SUSTAINABLE MACHINE INTELLIGENCE JOURNAL

Sustainable Machine Intelligence Journal

SCIENCES FORCE
International Academic Publisher

Detection of Depression from Arabic Tweets Using Machine Learning

Areej Alzoubi[1], Ahmad Alaiad[2], Khaled Alkhattib[3], Ahed J Alkhatib[4,*], Aseel Abu Aqoulah[5], Almo'men bellah Alawnah[6], and Ola Hayajnah[7]

[1] Department of Computer Information Systems Jordan University of Science and Technology Jerash, Jordan; azalzoubi19@cit.just.edu.jo
[2] Department of Computer Information Systems Jordan University of Science and Technology, Jerash, Jordan; aiaiad@just.edu.jo
[3] Department of Computer Information Systems Jordan University of Science and Technology, Jerash, Jordan; khatib@just.edu.jo
[4] Legal Medicine, Toxicology of Forensic Science and Toxicology Department Jordan University of Science and Technology, Irbid, Jordan; ajalkhatib@just.edu.jo
[5] Department of Computer Information Systems Yarmouk University, Irbid, Jordan; 2020162027@ses.yu.edu.jo
[6] Department of Industrial Engineering Jordan University of Science and Technology, Irbid, Jordan; akala-wanah19@eng.just.edu.jo
[7] Department of Computer Information Systems Jordan University of Science and Technology; hayaj-nahola@gmail.com

Event	Date
Received	05-01-2024
Revised	23-02-2024
Accepted	28-02-2024
Published	02-03-2024

Abstract: Depression has become the disease of the times and has caused suffering and disruption in the lives of millions of people around the world of all ages. Method: We obtained 16,581 Arabic tweets, whether they express depression or not, and the symptoms they contain for 1439 Arab Twitter users. We classified whether the user is depressed or not. We used many machine learning algorithms: DT, RF, Mutational Naïve Bayes, and AdaBoost , we also used feature extraction like BOW and TF-IDF. The result: Our experiments showed that Mutational Naïve Bayes with TF-IDF had the highest accuracy of 86% when rating tweets. Conclusion: Caring for the mental health of people is very important, as some measures must be taken to maintain the mental health of people in the early stages of infection.

Keywords: Depression, Twitter, Tweets, Mental Health, Machine Learning

1. Introduction

Feelings of sadness, frustration, loss of interest, and loss of pleasure in daily activities are familiar to everyone. In the last twenty years, lost years of life due to Depression have increased by 37.5% [1] [2]. The Depression Symptoms according to The American Psychiatric Association (APA) are: Depressed mood, Lack of interest or enjoyment of everything, Appetite and weight disorder, Slowed thinking and decreased physical movement, Loss of energy, Self-contempt, Impaired ability to think, Repeated thoughts of death and suicide, and Sleep disorder. Depression is one of the leading causes of disability among the world's population [3]. The rate of depression in the world is approximately 7% according to World Health Organization WHO, with significant financial costs to governments and health institutions [4]. According to a survey conducted by the Arab Barometer network for research, approximately 30% of Arabs have suffered from depression, where it is one of the fundamental reasons that lead an individual to think about suicide [5]. In the last few years, people have tended to use social networking sites to share their lives and feelings [6], as an effective way to express feelings of fear, anxiety, depression, and opinions [7] [8]. The Twitter platform helps to discover important things, people's interests and thinking trends, and a lot of information [19]. Therefore, Twitter was a source for collecting our data to analyze people's feelings and to reveal the depression among Arab Twitter users.

 *Correspondence: ajalkhatib@just.edu.jo

2. Related Work

The authors of [9] Aimed to analyze Facebook user posts in English to detect depression, where social networks are used to communicate with friends and share opinions and pictures that reflect the mood of users and express their feelings. Diagnosing depression using social media posts has become an important and global source. In this study, the aim was to investigate the possibility of detecting depression, the authors suggested machine learning technology as a powerful method that could be developed. The efficiency of the proposed method was evaluated using a set of psycholinguistic characteristics. The proposed method has the potential to significantly improve the accuracy of classification. It is worth noting that, through various experiments, the Decision Tree (DT) gives the highest accuracy between 60% and 80%, compared to other machine learning algorithms for detecting depression.

