Cyberbullying Detection in Syrian Slang on Social Media by using Data Mining

Reem Thabt Ali
Faculty of Information Technology
Syrian Virtual University
Damascus, Syria

Dr. Mohamad-Bassam Kurdy
Faculty of Information Technology
Syrian Virtual University
Damascus, Syria

Abstract— **Social media has become a fertile medium for the users to exchange their information in different languages and slangs, In spite of the clear advantages of social media, social media platforms have significant adverse consequences. The users have the ability to exploit social media to cause the annoyance and embarrassment for others, this embarrassment is known as cyberbullying which has become extremely common especially in Arab nations and causes serious influences on the victims. Thus, cyberbullying Detection on different social media platforms takes the concern of the researches, but the most studies proposed approaches to detect cyberbullying in English language and few researches were in other languages. Due to the widespread usage of Arabic languages specifically Syrian slang on social media platforms, the scarcity of studies in Cyberbullying Detection in Syrian slang on social media and the lack of the textual dataset in Syrian slang, all these reasons were a motivation to propose a method to detect a cyberbullying content in Syrian slang on Facebook. This approach based on Data mining Algorithms which are applied on dataset extracted from comments on Facebook posts, in addition, this paper introduces an algorithm to measure the severity of cyberbullying in a comment. Moreover, comparisons were made of used classifiers against the Accuracy, Recall, Precision, and F1-Measure metrics. A reasonable accuracy of 77% in detecting one of cyberbullying categories: sexual, physical sexual, religious, political, appearance (cyberbullying related to animals or something), racism, cultural, psychological, praying negatively for person, and general cyberbullying by Support Vector Machine classifier, while Adaptive Boosting algorithm achieved the highest Precision rate of 94%, and Stochastic Gradient Descent classifier recorded the best Recall and F1-Score metrics rates of 47%, 52% respectively.**

Keywords:- Data Mining, Cyberbullying, Syrian slang, Dialectal Arabic, Modern Standard Arabic.

I. INTRODUCTION

Disclaimer: Due to the nature of the paper, some examples contain highly cyberbullying language. They do not reflect the views of the authors in any way, and the point of the paper is to help fight this type of content on social media.

Nowadays, people around the world use different social networking platforms like Twitter, Facebook, and Instagram, etc. The posts, comments and replying are not shared only in a positive way, but also might be used inappropriately on social platforms, this is often referred as cyberbullying which is defined as willful and repeated harm inflicted through the medium of electronic text [1].

Cyberbullying is emerging as a serious social issue, because of its negative physiological effects on the victims,

according to American Academy of Child and Adolescent Psychiatry, bullying in different forms can lead to serious academic, social, emotional and legal difficulties [2]. As the remarkable rise in Cyberbullying in Arabic pages, most of previous researches used data mining techniques to detect cyberbullying which written in Modern Standard Arabic (MSA) on social media and less work was directed toward cyberbullying detection in Dialect;;al Arabic (DA) which is an accent for Arabic language. Due to the variety of Dialectal Arabic, the absence of the rules, the scarcity of the resources in Dialectal Arabic [3] and some words are considered offensive in a country, while good in another country, for instance the word "Yetqalash" in Yemen is a good word, while in Morocco it is a profane [4]. Moreover, there is a lack of Arabic language databases, specifically the datasets of Dialectal Arabic.

This paper would propose an approach to detect cyberbullying in Arabic Syrian slang by using different data mining models on Facebook which is the most used by Syrian according to new statistics which have resulted that 86% of people in Syria use Facebook while about 1.7% use Twitter [5], Moreover, this research would detect the severity of cyberbullying behavior in Syrian slang.

This paper outlines the related works in section 1.0, the methodology to detect cyberbullying presents in section 2.0, the comparison of the results among different data mining models and the evaluation of these results would be organized in section 3.0, a proposed algorithm to calculate the severity of cyberbullying in the text is discussed in section 4.0, finally, conclusions and future works are stated in section 5.0.

II. RELATED WORK

Although there are few papers on Cyberbullying detection in dialectal Arabic on social media, recently many researchers have been working on detecting cyberbullying category on social media platforms which written in different languages by using data mining or machine learning algorithms without detection of cyberbullying severity.

