

Social cyber-criminal, towards automatic real time recognition of malicious posts on Twitter

Yasser Ibrahim Abdelmonem Ali, Mohammed Abdel Razeq, Nasser El-Sherbeny

Mathematic and Computer Science Department, Faculty of Science, Azhar University, Cairo, Egypt

Article Info

Article history:

Received Jun 29, 2021

Revised Dec 1, 2021

Accepted Dec 9, 2021

Keywords:

Cyber-criminal

Dominant meaning

Malicious posts on Twitter

Term frequency

TF inverse document frequency

ABSTRACT

Easy access to the internet throughout the world has fully reformed the usage of social communication such as Facebook, Twitter, Linked In which are becoming a part of our life. Accordingly, cybercrime has become a vital problem, especially in developing countries. The dissemination of information with no risk of being discovered and fetched leads to an increase in cyber-criminal. Meanwhile, the huge amount of data continuously produced from Twitter made the discovery process of cyber-criminals is a tough assignment. This research will contribute in determined on the build the comparable vectors for (positive and negative classes) and then the classify incoming tweets to predicate his class (positive or negative). The proposed routines starting with the construct super comparable vectors (SCV) (positive and negative vectors), and the construct vector for the incoming tweet, and then calculate similarities with both SCV and compare calculated similarities to predicate class of incoming tweet. In this research, we used some common techniques for calculating the weight of terms in tweets to construct SCV. To ensure the successful operation of the proposed system, we performed a pilot analysis on a real example of an examination. Research Improves precision, recall, and F1 values by 87%, 59%, 69.99%, respectively.

This is an open access article under the [CC BY-SA](#) license.



Corresponding Author:

Yasser Ibrahim Abdelmonem Ali

Mathematic and Computer Science Department, Faculty of Science, Azhar university

Nasr City, 11884, Cairo, Egypt

Email: yasser.ibrahim@azhar.edu.eg

1. INTRODUCTION

The aim of cybersecurity is to secure processes, hardware, software, and information components of computer systems during online session from stealing, harm, interruption, and illegal use. In present era, social network is the part and parcel in everyday life. Easy access to internet through the world has fully reformed the usage of social communication such as Facebook, Twitter, Linked-In which are becoming a part of our life. Accordingly, cybercrime has become a vital problem, especially in a developing country. One of these cybercrimes is cybersecurity information which contains a variety of unofficial sources, such as social media platforms, chat rooms, blogs, and developer forums. This type kind of crime provides information about security vulnerabilities, threats, and attacks. In order to secure this information, we need real-time intelligence. On cybersecurity threats and vulnerabilities. The proposed approach will carefully examine a number of studies that have suggested models for future cybersecurity threats and are based on time series and moving averages.

There are over 150 million users write over 500 million tweets per day for the year 2019 [1]. considered text categorization is necessary to get relevant info from such a huge collection tweet and convert

the entire large-scale information into a small size subcategory. It can be controlled for analysis. As a result, attempts have been made to classify the text of the tweet into several areas to discover information on a particular topic. Typical examples are a geolocation user prediction to categorize tweets by geo-political location [2].

Also, Twitter has 330 million monthly users worldwide [3]. Considered as Turkey fifth country with 9 million approximate consider active from January 2019 [4]. Political predicting affiliation by rating tweets related to politics [5], and predicting crime categorizes Twitter posts based on emotion [6]. However, the rating of difficulty in twitter posts in nature because of the length of the twitter posts is limited, i.e., under 280 characters, and varied types of users engage in informal tweet writing [7].

Dionísio *et al.* [8]. presents new methodology to improve the search results. His problems are cyberattacking is common and big issue because there are many flows of data produce social media such as twitter, so most organization resort security information. These systems rely on synchronizing the latest, corrections, as well as threats presented by threat feeds. He tried to produce a new architecture separated in three stages. Collect data from Twitter by using Twitter application programming interface (API), apply filters and normalizes tweets on specific format. Apply binary classifier labels to classify fetched tweets into relevant or irrelevant, relevant It is possible that Tweets contain valuable information about an asset and are not related otherwise. Finally, the named entity recognition (NER) network processes the relevant tweets during the information extraction stage. For example, the information gathered could be utilized to send out a security alert. His approach results, the proposed pipeline achieves a true rate positive of 94% on average 91% rate negative for rating task and F1- score 92% average for task entity named recognition, via 3 infrastructures topical research.

Sabotke *et al.* [9] shows a vulnerability detection service based on Twitter suggested using the supported vector machine (SVM) classification. It could be able to be exploited in a real-life situation a fascinating feature The goal of this research is to look into aggressive interference as a way to trick the classifier. These recruiters may provide the information in a more organized and formal way at their Twitter headquarters. It takes advantage of some of the characteristics of these essays, such as their grammatical relationships and relationships. In conjunction with this proposal, Zhou *et al.* [10] uses the architecture of the NER to extract indicators of compromise (IoC) according to cybersecurity reports in contrast to the design. Like the one bearing her name. Badjatiya *et al.* [11] educate other customers' opinions of copyright on SemEval 2015 [12] Twitter. The outdoor equipment introduced to the challenge ranked first in both missions.

