained best results with the LR model, which scored the highest accuracy value of 93.0% and the average value of 92.5% for the F1-score (table 5). Next, we arrived at the accuracy value of 92% for the SVM model that also gives an average of 92.5% for Recall (table 3). In the last position, the KNN model achieves the mean precision value of 51.5% and the accuracy value of 51.0% (table 4).

The results presented above show that the LR and SVM models perform better than the LSTM and KNN models. Several reasons interpret these results, including firstly, the LSTM model works efficiently with a large dataset, so when the bigger dataset, the better the accuracy. On the other hand, for the KNN model, we need adaptation of some parameters which complicates the task of training the model. Finally, the choice between a positive and negative sentiment (2 choices only) gives priority to both LR and SVM models.

5. Conclusion and future work

Nowadays, Arabic Sentiment Analysis is of great importance in many areas such as politics, production, and services. Social networks are full of Arabic texts in which users express themselves on different topics where their opinions are of a great interest for ASA research. For the ASA domain, deep learning and machine learning tools and algorithms are indispensable for automatic data processing. In this context, we have used four classifier models: SVM, KNN, LR and LSTM. These models allow us to classify a set of opinions and reviews into the positive or negative ones.

The current study began with pre-processing the ARev database, which contains around 50000 comments and reviews; 85% of the dataset was chosen for training and 15% for testing. The performance of the classifiers used in this work was evaluated by four metrics: Precision, Recall, F1–Score, and Accuracy. The results show that the SVM and LR models outperform the KNN and LSTM models.

In future work, we will work through the AraBERT model applied on a very large database of comments and reviews written by the standard Arabic language and the Arabic dialects. In this respect, we will try to improve ASA by evaluating the efficiency of one of the state-of-art models, namely the AraBERT model.

Conflicts of interest

"The authors have no conflicts of interest to declare"

REFERENCES

[1] LIU, Bing. Sentiment analysis and opinion mining. Synthesis lectures on human language technologies, 2012, vol. 5, no 1, p. 1-167.
[2] Al Mansoori, S., Almansoori, A., Alshamsi, M., Salloum, S. A., & Shaalan, K. (2020). Suspicious Activity Detection of Twitter and Facebook using Sentimental Analysis. TEM Journal, 9(4), 1313.
[3] Kolkur, Seema, Gayatri Dantal, and Reena Mahe. "Study of different levels for sentiment analysis." International Journal of Current Engineering and Technology 5, no. 2 (2015): 768-770.
[4] Kemp S (2018) Digital in 2018: essential insights into internet, social media, mobile, and e-commerce use around the world
[5] Qwaider, Chatrine, Stergios Chatzikyriakidis, and Simon Dobnik. "Can modern standard Arabic approaches be used for Arabic dialects? Sentiment analysis as a case study." In Proceedings of the 3rd Workshop on Arabic Corpus Linguistics, pp. 40-50. 2019.
[6] Fouadi, Hassan, Hicham El Moubtahij, Hicham Lamtougui, Khalid SATORI, and Ali Yahyaouy. "Applications of Deep Learning in Arabic Sentiment Analysis: Research Perspective." In 2020 1st International Conference on Innovative Research in Applied Science, Engineering and Technology (IRASET), pp. 1-6. IEEE, 2020. DOI: 10.1109/IRASET48871.2020.9092163
[7] Rahab, H., Djoudi, M. and Zitouni, A., 2021. Sentiment Analysis of Arabic Documents: Main Challenges and Recent Advances. In Natural Language Processing for Global and Local Business (pp. 307-331). IGI Global., DOI: 10.4018/978-1-7998-4240-8.ch013
[8] Salama, R.A., Youssef, A. and Fahmy, A., 2018. Morphological word embedding for Arabic. Procedia computer science, 142, pp.83-93. DOI: 10.1016/ j.procs. 2018.10.463