One of the more notable research was conducted by [6], the researcher focused on detecting cyberbullying on Twitter and its growth in Indonesia,. This study provides statistics of which cyberbullying category is often used by the abusers in Indonesian language on their victims, so detection was implemented by using Naïve Bayes Classifier which is simple algorithm and has high speed in training and classifying[7,8]

the texts into one of these four cyberbullying categories: psychology, sexuality, cyberbullying related to animals and general cyberbullying. The data have been collected from twitter by Twitter API through the duration of November-December 2016 and the collecting dataset contains 1245 tweets, an automatic preprocessing of data has been applied for the dataset, such as TF-IDF weighting, stop words, Case folding and N-Gram .The preprocessing step was achieved by machine learning program called WEKA [9], then data validation has been done by using 10 Fold Cross Validation from machine learning WEKA, finally the classification step by using Naïve Bayes Classifier. The result of this research is that most widely used cyberbullying category by abusers on Twitter in Indonesian language during the study period was psychology category, which reached 45.56%.

Another work done by [10] which presents a solution for the issue of cyberbullying in both Arabic and English languages, the proposed solution used machine learning algorithms which required a huge dataset to be collected, this dataset was obtained from both Facebook and Twitter, the data was scrapped by two tools, the first tool is Twitter scrapper which was written using PHP, the tweets were collected using this scrapper from Lebanon, Syria, Gulf Area and Egypt and this database reached 4.93 GB of size, the second one is Facebook scrapper which was written in Python and the database from Facebook reached 0.98 GB of size . After that the Version 3.9.1 of WEKA was used to clean and preprocess the gathering data, the texts which written in other languages than English or Arabic were ignored, as a result of that dataset contained 35273 Arabic texts and 91431 English tweets. Arabic content was labeled manually by adding another attribute which is "bullying" which has either "yes" for cyberbullying content or "no" otherwise. In this research, Naïve Bayes and SVM models were chosen to classify the text, this decision has been made according to [4,11,12 and 13], which ensured these two algorithms are the best for classification of the texts, the study based on training and classifying the dataset for several times to reach the best results. Firstly the proposed system was trained by Naïve Bayes algorithm and the system was able to detect 801 out of 2196 texts in the first run, and this result achieved the aim of the research that cyberbullying in Arabic is possible, Naïve Bayes model achieved a precision of 36.5 % for "yes" class and 94.5 % for "no" class. For training by Support Vector Machine (SVM) model, an Affective Tweets package was utilized [14] and particularly the Tweet ToSentiStrengthFeatureVector filter which depends on SentiStrength technique which developed by [15], this filter has the ability to support English and other languages, moreover, this filter was customized for Arabic language by replacing the files of English lexicon with Arabic files which contain the Arabic weighted profane words from multiple Arabic countries. SVM model achieved greater precision for the "yes" class 81.5% than Naïve Bayes classifier, but approximately the same precession for the "no" class 94.4%. Whereas, the other measures like ROC, TP, FP, F-score and recall were approximately the same.

In another paper [16], the authors built a large dataset which consists of offensive Arabic words from different dialects and topics. This dataset was collected by using Twitter API to gather the tweets over the time from 15-April - 2019 to 6- May- 2019 and reached 660k tweets of size. In addition, it is obvious that pattern which has vocative particle يا("ya" – meaning "O")[1] which is used basically in directing the speech to the people was involved mostly in obtained Arabic offensive tweets. The next step after obtaining the dataset of Arabic offensive words was to annotate these tweets by an experienced native speaker with a good knowledge of different dialectal Arabic, the tweets were labeled to one of these categories: offensive, vulgar, hate speech, or clean. Since the offensive tweets involve implicit insults or attacks against others, vulgar category contains profanity words like sexual parts of the body, hate speech category was the tweets which contain racism, religious and ethnic words, and clean label was the tweets which do not contain vulgar or offensive language and take in consideration the tweets which composed of some offensive words, but the whole tweet was not considered as offensive. After that the dataset was preprocessed and cleaned such as tokenization by Farasa Arabic NLP toolkit [16], removing URLs, hashtags, digits, mentions, retweets, Arabic normalization, normalized letter repetition to keep at most two repeated letters and finally elimination the diacritics (tashkeel).