Badjatiya *et al.* [11] presents an example which is not assertive as, but is not limited to, the path of learning for the advice of the fossil idiot who loves you to hate. The authors report that learning techniques play an important role in eating. Regarding the applications of learning applications, applications, applications, applications, and applications. Wagner *et al.* [13] long short-term memory (LSTM) architecture implemented sequencing to suspend medical entities for public health monitoring. The architecture offered is superior to the previous case. As far as the future is concerned, in relation to the infrastructure proposed by Lample *et al.* [14]. This is how we adopted our NER approach [15].

Aslan *et al.* [16] works on Twitter as example, they use machine learning techniques, which investigate if accounts social media were importance in terms of cybersecurity. They used the Python programming language Crawler Twitter API to fix their dataset for use in their research. Alves *et al.* [17] they introduced many of tested the quantitative evaluation with everyone in mind Twitter posts from 80 accounts over 8 months (a total of 195,000 tweets), it shows that their methodology came at the right time and successfully find the most security-related Tweets related to the IT infrastructure example (value of positive measure greater than 90%), incorrectly choose. A little number of related tweets (value of false positive measure less than 10%). Duarte *et al.* [18] introducing a new methodology developed for text analysis in English or other widely spoken languages, such as Portuguese using big data [19]. Javed *et al.* [20] introduces what because of this type of shortcut, we can block such harmful websites by clicking on them as locker websites. Khandpur *et al.* [21] shows a new study on detecting cyber-attacks by analysing Twitter data. Sohime *et al.* [22] shows if the investigation period is long enough, cyber attackers have an advantage, which is a difficulty for a security analyst. Keep up with the latest risks.

We aim in this research to produce a mixed methodology to overcome the above problem, we can summarize as the following:

- Collect binary classification dataset.
- Collect full description from Twitter application programming interface (API) for each Tweet, separation dataset into positive and negative pockets.
- Construct super comparable vector (SCV) for each pocket to generate vector positive, and vector negative.
- Construct vector from entire Tweet.

- Calculate similarities between positive and negative vectors.
- Compare between results.

The rest of this study is organized as follows: section 2 shows the methodology and the proposed algorithm with how to enhance binary classification using Word2Vec model. Section 3, presents the experiment and show the improved happened after applying our proposed algorithm. Conclusion and future projects are presented in section 4.

2. METHODOLOGY

Our methodology focuses on how to improve the accuracy results of the classification method. Therefore, it uses twitter social network and binary classified dataset to which will be described in the later sections. We will take a tour in our proposed algorithm which research try implement to answer suggest questions. Also, we will discuss the suggestion techniques which research will be implement.

2.1. Classification algorithm

This section tries to overcome the following challenges which are the main part of this methodology:

- How to enhance binary classification using Word2Vec Model?
- How to construct super comparable vector SCV?

In Figure 1, shows the proposed algorithm stages. First this algorithm collects meta-data (full tweet text and tags) about tweets by using Twitter API service (this service provided by Twitter itself) for working dataset [23].