[9] Faris, Hossam, Ibrahim Aljarah, Maria Habib, and Pedro A. Castillo. "Hate Speech Detection using Word Embedding and Deep Learning in the Arabic Language Context." In ICPRAM, pp. 453-460. 2020. DOI: 10.5220/0008954004530460
[10] El Mekki, Abdellah, Abdelkader El Mahdaouy, Ismail Berrada, and Ahmed Khoumsi. "Domain Adaptation for Arabic Cross-Domain and Cross-Dialect Sentiment Analysis from Contextualized Word Embedding." In Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, pp. 2824-2837. 2021. DOI: 10.18653/v1/2021.naacl-main.226
[11] Bolbol, Noor Khalid, and Ashraf Yunis Maghari. "Sentiment Analysis of Arabic Tweets Using Supervised Machine Learning." In 2020 International Conference on Promising Electronic Technologies (ICPET), pp. 89-93. IEEE, 2020. DOI: 10.1109/ICPET51420.2020.00025
[12] M. A. El-Affendi, K. Alrajhi and A. Hussain, "A Novel Deep Learning-Based Multilevel Parallel Attention Neural (MPAN) Model for Multidomain Arabic Sentiment Analysis," in IEEE Access, vol. 9, pp. 7508-7518, 2021, doi: 10.1109/ACCESS.2021.3049626.
[13] Alharbi, A., Kalkatawi, M. & Taileb, M. Arabic Sentiment Analysis Using Deep Learning and Ensemble Methods. Arab J Sci Eng 46, 8913–8923 (2021). https://doi.org/10.1007/s13369-021-05475-0
[14] Eletter, S. F. (2020). Sentiment analysis towards actionable intelligence via deep learning. TEM Journal, 9(4), 1663-1668.
[15] Abdelli, Adel, Fayçal Guerrouf, Okba Tibermacine, and Belkacem Abdelli, "Sentiment Analysis of Arabic Algerian Dialect Using a Supervised Method", In International Conference on Intelligent Systems and Advanced Computing Sciences (ISACS) IEEE, 2019, pp. 1-6, doi: 10.1109/ISACS48493.2019.9068897.
[16] Mohamed Aly and Amir Atiya. Labr: A large-scale Arabic book reviews dataset. In Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers), volume 2, pages 494–498, 2013.
[17] Hady ElSahar and Samhaa R El-Beltagy, "Building large Arabic Multi-domain resources for sentiment analysis", In International Conference on Intelligent Text Processing and Computational Linguistics, Springer, 2015, pp. 23–34, doi: 10.1007/978-3-319-18117-2_2.
[18] M'hamed Mataoui, Omar Zelmati, and Madiha Boumechache, "A proposed lexicon-based sentiment analysis approach for the vernacular Algerian Arabic", Research in Computing Science 110, no 1, pp. 55 –70, 2016, doi:10.13053/rcs-110-1-5.
[19] Barnaghi P, Ghaffari P, Breslin JG. Opinion mining and sentiment polarity on Twitter and correlation between events and sentiment. In: 2016 IEEE second international conference on big data computing service and applications (Big Data Service); 29 March–1 April 2016; 2016. p. 52–7.
[20] Faris H, Aljarah I, Habib M, Castillo P. Hate speech detection using word embedding and deep learning in the Arabic language context. In: Proceedings of the 9th international conference on pattern recognition applications and methods; 2020.p. 453-60.
[21] Corinna Cortes and Vladimir Vapnik. Support-vector networks. Machine learning, 20(3):273–297, 1995.
[22] Ian Goodfellow, Yoshua Bengio, Aaron Courville, and Yoshua Bengio. Deep learning, volume 1. MIT press Cambridge, 2016.
[23] M. A. Firyulina and I. L. Kashirina, "Classification of cardiac arrhythmia using machine learning techniques," doi: 10.1088/1742- 6596/ 1479/1 /012086.
[24] B. B. Frey, The SAGE encyclopedia of educational research, measurement, and evaluation. Sage Publications, 2018.
[25] B.-M. Hsu, "Comparison of Supervised Classification Models on Textual Data," Mathematics, vol. 8, no. 5, p. 851, 2020, doi: 10.3390/math8050851.
[26] I. A. A. Amra and A. Y. A. Maghari, "Student's performance prediction using KNN and Naïve Bayesian," in 2Ol7 8th International Conference on Information Technology (ICIT), 2017, pp. 909–913, doi: 10.1109/ICITECH.2017.8079967.
[27] S. Zhang, X. Li, M. Zong, X. Zhu, and R. Wang, "Efficient kNN Classification with Different Numbers of Nearest Neighbors," IEEE Trans. Neural Networks Learn. Syst., vol. 29, no. 5, pp. 1774–1785, 2018, doi: 10.1109/ TNNLS. 2017.2673241.
[28] D. A. Adeniyi, Z. Wei, and Y. Yongquan, "Automated web usage data mining and recommendation system using K-Nearest Neighbor (KNN) classification method," Appl. Comput. Informatics, vol. 12, no. 1, pp. 90–108, 2016.
[29] B. Trstenjak, S. Mikac, and D. Donko, "KNN with TF-IDF based framework for text categorization," Procedia Eng., vol. 69, pp. 1356–1364, 2014.
[30] Yacouby, Reda, and Dustin Axman. "Probabilistic Extension of Precision, Recall, and F1 Score for More Thorough Evaluation of Classification Models." In Proceedings of the First Workshop on Evaluation and Comparison of NLP Systems, pp. 79-91. 2020.

Authors Profile

Hassan Fouadi received the master's degree in computer science from High School of Technology at the University Sidi Mohamed Ben Abdellah, Fez, Morocco, where he is currently pursuing the Ph.D. degree. His research interests include sentiment analysis, machine learning, deep learning, and natural language processing.

Prof. Hicham El Moubtahij is currently a Professor of Computer Science at the University of Ibn Zohr, Agadir, Morocco. He received his Ph.D. in Computer Science from the University of Sidi Mohamed Ben Abdellah, Fez, Morocco, in 2017. He is now a member of the Systems and Technologies of Information Team at the High School of Technology at the University of Ibn Zohr, Agadir. His current research interests include machine learning, deep learning, Arabic handwriting recognition, Text Mining, and medical imagery.

Hicham Lamtougui received the master's degree in computer science from the Faculty of Science Dhar El Mahraz, Sidi Mohamed Ben Abdellah University, Fez, Morocco, where he is currently pursuing the Ph.D. degree. His research interests include machine learning, deep learning, Arabic handwriting recognition, and Text Mining.

Ali Yahyaouy is a professor at the Faculty of Sciences Dhar El Mahraz, Sidi Mohamed Ben Abdellah University, Morocco. He obtained his PhD degree at University of Technology of Belfort-Montbéliard, France and Sidi Mohamed Ben Abdellah University, Morocco in 2010. He is a member of the LISAC laboratory in Fez. He has been the coordinator of the University Qualification ''Software Engineering and Multimedia'' since 2014 and of the International Francophone Master ''Web Intelligence and Data Science'' since 2017. His research activities mainly concern Machine learning, Multi-agent Systems, Smart Grids and ITS.