The authors of [10] used social media data to build an intelligent system that can efficiently deal with the problem of early detection of depression. This system, based on machine learning techniques, can manipulate data flows provided by users over time. This system can determine if the processing data is sufficient to classify the users. ERD tasks depend on the risky decisions that may affect an individual's life. In this paper, the authors introduce the SS3 model, which is a supervised instructional model capable of classifying texts. This SS3 module is designed to be used to deal with an ERD problem. The model was evaluated in CLEF's eRisk2017 pilot task for early detection of depression. The experimental results showed that this model can outperform other models despite being less computationally expensive and can explain its rationale.

The authors of [11] present this research that builds on the eRisk2017 pilot mission aimed at detecting early depression. The eRisk2017 dataset is transcripts collected from 887 Reddit users' messages. The main objective of the mission is to classify users into a state of risk of depression and status other than a risk of depression. This research takes into account a set of different features of the task of detecting depression among Reddit users by processing written messages. A bigram, bag of words, and embedding were used on the eRisk2017 dataset to assess the ability to apply morphological and morphological features. A different set of features that have a role in detecting depression have been discovered in social media platforms. Text messaging worked on Reddit. Bag of words, word embedding, and bigram work were discovered for compatibility with the eRisk2017 database. In Random Forest the TF-IDF with morphology achieved the highest scores on test data at 63% F1 and 92% accuracy. In SVM the embedding features were obtained at a higher recall of 63% than TF-IDF but give an accuracy of 92% and precision of 63.8%. This research yielded good results compared to the 2017 CLEF / eRisk Task Report.

The authors of [12] suggested that it is possible to take advantage of linguistic relationships and depression, to detect depressed social media users from their posts. The method in this study relies on feature sets that are the standard bag of words and surface features to more linguistically speaking. A database containing Reddit posts was used. The authors proposed a supervised model based on linguistic characteristics for screening social media users with depression. This model was evaluated and applied to two tasks: The first task was to check the user's posts, to determine if this user was depressed or not, and the second task was to determine the history of the user's writing for his participation, to facilitate early detection of depression symptoms. The model achieved the best results using the random forest to detect social media users with depression with the precision of 75%, recall 52%, and 92% accuracy.

The authors of [13] aimed to discover depression early by collecting publications from various social media platforms. In this study, the Twitter platform was used, and a set of features associated with depression were extracted. To detect Twitter users with depression, the authors proposed a depression dictionary learning model with multiple modes. This model provided an advantage over baselines (+ 3% to + 10%). The Twitter database was analyzed to be able to discover the basic behaviors among depressed and non-depressed social media users. Multimodal Depressive Dictionary Learning achieved the best results with an accuracy of 85% compared to Naive Bayes 79%. Depression detection models can detect depression, in addition to having the ability to detect depression early so that treatment can be timely.

The authors of [14] proposed an F1 depression detection time scale. This metric was applied to various models from the eRisk 2017 joint missions to detect depression. The best value for F1, 71.2, was obtained through a GRU sequence model with words denoting DepWord and features specific to UMLS (MetaMap). The authors of [15] this study used data from social media platforms to uncover methods for the early detection of MDDs using machine learning algorithms. An integrated analysis of the data set on what users write on social media sites, such as texts. The authors proposed two approaches that rely on single and dual machine learning. The first method uses a single combined RF classifier with two threshold functions, while the second method uses two separate RF classifiers, one for detecting depressed subjects and the other for detecting non-depressed users. Features are identified by textual similarities and writing. The proposed dual model performs better than the proposed single model and can improve existing detection models by close to 10%. The authors of [16] Assessed the effect of psychological language on machine learning algorithms that categorized depressed social media users based on their posts, incorporated psycholinguistic features into a rule-based model, and then evaluated its effect on classification, along with three other Workbooks. The results excel in the Reddit Depression Diagnostic Database. Best results were obtained when using TF-IDF with a passive-aggressive classifier with an accuracy of 96%, Precision of 82%, recall 64%, and F1 72%. This is because the passive-aggressive classifier is a sequentially learning classifier. Multinomial Naive Bayes the accuracy 94% Precision 61%, recall 47%, and F1 53%. The analysis of social media posts can provide useful means for the ability to understand the users' mental health status and early detection of depression. Preliminary results were presented on the use of written psycholinguistic features as an improvement in the rating approach for depressed social media users. The authors of [17] proposed a method capable of mapping data on social media to screen for depression. Using the proposed system, a social moderator patient portal (SMPP) application was developed, the task of this application is to detect signs of depression in Facebook platform users, and this application relies on classification algorithms for machine learning. A dataset of 4,350 Facebook users who were classified with depression was examined using the Center for Epidemiological Studies of Depression Scale (CES-D). Then, a set of features that can reveal users with depression were identified. This model may be useful for psychiatrists in diagnosing depression and analyzing its behavior. The authors of [18] proposed a model based on data analysis to screen people with depression. Data was collected from user posts on Twitter and Facebook. The standard method for screening for depression was the structured or semi-structured interview method (SDI). Machine learning was used to process data collected from SNS users. NLP Classified Natural Language Processing, using the Naïve Bayes algorithm, to more efficiently detect depressed patients. Results were obtained from a machine learning model using Nave Bayes; Accuracy is 74%, Precision is 100% and Recall 60%.