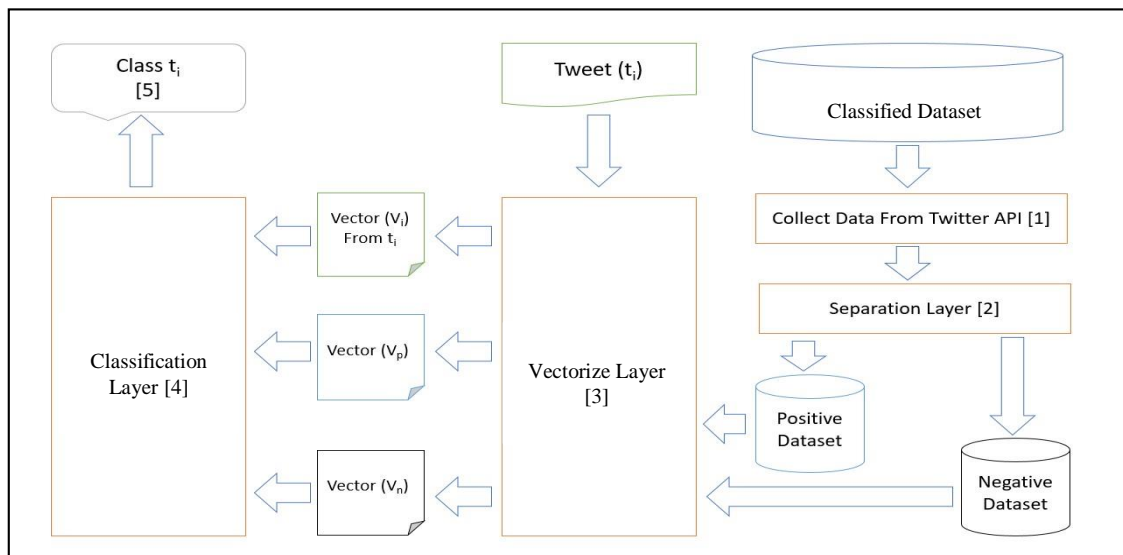


Figure 1. Algorithm stages

This dataset consists of three classified datasets D_1, D_2, D_3 . Each raw dataset contains tweet-id, classified status only, so we must collect other data from Twitter API. The next step is to separate dataset into two pockets one for positive tweets and other for negative pockets to prepare to the next step for each pocket (positive and negative) we will construct super vector comparable, we will use some different methodologies to construct super comparable vector we will describe its later, now we will be ready to compare each dataset with two super vector comparable (positive and negative) by using output of the following pockets:

$$V_i \text{ for tweet } i, V_p \text{ for super positive vector and } V_n \text{ for super negative vector}$$

As showing in vectorize layer to prepare the next step by calculating similarity between previous vectors we have new class for each tweet in D_i , we will compare a new class and actual class to calculate efficiency for each methodology. We will show our architecture in detail in the following sections. But now we can summarize our architecture as the following algorithm:

- Collect binary classified dataset.
- Collect full description from Twitter API for each Tweet.
- Separation dataset into positive and negative pockets.
- Construct super comparable vector (SCV) for each pocket to generate V_p (for positive pocket) and V_n (for negative pocket).
- Construct vector V_i for tweet t_i .
- Calculate S_{pi} and S_{ni} (S_{pi} similarity between V_i and V_p & S_{ni} similarity between V_i and V_n).
- Compare between S_{pi} and S_{ni} to predicate class of t_i .

2.2. Classification method

We will use in our proposal the following technique in vectorize and classification layers:

- Term frequency technique (TF).
- Term frequency inverse document frequency technique (TF-IDF).
- Dominant meaning technique.

We will add extra step in each above technique, by remove duplicate dimension from both super vector comparable positive and negative, this addition will give a change in results significantly.

For construct super vectors comparable for positive and negative pockets by merging all tweets inside each pocket (positive and negative) and calculate weight of each term by some technique, will describe later. This process will produce two vectors positive V_p and negative V_n predication. We will use same technique to construct super comparable to construct V_i for tweet t_i , to prepare next step. Calculate S_{pi} and S_{ni} similarity between t_i and both V_p and V_n by user Euclidean distance equation as the following:

$$S_{pi} = \text{similarity}(t_i, V_p), S_{ni} = \text{similarity}(t_i, V_n),$$

where $i=1,2,3,\dots,k$ and k is dataset size.

$$\text{similarity}(\vec{x}, \vec{y}) = \frac{\vec{x} \cdot \vec{y}}{|\vec{x}| |\vec{y}|}$$

Depend on S_{pi} and S_{ni} we will get the class of t_i to calculate efficacy. Now let us take some brief about the techniques which using in this paper:

- Term frequency (TF) technique by calculate ratio of number of times term occurrence and total number of terms in document.

$$TF(t) = \frac{\text{Number of times term } t \text{ appears in a tweet}}{\text{Total number of terms in a tweet}}$$

- Term frequency inverse document frequency (TF-IDF) technique like TF with cancellation all terms which occurred in all documents.

$$TF - IDF(t) = TF(t) \cdot \log\left(\frac{\text{Number of tweet term occurred}}{\text{Total number of tweets}}\right)$$

- Dominant meaning method, which is well-known “The set of keywords that fit the intended meaning of the target word” [24]. The question is viewed as a goal meaning, as well as certain words that fall within that meaning’s scope. It freezes the intended meaning, known as the keyword, then adds or removes slave words that explain the meaning [25], [26]. To constructing dominant meaning hierarchical, assume we have a dataset made up of n , that $C = \{C_i\}_{i=1}^n$ is, for each C_i from C represented by a collection of documents which trying to describe concept C_i , suppose that the collection consists of m documents, that is $\{C_i = D_j^i, j = 1, 2, 3, \dots, m_i\}$, each document in this collection consists of a set of k words or $D_j^i = w_{lj}^i, l = 1, 2, 3, \dots, k_j$ terms. The w_{lj}^i s represent word repetition w_{lj}^i occurs in document D_j^i which slave’s words of concept C_i . This frequency is calculated as the number of times that the w_l occurs in the D_j^i . The following steps represent the process to choose topN words which can indicate the dominant meaning of concept C_i , suppose that word w_c^i symbolizes concept C_i .
 - a) Compute each w_{lj}^i for all i, j .
 - b) Suppose that $C_{i,j}$ is the frequency of concept C_i , which appears in document D_j^i where $j = 1, 2, 3, \dots, m_i$.
 - c) Calculate maximum of C_{ji} for all $i, F_c^i = \text{Max}\{C_{ji}\}_{j=1}^{m_i}$.
 - d) Calculate maximum value of w_{lj}^i for all $l, j, F_{wj}^i = \text{Max}\{w_{lj}^i\}_{j=1}^{m_i}$.
 - e) Choose P_c^i , which satisfies $0 \leq F_{wj}^i \leq F_c^i$.
 - f) Finally, consider the dominant meaning probability:

$$P_{ij} = P_{ij}(w_j|C_i) = \frac{1}{m_i} \sum_{l=1}^{m_i} \frac{w_{lj}^i}{F_c^i}, i = 1,2, \dots, n \text{ and } j = 1,2, \dots, k_j$$

Now for each concept C_i , we rank the terms of collection $\{P_{i1}, P_{i2}, \dots, P_{im}\}$ in decreasing order. As a result, the Dominant Meaning of the concept C_i can be represented by the set of words that corresponds to the set $\{P_{i1}, P_{i2}, \dots, P_{iN}\}$; that is:

$$W_i = \{w_{i1}, w_{i2}, \dots, w_{iN}\}$$

The set W_i is representing more intended meaning for concept C_i .

3. RESULTS AND DISCUSSION

To try answer on questions How to enhance binary classification using Word2Vec Model? And how to construct comparable vector? We conduct experiment on collection tweets of binary classified. We will talk in detail about dataset used to calculate efficiency. Also, we will take a tour and discuss in detail how could answer about suggested questions, and will try answer question "what's our recommendation technique which gives more accurncy results?" later last section.

3.1. Dataset

The datasets consist of three binary classified datasets (D_1, D_2, D_3) [23], these datasets consist of 4614, 2127 and 1081 tweets, respectively. D_i represented by coma separate value file format (csv file) with basic data (tweet-id, classification class). Table 1 shows describe the above training datasets. We will use D_i where $i = 1, 2, 3$ to construct super vectors comparable and we will compare it's with all other datasets excluding itself (D_i) and we will describe the results in the next section.

Table 1. Number of tweets for both pockets

Dataset	D_1	D_2	D_3
Positive	2391	634	453
Negative	2223	1493	628
Total	4614	2127	1081

3.2. Exprimatal results

We will partition this section into time reality and proposed system efficiency, via discussion the results of time and efficiency. As we see the results show the purpose of using reduction technique as extra step, significantly more improves. In this research, the traditional method to calculate a weight of any term in specific context is term frequency (TF) technique. As we see the results show the purpose of using reduction technique as extra step, significantly more improves.

- TF V.S. TF + reduction Table 2 shows the results of precision and recall results for all techniques. The result of first technique TF pure (without adding any extra steps) gives 80% and 91% for precision and recall respectively, also TF extra (TF with reduction duplicate dimensions from both super vectors comparable) gives 87% and 60% for precision and recall respectively as shown in Table 2 and Table 3. The improvement in precision and recall for both TF & TF + reduction techniques is taken around 7% and 31% respectively as shown in Table 4 and Table 5. In contrast, the are improvement in time performance for both techniques as shown in Figure 2. Figure 2 shows time performance for TF V.V. TF + reduction which tell us TF + reduction more time performance than TF.
- Dominant meaning (DM) V.S. DM + reduction Table 2 shows the results of precision and recall results for all techniques. The result of first technique DM pure gives 80% and 91% for precision and recall respectively, also DM + Reduction gives 87% and 60% for precision and recall respectively as shown in Table 2 and Table 3. The improvement in precision and recall for both DM & DM + Reduction techniques is taken around 7% and 31% respectively as shown in Table 4 and Table 5. In contrast, the are improvement in time performance for both techniques as shown in Figure 3. Figure 3 shows time performance for DM v.s. DM + reduction which tell us DM + reduction more time performance than DM.
- TF-IDF V.S. TF-IDF + reduction Table 2 shows the results of precision and recall results for all techniques. The result of first technique TF-IDF pure (without adding any extra steps) gives 83% and 87% for precision and recall respectively, also TF-IDF reduction gives 91% and 59% for precision and recall respectively as shown in Table 2 and Table 3. The improvement in precision and recall for both TF-IDF & TF-IDF + reduction techniques is taken around 4% and 32% respectively as shown in Table 4 and Table 5. In contrast, the are improvement in time performance for both techniques as shown in

Figure 4. Figure 4 shows time performance for TF-IDF v.s. TF-IDF + Reduction which tell us TF-IDF + reduction more time performance than TF-IDF.