Different classifiers were employed in this study, SVM model with a radial function kernel was mainly used with lexical features and pre-trained static embedding [17, 18] while Adaptive Boosting and Logistic regression classifiers were employed when using Mazajak embedding [19]. SVM gave the best precision of 88.6%, Logistic Regression achieved 84.7% of precision, while Adaptive Boosting model achieved 74.3 of precision.

Another work has been achieved by [20], the authors provide a description to develop an algorithm to detect the cyberbullying in SMS messages, for this study, real text messages were collected from eleven cell phones related to the participants in this research, over 2016 the collecting data reached 80k messages, then those messages were separated into conversations which consist of three basic components: sender, receiver, and time stamp, after that the sender and receiver were matched to decide whether both of them relate to the same conversation or not. Each text message was tokenized into words, a comparison between each word in the message and the words in a specialized dictionary of bullying words which has been created by the authors, and the word takes the score of the bully word and adds it to the total message score. If the message score passes a threshold, then the messages is considered as cyberbullying. Once the scores are calculated for all the messages, then the algorithm adds all the message scores in a conversation and if the total conversation score exceeded the threshold, then the conversation is entirely classified as cyberbullying.

---

[1] Arabic words are provided along with their Bulkwalter transliteration and English translation.

The proposed algorithm achieved a precision of 64.58%, to improve this result, the algorithm was running through other three datasets, after running this algorithm through these sets of data, the research found that the proposed algorithm is capable of reaching a precision of at least 90%.

Another study was presented by [21] which does not focus only on the cyberbullying detection but also on the detecting of the cyberbullying severity, in this study the authors selected the dataset of harassment which provided by [22], this dataset was classified into five categories: sexual, racial, appearance, intelligence and political. Then
The dataset was preprocessed by converting the tweets to the lowercase, decrease the number of the repeated letters, eliminating the URLs and namely mentions, applying tokenization by using CMU TweetNLP library [23], in the end removing stop Words and applying stemming [24].

In order to estimate the cyberbullying severity of the dataset, the annotated cyberbullied tweets were categorized into four different levels: low, medium, high, and non-cyberbullying, then the researchers classified: sexual and appearance cyberbullied tweets as high-level cyberbullying intensity, tweets related to political and racial as medium-level, and intelligence cyberbullied tweets as low-level. After that, a framework was developed to extract the feature by applying part-of-speech (POS) with Twitter tagger which depends on CMU TweetNLP library [23] to disambiguate the word sense, then the feature generation step was done by applying the classification of the document level and measurement of the semantic orientation of each word in the dataset [25].

After the previous steps, the study tested multiple machine learning algorithms to choose the most effective classifier to detect the cyberbullying content and its severity, the chosen models were: Naïve Bayes, Support Vector Machine (SVM), Decision Tree, Random Forest, and K-Nearest Neighbors (KNN), these models achieved respectively a precision of 85.5%, 88.5%, 89%, 89.8%, and 86.9%.

III. METHODOLOGY

The main aim of the proposed methodology to detect cyberbullying on social media is to reduce the bullying behavior [26], this proposed method can be employed to give an assistance to fight cyberbullying and monitor the written texts in Syrian slang on social media platforms. Moreover, the detection of cyberbullying on social media might be considered as a proactive step to support and advise the victims of the bullying [27].
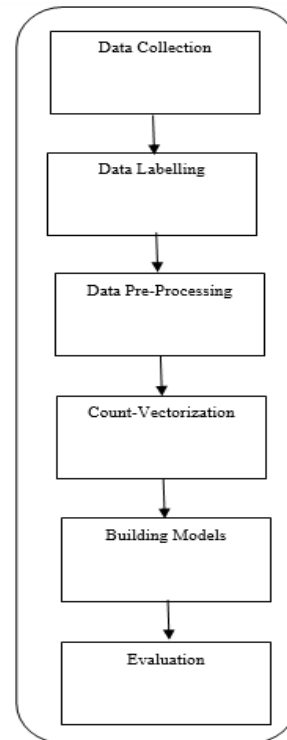

Fig. 1 Proposed Methodology

At the present days, Facebook is the most popular online social network, about seventy-nine percent of internet users have accounts on Facebook, and on average these users access Facebook eight times a day [28], for this reason, the required dataset for this study was collected from Facebook specifically the comments on the posts in Arabic Syrian which written in Syrian slang.