3. Proposed Method

A. Building Datasets

We collected **1439 Arab user accounts** who are posting specific words such as:" أنا كئيب/ة, أنا مكتئب/ة"I am depressed", and أعاني الاكتئاب"I suffer from depression". We collected Arabic Tweets using the Apify website, from January 1, 2019, to October 31, 2020, because we started collecting data in November 2020.

B. Data Preprocessing

To arrive at clean datasets, we followed these points:

Remove Arabic diacritics, these include Dammah, Fatahah, Kasrah, Maddah, Shaddah, Tanwin, Sukun.

Remove Word elongation (kashida), such as الاحزااااان, الموووووت

Remove Punctuation, numbers, and emojis

Remove Non-Arabic characters

Remove Multiple whitespace and empty lines

Remove Users names and URLs

Remove duplicate tweets

Convert different forms of characters to the same form:

ا → أ آ ا إ

ي → ى ي

ء → ئ ؤ ء

ك → گ

ه → ه ة

C. Classification Depressed Arab Twitter Users.

The Arabic Tweets for each user were divided for each user into two consecutive weeks. This helps determine the psychological state of the user whether he suffers from depression, as he must suffer from three symptoms of depression, one of these symptoms must be depressed mood, for at least two consecutive weeks in three months, according to the APA.

D. Binary Classification Machine Learning

This section investigated various Machine Learning algorithms in binary classification. Adaboost, Naïve Bayes NB, RF, and DT, they are all trained on the training split 75% and evaluated on the testing split 25% with different Features Extraction: BOW, TF-IDF, N-gram, and Global Vectors for Word. For all the experiments presented, we use Google Colab with a GPU hardware accelerator.

E. Determine Depressed Tweets and the Symptoms using My Depression

To facilitate the task of categorizing tweets, we built a dictionary of words for depression in Arabic similar to the method of [20] [21]. We used the My Depression dictionary to classify the tweets of Arab Twitter users. For the tweet to be classified as a depression tweet, it must contain at least one keyword used by an Arab depression patient, as the depression tweet may contain more than one keyword indicating depression. We got the data as shown in Table 1 which shows information about our datasets.

TABLE 1.

DATASET INFORMATION

Dataset Information	Number
Total number of Twitter users Accounts	1439
Total number of Depressed users	600
Total number of Tweets	16581
Depression Tweets	8458
Non- Depression Tweets	8123

4. Experimental Results

We have done several experiments with machine learning algorithms; we got different results in our experiments. Table 2 presents the results we obtained from this experiment. Table 3 presents the results we obtained from this experiment. Table 4 presents the results we obtained from this experiment. Table 5 presents the results we obtained from this experiment. Table 6 presents the results we obtained from this experiment.

TABLE 2

BINARY CLASSIFICATION RESULTS BASE ON BAG OF WORD

Machine Learning	Accuracy	Precision	Recall	F1-Score
Mutational Naïve Bayes	78%	77%	91%	73%

Random Forest	83%	80%	90%	83%
Decision Tree	**84%**	**81%**	**91%**	**84%**
AdaBoost	71%	86%	51%	70%

TABLE 3

BINARY CLASSIFICATION RESULTS BASE ON TERM FREQUENCY- INVERSE DOCUMENT FREQUENCY

Machine Learning	Accuracy	Precision	Recall	F1-Score
Mutational Naïve Bayes	**86%**	75%	89%	79%
Random Forest	84%	84%	84%	84%
Decision Tree	83%	82%	86%	83%
AdaBoost	71%	86%	53%	70%