Table 2. Precision, recall, and F1 summary for all techniques

Technique	Precision	Recall	F1
TF	80.00%	91.00%	82.23%
TF Reduction	87.00%	60.00%	70.34%
DM	80.00%	91.00%	82.23%
DM Reduction	87.00%	60.00%	70.34%
TF-IDF	83.00%	91.00%	86.36%
TF-IDF Reduction	87.00%	59.00%	69.99%

Table 3. Precision, recall, and F1 summary for all techniques

Technique	Dataset	D ₁			D ₂			D ₃		
		Precision	Recall	F1	Precision	Recall	F1	Precision	Recall	F1
TF	D1	-	-	-	64.00%	76.00%	69.49%	76.00%	88.00%	81.56%
	D2	75.00%	91.00%	82.23%	-	-	-	74.00%	91.00%	81.62%
	D3	80.00%	81.00%	80.50%	68.00%	72.00%	69.94%	-	-	-
TF Reduction	D1	-	-	-	76.00%	50.00%	60.32%	80.00%	49.00%	60.78%
	D2	87.00%	39.00%	53.86%	-	-	-	82.00%	36.00%	50.03%
	D3	85.00%	60.00%	70.34%	76.00%	53.00%	62.45%	-	-	-
DM	D1	-	-	-	64.00%	76.00%	69.49%	76.00%	88.00%	81.56%
	D2	75.00%	91.00%	82.23%	-	-	-	74.00%	91.00%	81.62%
	D3	80.00%	81.00%	80.50%	68.00%	72.00%	69.94%	-	-	-
DM Reduction	D1	-	-	-	76.00%	50.00%	60.32%	80.00%	49.00%	60.78%
	D2	87.00%	39.00%	53.86%	-	-	-	82.00%	36.00%	50.03%
	D3	85.00%	60.00%	70.34%	76.00%	53.00%	62.45%	-	-	-
TF-IDF	D1	-	-	-	72.00%	81.00%	76.24%	83.00%	90.00%	86.36%
	D2	78.00%	91.00%	84.00%	-	-	-	77.00%	91.00%	83.42%
	D3	81.00%	83.00%	81.99%	70.00%	73.00%	71.47%	-	-	-
TF-IDF Reduction	D1	-	-	-	76.00%	50.00%	60.32%	80.00%	49.00%	60.78%
	D2	87.00%	39.00%	53.86%	-	-	-	82.00%	35.00%	49.06%
	D3	86.00%	59.00%	69.99%	78.00%	52.00%	62.40%	-	-	-

Table 4. The improvements values in precision

Technique	Precision	Improves
TF	80.00%	7.00%
TF Reduction	87.00%	-
DM	80.00%	7.00%
DM Reduction	87.00%	-
TF-IDF	83.00%	4.00%
TF-IDF Reduction	87.00%	-

Table 5. The improvements values in recall

Technique	Recall	Improves
TF	91.00%	31.00%
TF Reduction	60.00%	-
DM	91.00%	31.00%
DM Reduction	60.00%	-
TF-IDF	91.00%	32.00%
TF-IDF Reduction	59.00%	-