This section debates the research methodology that would be used in cyberbullying classifying with the right cyberbullying category and measuring the severity of the detecting cyberbullying. All steps for the proposed methodology are illustrated in "Fig. 1" above, and discussed in the following sections.

A. Data Collection:

As stated previously, the choice was to scrap data from Facebook, this decision was made because Facebook is most widely used by Arab nation, especially in Syrian community.

For the acquirement of data from Facebook, Facepager tool [29] that is an automated data retrieval application was used to fetch public data from different social media platforms like YouTube, Facebook, Twitter, and Amazon, the technique of comments extraction from Facebook shown in "Fig. 2" above.
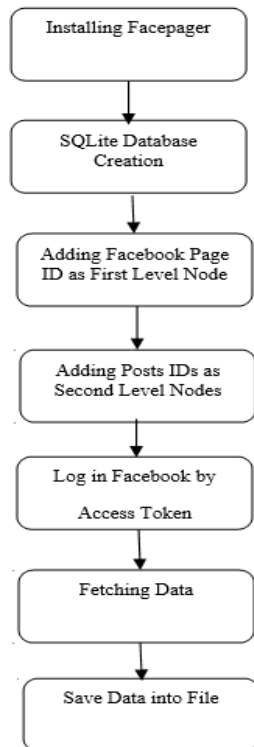
Fig. 2 Data Collection

Facepager tool makes the scrapping procedure easier and faster, after installing Facepager application, a SQLite database should be created to store the Facebook graph for the scrapped comments, then Facebook page identifier was added as the first node, and the identifiers of the posts were added as second level nodes under page identifier where the comments would be collected.

In this study the required data was obtained from popular Syrian page called "Yhrek Debak" on period January-May 2021, an access token was required to crawl the data from Facebook posts, for this reason an application was created on the Facebook Developers website and associate it with Facebook page "Yhrek Debak" to establish the connection, finally logging in to Facebook, the comments were obtained successfully and stored as Excel file which contains 4235 comments.

Furthermore, to enrich the collecting data, a simple questionnaire was published to get more examples of cyberbullying content in Syrian slang from native speakers.

B. Data labelling:

Arabic language is considered a sensitive and difficult language, in order to ensure the proper labelling of the crawled data, Arabic Syrian slang comments were labelled manually by adding another attribute called "bullying", this attribute was assigned to one of the following cyberbullying categories:

1)Sexual cyberbullying: this category refers to the content that embarrasses the people with sexual insults such as "Slut" "عاهرة".

2) physical sexual cyberbullying: is a type of cyberbullying contains explicitly the private sexual parts of the body.

3) Religious cyberbullying: this Category consists of the phrases target the religion and ideology of the victim, like "مسلم متشدد" which means "Strict Muslim"

4) Political cyberbullying: which contains the bullying words regarding the attitudes in policy such as "دنب الحكومة" which mean"the tail of government".

5) Appearance cyberbullying; this type related to the words which describe the appearance of the victims, for instance "أنتي فزيعة" which means "you are ugly".

6)Cyberbullying related to animals or something: this type associates with description of the people with type of animals, for instance "بتشبه القرد" which means" you look like the monkey " or describing any part of the victim's body with something else such as "أنفك متل البطاطاية" this means "your nose looks like potato" in other words "your nose is so big".

7)Racism cyberbullying: this cyberbullying is directed to the race or ethnicity of individuals [30].

8)Sectarian cyberbullying: these cyberbullying phrases related to the sects of the people, such as "شيعي مرتد".

9) Cultural cyberbullying: In this category, the abusers attempt to attack a culture of the victim in some way.