TABLE 4

BINARY CLASSIFICATION RESULTS BASE ON N-GRAM IN TERM FREQUENCY- INVERSE DOCUMENT FREQUENCY LEVEL

Machine Learning	Accuracy	Precision	Recall	F1-Score
Mutational Naïve Bayes	**60%**	57%	91%	54%
Random Forest	**60%**	81%	29%	56%
Decision Tree	59%	82%	26%	54%
AdaBoost	54%	40%	100%	53%

TABLE 5

BINARY CLASSIFICATION RESULTS BASE ON N-GRAM IN CHARACTER LEVEL IN TERM FREQUENCY- INVERSE DOCUMENT FREQUENCY LEVEL

Machine Learning	Accuracy	Precision	Recall	F1-Score
Mutational Naïve Bayes	**78%**	77%	84%	78%
Random Forest	73%	78%	68%	73%
Decision Tree	68%	69%	70%	68%
AdaBoost	72%	78%	65%	72%

TABLE 6.

BINARY CLASSIFICATION RESULTS BASE ON GLOBAL VECTORS FOR WORD REPRESENTATION

Machine Learning	Accuracy	Precision	Recall	F1-Score
Mutational Naïve Bayes	50%	52%	65%	50%
Random Forest	**78%**	78%	80%	78%
Decision Tree	75%	76%	76%	75%
AdaBoost	**78%**	72%	94%	77%

5. Conclusions

We detected depression among Arab Twitter users, through their tweets; Of the 1,439 users, 600 experienced depressions. To reach this result, we first built a dictionary containing the most frequently used Arabic words among Arab depressed patients to express their suffering from the nine symptoms of depression, according to the American Psychiatric Association. We used the My Depression dictionary to help us classify 22568 Arabic tweets collected from 1,439 Arab users, the rating was confirmed by the three psychologists. After eliminating the repetition of tweets, we obtained the My Depression database of Arabic tweets, which includes 16,581 Arabic tweets, categorized into depressing tweets and specifying the symptoms they contain or a normal tweet, for the My Depression database of 1439 Arab users on Twitter, classified Depressed users or not. Our experiments showed that

Mutational Naïve Bayes with TF-IDF had the highest accuracy of 86% when rating tweets. Caring for the mental health of people is very important, as some measures must be taken to maintain the mental health of people in the early stages of infection.

Acknowledgemnt

We would like to express our great appreciation to Dr.Ammar Alsaleh, Dr.Yousef Mselm, and Dr.Teeser Showash for their valuable recommendations and verifications of the terms used in the My Depression lexicon, and the classification tweets by this lexicon. We would also like to thank the Murad Center for Psychological Counseling, Kalima Center for Cognitive Behavioral Sciences, and My Heaven Clinic, for help me to communicate with patients

Author Contributions

All authors contributed equally to this work.

Funding

This research was conducted without external funding support.

Ethical approval

This article does not contain any studies with human participants or animals performed by any of the authors.

Conflicts of Interest

The authors declare that there is no conflict of interest in the research.

Institutional Review Board Statement

Not applicable.

Informed Consent Statement

Not applicable.

Data Availability Statement

Not applicable.

References

[1] Murray CJ, Vos T, Lozano R, et al. "Disability-adjusted life years (DALYs) for 291 diseases and injuries in 21 regions, 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010". Lancet 2012 Dec 15;380(9859):2197-2223.

[2] M. Deady, D. A. Johnston, N. Glozier, et al., "Smartphone application for preventing Depression: study protocol for a workplace randomized controlled trial," BMJ Open, vol. 8, no. 7, 2018.

[3] D. Maupomé and M. J. Meurs, "Using topic extraction on social media content for the early detection of depression," CEUR Workshop Proc., vol. 2125, pp. 2–6, 2018.

[4] C. S. Dewa, N. Chau, and S. Dermer, "Examining the comparative incidence and costs of physical and mental health-related disabilities in an employed population," J. Occup. Environ. Med., vol. 52, no. 7, pp. 758–762, 2010, doi: 10.1097/JOM.0b013e3181e8cfb5.

[5] E. Granger, "Wheidima Carneiro de Melo Abdenour Hadid Center for Machine Vision and Signal Analysis (CMVS), University of Oulu , Finland de technologie sup ´ erieure , Montreal , Canada," 2019 IEEE Int. Conf. Image Process., pp. 4544–4548, 2019.