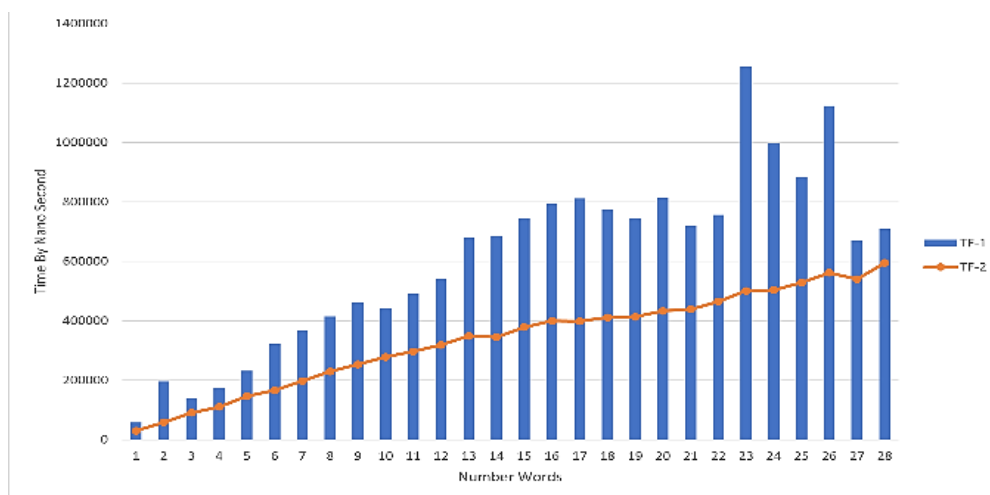


Figure 2. TF 1 and 2 time comparison chart

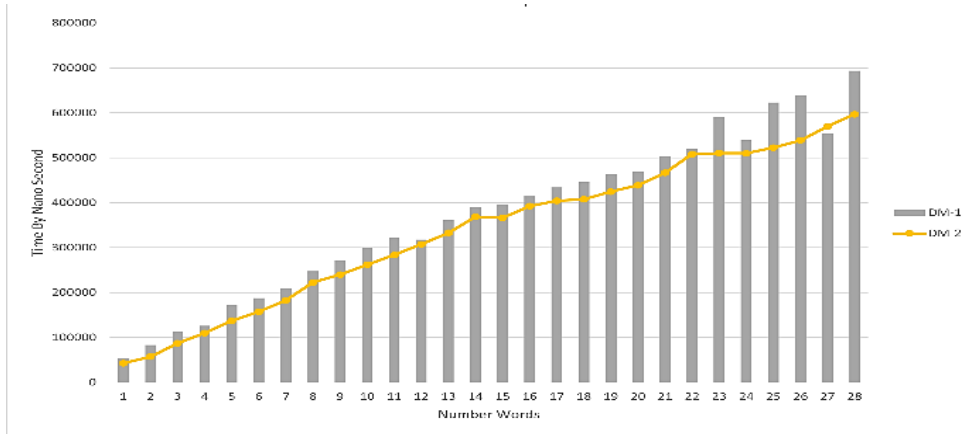


Figure 3. DM 1 and 2 time comparison chart

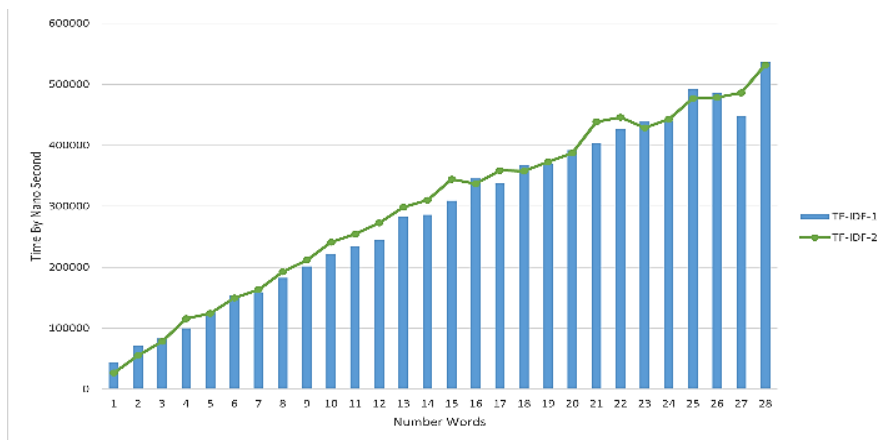


Figure 4. TF-IDF 1 and 2 time comparison chart

As shown in Figure 5 and Table 3, the improvement in time is increased quickly in the TF and TF + reduction technique, DM and TF-IDF pure more time efficient than TF pure, but TF reduction close from other techniques. It is interesting to note that the highest improvement in TD-IDF in both pure and reduction from time performance. Finally, above discussion and measurement values which lead us to our proposed methodology enhance precision 87% and recall 59%, in future we will work to enhance above vales by applying or mix another technique.

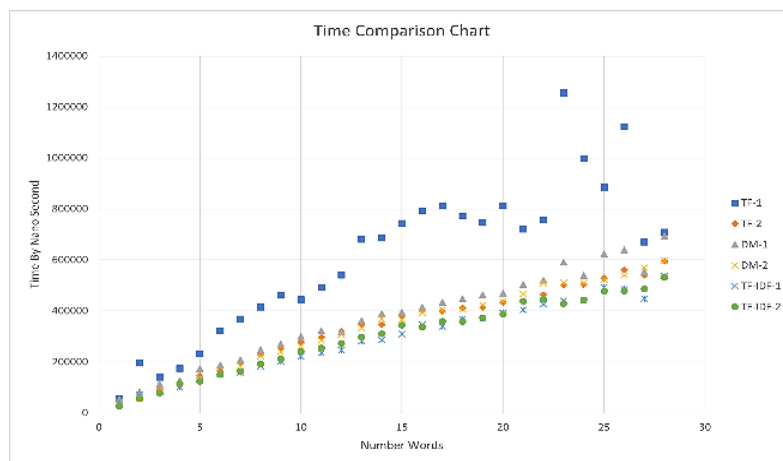


Figure 5. All techniques time comparison chart




4. CONCLUSION

The change in degradation for each year over the past four decades is gradually increased. In this research we used some common techniques for calculate a weight of terms in tweet to construct comparable vectors and compute similarity between these vectors and input tweet to predicate his class (positive or negative). The experimental results tell us the best improvement time in dominant meaning (DM) and term frequency inverse document frequency (TD-IDF) more than term frequency (TF) from excremental results.