10)   Psychological cyberbullying: is the bullying related to psychology, for example"انت واحد ما بتفهم"
 means "you are stupid".

11)   praying negatively for person: such as "الله يلعنك" which means "god curses you".

12)   General cyberbullying: is the cyberbullying content which does not belong to the previous categories, and it is not correct to classify it as one of them, for instance"تفو عليك" which means "spit on you".

13)   Non cyberbullying: the text does not contain any words related to cyberbullying behavior.

This classification of cyberbullying behavior has been validated by an expert in Arabic language and native speakers.

C. Data Preprocessing:
In this stage as is presented in "Fig. 3" below, different data preprocessing techniques is performed with the obtained data, this step is so important, due to its role in reduction the data complexity and improving the quality of the results of the later step, the data preprocessing phase was implemented by Python language and this step consists of the following steps:

1)Uploading Data: the crawled dataset was imported to apply the preprocessing steps which done by IO which is a part of standard Python library.

2)Reading Data: the uploading data was read by dataframe which is a column oriented tabular Python data structure [31].

3)Noise Elimination: this step was applied for the collecting data by special codes which written using Python language, all non-Arabic letters, digits, punctuations, URLs, mentions, emotional stickers, symbols, hashtags, Stop words, repeating letters (كتيرررر), Arabic Diacritics like ('ُ') and Tatweel ('__') were removed.

4)Normalization: for example (Hamza 'أ' converted to 'ا' and Alef Maksours 'ى' is normalized to 'ي'), and this step was achieved by a code written in Python.

5)Tokenization: Arabic text was broken into words and discrete units which are known as token by using a tokenizer written in Python.

6)Stemming: several Arabic stemmers which implemented in Python, such as Tashaphyne and ISRI
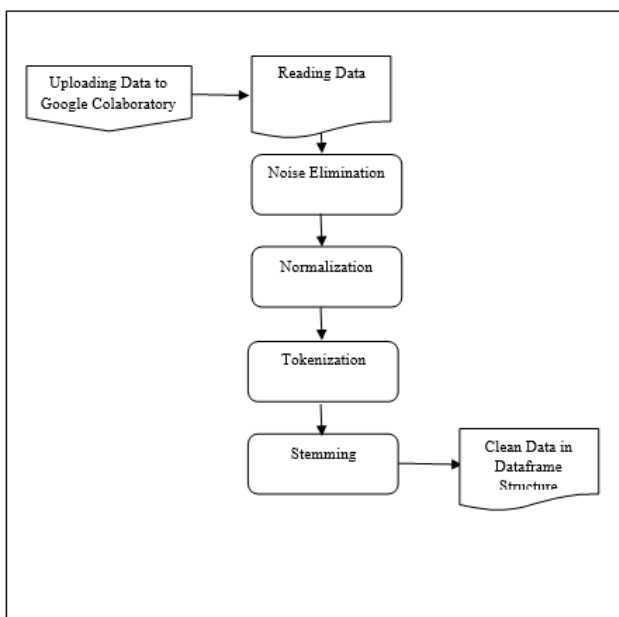

Fig. 3 Preprocessing steps

Arabic were tested on some words in Syrian slang, but the results were not precise adequately, for this reason, a specialized stemmer was built in Python to achieve this step.

D. Count-Vectorization:

After preprocessing and labelling data, the dataset has been split into training dataset that contains 80% of the entire obtained dataset and used for learning the classifier, and testing data which consist of 20% of the whole crawled dataset which acts an important role in providing unbiased

evaluation of a final model fitted on the training set. In this study the training set has 3388 comments while 847 comments have been involved in testing set, after that, Count Vectorization technique was applied to extract the features from the crawled dataset [32] to transform the textual comments into numerical values and return the corresponding matrix which can be understood and used directly by models to classify the comments with the correct cyberbullying category.

E. Building models:

Due to the significance of choosing the most effective model to detect the cyberbullying behavior and classifying it with the proper cyberbullying category, nine data mining algorithms have been employed in training and classification procedure, and all these models were built by using a Python library called Sklearn. The classifiers adopted in the current study are as follows:

1) Support Vector Machine (SVM): this is a supervised machine learning model and mostly employed in text classification, in addition, this classifier is traditionally used for binary classification [33], so it is required to be modified to work with multi-class classification [34] because several classes have been considered in this study.