[6] N. S. Alghamdi, H. A. Hosni Mahmoud, A. Abraham, S. A. Alanazi, and L. Garcia-Hernandez, "Predicting depression symptoms in an arabic psychological forum," IEEE Access, vol. 8, pp. 57317–57334, 2020.

[7] H. Alhazmi and M. Alharbi, "Emotion analysis of arabic tweets during COVID-19 pandemic in Saudi Arabia," Int. J. Adv. Comput. Sci. Appl., vol. 11, no. 10, pp. 619–625, 2020, doi: 10.14569/IJACSA.2020.0111077.

[8] K. S. Al-Khalifa et al., "Pattern of dental needs and advice on Twitter during the COVID-19 pandemic in Saudi Arabia," BMC Oral Health, vol. 21, no. 1, pp. 1–7, 2021, doi: 10.1186/s12903-021-01825-4.

[9]M. R. Islam, M. A. Kabir, A. Ahmed, A. R. M. Kamal, H. Wang, and A. Ulhaq, “Depression detection from social network data using machine learning techniques,” Heal. Inf. Sci. Syst., vol. 6, no. 1, pp. 1–12, 2018, doi: 10.1007/s13755-018-0046-0.
[10] S. G. Burdisso, M. Errecalde, and M. Montes-y-Gómez, “A text classification framework for simple and effective early depression detection over social media streams,” Expert Syst. Appl., vol. 133, pp. 182–197, 2019, doi: 10.1016/j.eswa.2019.05.023.
[11] M. Stankevich, V. Isakov, D. Devyatkin, and I. Smirnov, “Feature engineering for depression detection in social media,” ICPRAM 2018 - Proc. 7th Int. Conf. Pattern Recognit. Appl. Methods, vol. 2018-January, no. Icpram, pp. 426–431, 2018, doi: 10.5220/0006598604260431.
[12] F. Benamara, V. Moriceau, J. Mothe, F. Ramiandrisoa, and Z. He, “Automatic detection of depressive users in social media,” Conf. en Rech. d’Informations Appl. CORIA 2018 - 15th French Inf. Retr. Conf. Proc., 2020.
[13] C. Lin et al., “SenseMood: Depression detection on social media,” ICMR 2020 - Proc. 2020 Int. Conf. Multimed. Retr., pp. 407–411, 2020, doi: 10.1145/3372278.3391932.
[14] F. Sadeque, D. Xu, and S. Bethard, “Measuring the latency of depression detection in social media,” WSDM 2018 - Proc. 11th ACM Int. Conf. Web Search Data Min., vol. 2018-Febuary, pp. 495–503, 2018, doi: 10.1145/3159652.3159725.
[15] F. Cacheda, D. Fernandez, F. J. Novoa, and V. Carneiro, “Early detection of depression: Social network analysis and random forest techniques,” J. Med. Internet Res., vol. 21, no. 6, 2019, doi: 10.2196/12554.
[16] A. T. B, R. Antunes, and J. Lu, “from Psycholinguistic Patterns in Social Media Texts,” pp. 402–409, 2020, doi: 10.1007/978-3-030-45442-5.
[17] J. Hussain et al., “Exploring the dominant features of social media for depression detection,” J. Inf. Sci., vol. 46, no. 6, pp. 739–759, 2020, doi: 10.1177/0165551519860469.
[18] N. Al Asad, M. A. Mahmud Pranto, S. Afreen, and M. M. Islam, “Depression Detection by Analyzing Social Media Posts of User,” 2019 IEEE Int. Conf. Signal Process. Information, Commun. Syst. SPICSCON 2019, pp. 13–17, 2019, doi: 10.1109/SPICSCON48833.2019.9065101.
[19] E. Alomari, I. Katib, A. Albeshri, and R. Mehmood, “Covid-19: Detecting government pandemic measures and public concerns from twitter arabic data using distributed machine learning,” Int. J. Environ. Res. Public Health, vol. 18, no. 1, pp. 1–36, 2021, doi: 10.3390/ijerph18010282.
[20] Mohamed Abdel-Basset; Nour Moustafa; Hossam Hawash, "Deep Neural Networks," in Deep Learning Approaches for Security Threats in IoT Environments , IEEE, 2023, pp.27-54, doi: 10.1002/9781119884170.ch2.
[21] B. Bataineh, R. Duwairi, and M. Abdullah, “ArDep: An Arabic lexicon for detecting depression,” ACM Int. Conf. Proceeding Ser., pp. 146–151, 2019, doi: 10.1145/3369114.3369136.