REFERENCES

- [1] S. Aslam, Twitter by Numbers: Stats, Demographics & Fun Facts, Omnicore, 2021. Accessed: Jul. 14, 2020. [Online]. Available: <https://www.omnicoreagency.com/twitter-statistics/2020>
- [2] B. Han, P. Cook, and T. Baldwin, "Text-based twitter user geolocation prediction," *Journal of Artificial Intelligence Research*, vol. 49, pp. 451–500, 2014, doi: 10.1613/jair.4200.
- [3] H. T. Phan, V. C. Tran, N. T. Nguyen, and D. Hwang, "Improving the Performance of Sentiment Analysis of Tweets Containing Fuzzy Sentiment Using the Feature Ensemble Model," in *IEEE Access*, vol. 8, pp. 14630-14641, 2020, doi: 10.1109/ACCESS.2019.2963702.
- [4] A. Okay, P. A. Gole, and A. Okay, "Turkish and slovenian health ministries' use of twitter: A comparative analysis," *Corporate Communications: An International Journal*, vol. 26, no. 1, pp. 176-191, 2020, doi: 10.1108/CCIJ-01-2020-0019.
- [5] M. D. Conover, B. Goncalves, J. Ratkiewicz, A. Flammini, and F. Menczer, "Predicting the Political Alignment of Twitter Users," *2011 IEEE Third International Conference on Privacy, Security, Risk and Trust and 2011 IEEE Third International Conference on Social Computing*, 2011, pp. 192-199, doi: 10.1109/PASSAT/SocialCom.2011.34.
- [6] X. Wang, M. S. Gerber, and D. E. Brown, "Automatic crime prediction using events extracted from twitter posts," in *Int. Conf. Social Computing, Behavioral-Cultural Modeling, and Prediction*, 2012, pp. 231–238, doi: 10.1007/978-3-642-29047-3_28.
- [7] D. Ramage, S. Dumais, and D. Liebling, "Characterizing microblogs with topic models," *Proceedings of the Fourth International AAAI Conference on Weblogs and Social Media*, 2010.
- [8] N. Dionísio, F. Alves, P. M. Ferreira, and A. Bessani, "Cyberthreat Detection from Twitter using Deep Neural Networks," *2019 International Joint Conference on Neural Networks (IJCNN)*, 2019, pp. 1-8, doi: 10.1109/IJCNN.2019.8852475.
- [9] C. Sabotke, O. Suci, and T. Dumitras, "Vulnerability disclosure in the age of social media: Exploiting twitter for predicting real-world exploits," *Proceedings of the 24th USENIX Conference on Security Symposium*, 2015, pp. 1041–1056.
- [10] S. Zhou, Z. Long, L. Tan, and H. Guo, "Automatic identification of indicators of compromise using neural-based sequence labelling," *arXiv preprint arXiv:1810.10156*, 2018.
- [11] P. Badjatiya, S. Gupta, M. Gupta, and V. Varma, "Deep learning for hate speech detection in tweets," in *Proceedings of the 26th International Conference on World Wide Web Companion*, 2017, pp. 759–760, doi: 10.1145/3041021.3054223.
- [12] M. Pontiki, D. Galanis, H. Papageorgiou, S. Manandhar, and I. Androutsopoulos, "Semeval-2015 task 12: Aspect based sentiment analysis," in *Pro. 9th Int. Workshop on Semantic Evaluation (SemEval 2015)*, 2015, pp. 486–495, doi: 10.18653/v1/S15-2082.
- [13] C. Wagner, A. Dulaunoy, G. Wagener, and A. Iklody, "Misp: The design and implementation of a collaborative threat intelligence sharing platform," in *Proceedings of the 2016 ACM on Workshop on Information Sharing and Collaborative Security*, 2016, pp. 49–56, doi: 10.1145/2994539.2994542.
- [14] G. Lample, M. Ballesteros, S. Subramanian, K. Kawakami, and C. Dyer, "Neural architectures for named entity recognition," *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, 2016, pp. 260-270, doi: 10.18653/v1/N16-1030.
- [15] A. J. Yepes and A. MacKinlay, "Ner for medical entities in twitter using sequence to sequence neural networks," in *Proceedings of the Australasian Language Technology Association Workshop 2016*, 2016, pp. 138–142.
- [16] C. B. Aslan, R. B. Saglam, and S. Li, "Automatic detection of cyber security related accounts on online social networks: Twitter as an example," in *Proc. 9th Int. Conf. social media and Society*, 2018, pp. 236–240, doi: 10.1145/3217804.3217919.
- [17] F. Alves, A. Bettini, P. M. Ferreira, and A. Bessani, "Processing tweets for cybersecurity threat awareness," *Information Systems*, vol. 95, 2021, doi: 10.1016/j.is.2020.101586.
- [18] F. F. Duarte, O. M. Pereira, and R. L. Aguiar, "Discovery of newsworthy events in twitter," in *Proceedings of the 3rd International Conference on Internet of Things, Big Data and Security (IoTBDs 2018)*, 2018, pp. 244–252, doi: 10.5220/0006712702440252.
- [19] A. Seth, S. Nayak, J. Mothe, and S. Jadhay, "News dissemination on twitter and conventional news channels," in *19th International Conference on Enterprise Information Systems (ICEIS 2017)*, vol. 1, 2017, pp. pp. 43-52, doi: 10.5220/0006264100430052.
- [20] A. Javed, P. Burnap, and O. Rana, "Prediction of drive-by download attacks on twitter," *Information Processing & Management*, vol. 56, no. 3, pp. 1133–1145, 2019, doi: 10.1016/j.ipm.2018.02.003.
- [21] R. P. Khandpur, T. Ji, S. Jan, G. Wang, C.-T. Lu, and N. Ramakrishnan, "Crowdsourcing cybersecurity: Cyber attack detection using social media," in *Proceedings of the 2017 ACM on Conference on Information and Knowledge Management*, 2017, pp. 1049–1057, doi: 10.1145/3132847.3132866.
- [22] F. H. Sohime, R. Ramli, F. A. Rahim, and A. A. Bakar, "Exploration Study of Skillsets Needed in Cyber Security Field," *2020 8th International Conference on Information Technology and Multimedia (ICIMU)*, 2020, pp. 68-72, doi: 10.1109/ICIMU49871.2020.9243448.
- [23] N. Dionísio, Working Dataset, Github, 2020. Accessed: Feb. 9, 2020. [Online]. Available: <https://github.com/ndionysus/twitter-cyberthreat-detection/tree/master/data>
- [24] M. A. Razek, C. Frasson, and M. Kaltenbach, "Dominant meanings approach towards individualized web search for learning environments," in *Advances in web-based education: personalized learning environments*, pp. 46–69, 2006, doi: 10.4018/978-1-59140-690-7.ch003.
- [25] M. A. Razek, "Towards more efficient image web search," *Intelligent Information Management*, vol. 5, no. 6, pp. 196-203, 2013, doi: 10.4236/iim.2013.56022.
- [26] Y. Ibrahim, M. A. Razek, and K. A. ElDahshan, "Towards more Efficient for Web Image Search Engine using Dominant Meaning Technique," *International Journal of Computer Trends and Technology (IJCTT)*, vol. 54, no. 2, pp. 91-96, 2017, doi: 10.14445/22312803/IJCTT-V54P114.