2) Multinomial Naïve Bayes: this classifier is regarded as an effective classifier that used for text classification problems in variety of social media studies [34], for this reason Naïve Bayes algorithm has been chosen in this research.

3)Decision Tree: This is a supervised classifier. Because of the capabilities of Decision Tree algorithm to learn disjunctive expressions that make it an appropriate choice for text classification [36].

4)Random Forest (RF): this algorithm is a supervised classification algorithm, according to [37] Random Forest is an effective algorithm to assess the missing values of data and provide a reasonable accuracy even if a large portion of missing data is involved in dataset.

5)K-Nearest Neighbors (KNN): is a supervised model which memorizes observations from within a labelled to predict the labels for new and unlabeled observations. Moreover, this classifier is suitable for multi-class problems [34].

6)Multinomial Logistic Regression: this model is regarded as the one the most popular model in machine learning models, it is a linear model has become an important model used in multiclass classification, and is regarded as a statistical approach which used in different studies in machine learning and data mining [38].

7) Adaptive Boosting: this classifier is considered as an ensemble classifier (consists of various classifying algorithms), and uses iterative approach to learn from weak classifier's errors and make them stronger, according to [39] Adaptive Boosting algorithm gives a

good accuracy in cyberbullying Detection on social media.

8)Stochastic Gradient Descent (SGD): is a linear classifier (SVM, logistic regression) optimized by the SGD. These are two different concepts. While SGD is an optimization method, Logistic Regression or linear Support Vector Machine is a machine learning algorithm [40].

9)Bagging: It is also an ensemble meta-estimator classifier where base classifiers are fitted on each random subsets of the original dataset and then compute their individual predictions, either by voting or averaging to output a final expectation.

All previous classifiers were run through the obtained dataset to learn them and then to predict the proper category of cyberbullying, the time which has been taken in the training and prediction procedures illustrated in "Fig. 4" below, as it shown in "Fig. 4", the best training time was recorded by K-Neighbors classifier while the worst training time relates to Bagging model, the best prediction time belongs to Support Vector Machine classifier, whereas the worst prediction time associate with Random Forest classifier.
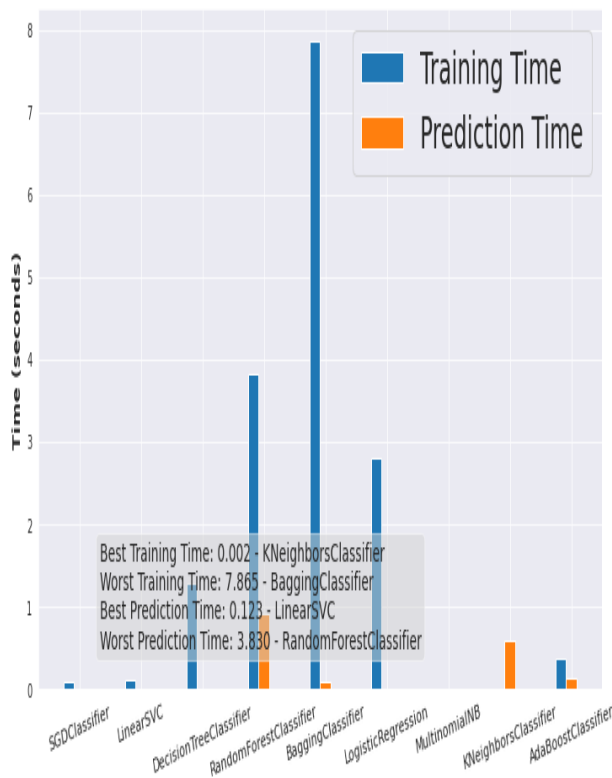

Fig. 4 Time Complexity of Algorithms

F. Evaluation:
Several metrics exist to measure the performance of the Data Mining or Machine Learning classifiers, these metrics are based on "Confusion Matrix" [41] which consists of True Positives (TP): these are the accurately predicted positive values, True Negatives (TN): the values which labelled accurately as negative values, False Positive (FP): the instances which predicated incorrectly as belonging to the

positive class, in this case when the comment is wrongly labelled as cyberbullying behavior, False Negatives (FN) represent the cases which labelled as relating to negative class but should have been labelled as positive class, the performance metrics which used in this study were described below:

1) Accuracy: is the most popular utilized metric and defined as the ratio of correctly returned valued to the total values.

$$(TP)+(TN)/(TP)+(FP)+(TN)+(FN) \qquad (1)$$

2)Precision: is determined as the proportion of relevant observations which are correctly predicted positive values out of the all predicted positive values.

$$(TP)/(TP)+(FP) \qquad (2)$$

3)Recall: is the ratio of returned relevant values to the total values in the entire class.

$$(TP)/(TP)+(FN) \qquad (3)$$

4)F1-measure: is a special case of F-Measure which defined as the weighted mean of Precision and Recall, and was founded to overcome the problem of the negative correlation between Precision and Recall [42], this problem was solved by $\beta=1$ which is a parameter to control the balance between Recall and Precision where $0 \leq \beta \leq \infty$ [43].

$$(2 \times precision \times recall)/(precision+recall) \qquad (4)$$

IV. RESULT OBTAINED

A comparison of the performance was achieved among the used data mining classifiers according to the following evaluation metrics: Accuracy, Precision, Recall and F1-Measure, these metrics were presented in [Table I] below. It shows, Stochastic Gradient Descent (SGD) and Support Vector Machine classifiers (SVM) achieved the highest Accuracy of 77% as compared to other classifiers, Adaptive Boosting achieved the greatest Precision rate of 94%, as it shown in [Table I] Stochastic Gradient Descent (SGD) model recorded the highest Recall rank of 49% and F1-Measure rate of 53%. It is obvious that results obtained by Stochastic Gradient Descent (SGD) and Support Vector Machine models are closer to each other somewhat than other evaluation metrics which achieved by the rest data mining models in this study.

Published by :

http://www.ijert.org

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 11 Issue 04, April-2022

TA BLE I. A COMPARISON OF PERFORMANCE AMONG CLASSIFIERS.

Algorithm	*Accuracy*	*Precision*	*Recall*	*F1-Score*
SGD	0.762	0.728	**0.473**	**0.523**
SVM	**0.768**	0.754	0.452	0.507
Bagging	0.733	0.737	0.403	0.465
Radom Forest	0.734	0.777	0.383	0.461
Decision Tree	0.726	0.692	0.375	0.422
Logistic Regression	0.721	0.821	0.235	0.282
Multinomial NB	0.716	0.800	0.206	0.239
K- Neighbors	0.631	0.807	0.147	0.173
Adaptive Boosting	0.578	**0.944**	0.089	0.077

The "Fig. 5" below presents graphically the classification summary of the classifiers, as seen the Accuracy, Recall and F1-Measure achieved by Stochastic Gradient Descent (SGD) and Support Vector Machine classifiers (SVM) was the highest, while the best Precision recorded by Adaptive Boosting and Logistic Regression models respectively.
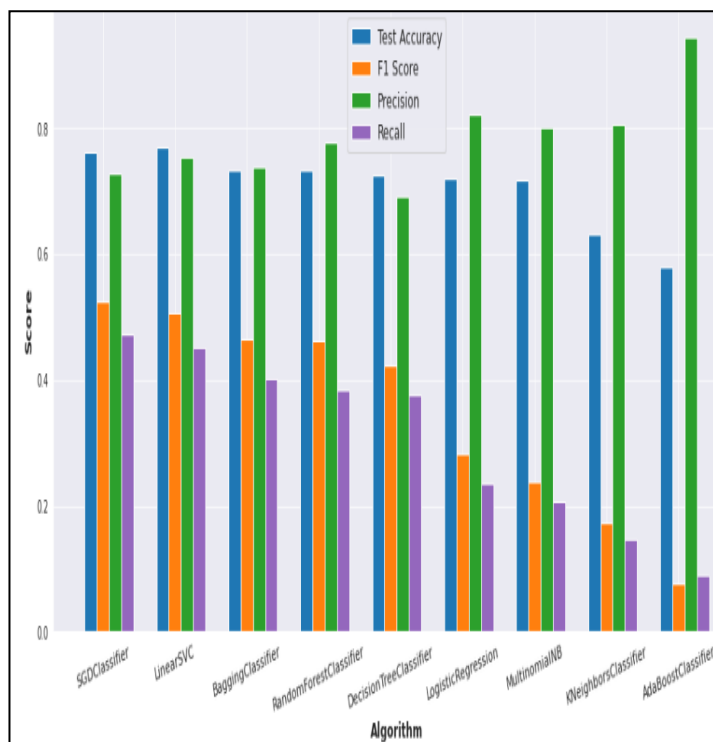