BIOGRAPHIES OF AUTHORS

Yasser Ibrahim Abdelmonem Ali    is Lecturer Assistant at Math. & Computer Science Department, Faculty of Science, Azhar university, Nasr City, Cairo, Egypt. He born in Cairo, Egypt, holds a Bachelor's in Pure Mathematics and Computer Science with Very Good from degree. He Holds a master's degree in Computer Science from Math. & Computer Science Department, Faculty of Science, Azhar university with Image Retrieval. Pупlished paper with Title "Towards more Efficient for Web Image Search Engine using Dominant Meaning Technique". He can be contacted at email: yasser.ibrahim@azhar.edu.eg.



Dr. Mohammed Abdel Razeq    is a Professor of Computer Science at Azhar University. He holds a Ph.D. in Computer Science-Artificial Intelligence-from University of Montreal, Canada in 2004. His research focuses on the design of a new application using artificial intelligence techniques on e-learning, Medicine, Cybersecurity, Internet of Thing, and others. He has more than 80 papers published in international journals and Conferences. He serves as an editor member for many Journals and as a reviewer of many international conferences. As a postdoctoral fellow at NSERC, Canada, He had worked in creating intelligent signing system to manipulate a huge database containing customers' purchases at Retail Company. He has been added to Who is Who in the world in 2009. He is working as a consultant for quality assurance for traditional and online education at many institutions: Minster of Higher Education Egypt, King Abdul Aziz University-Saudi Arabia, and National Authority for Quality Assurance and Accreditation of Education (NAQAAE). He can be contacted at email: abdelram@azhar.edu.eg.



Dr. Nasser A. El-Sherbeny    born in Mansoura, Egypt, holds a Bachelor's in Mathematics with Excellent degree. He had his master's degree in Mathematics from Mathematics Department, Faculty of Science, Mansoura University. He had his DEC2 from Faculte' Polytechnique de Mons, Mons, Belgique. He had his Ph.D. degree in Mathematics from Mathematics Department, Faculty of Science, Mons University, Mons, Belgium. Currently works as a Prof. in Mathematics, Mathematics Department, Faculty of Science, Al-Azhar University, Cairo, Egypt, published several scientific researchers in Optimization Networks. He can be contacted at email: nasserelsherbeny@yahoo.com.