Fig. 5   Classification Summary of Algorithms

CALCULATION OF CYBERBULLYING SEVERITY

After classification the comment with its cyberbullying category by the previous data mining models, the intensity of cyberbullying in that comment which written in Syrian slang was measured by running the proposed algorithm in the study [21] through the comment, but this algorithm was customized to fit Syrian slang by building a dictionary of bullying words in Syrian slang and each bullying word was assigned with its score according to the cyberbullying category of this word as it shown in [Table II].

When the algorithm starts, each comment splits into separate tokens (words) and each token is compared to the words in the bullying dictionary to take its corresponding score and then add this score to the total comment score which has an initial value of zero, then total comment score is compared to the cyberbullying ranges which presents in the [TABLE III] to determine the cyberbullying severity of the comment: high, medium, low.

The intensity of cyberbullying belongs to the numerical domain [0,1], and these domains of cyberbullying severity were approved by expert in Arabic language, in addition, this study exploits the classification of cyberbullying severity stated in [22] which classified the sexual and appearance cyberbullying as high level, racial and political cyberbullying as medium level. It was taken in consideration some Arabic words which give the text more emphasis and in their turn increase the cyberbullying severity, such as:" كتير ، جدا، كثير " which mean "very much".

TABLE II.   SEVERITY OF EACH CYBERBULLYING CATEGORY

Cyberbullying Category	*Severity*	*Score*
Sexual	High	0.5
Physical Sexual	High	0.5
Appearance	High	0.5
Cyberbullying related to animals or something	High	0.5
Political	Medium	0.4
Psychological	High	0.5
Religious	Medium	0.3
Racism	Medium	0.4
Sectarian	Medium	0.4
Praying Negatively for Person	Medium	0.3
Cultural	Low	0.2
General	Low	0.2
Emphasis Words	Low	0.1

TABLE III.   NUMERICAL DOMAINS OF CYBERBULLYING SEVERITY

Numerical Domain	Arabic Name	English Name
[0.5,1]	مرتفع	High
[0.3,0.4]	متوسط	Medium
[0.1,0.2]	منخفض	Low

V.   CONCLUSION AND FUTURE WORK

This research presented the first dataset of cyberbullying content in Syrian slang and proved that cyberbullying in Syrian Slang is detectable in spite of the difficulties of the Arabic Language and its different dialects, and that is regarded as a motivation to take more steps in the future in cyberbullying Detection in other Dialectal Arabic.

The collecting dataset for this study could be a base for other studies in the same field and could be enhanced by extending this dataset through crawling more data from different social media platforms, and employed them in the future study. In addition, the proposed methodology could be upgraded not only to detect cyberbullying in Syrian slang but also in other Dialectal Arabic from different countries and block the text which contains cyberbullying behavior. Furthermore, development an ensemble model which improves the performance of the proposed method to detect cyberbullying, or employing deep learning algorithms instead of data mining and machine learning models in cyberbullying detection and make a comparison among the results.

REFERENCES

[1] J.W. Patchin and S. Hinduja, "Bullies move Beyond the Schoolyard;a Preliminary Look at Cyberbullying," Youth Violence and JuvenileJustice, vol. 4, no. 2, pp. 148–169, 2006.

[2] American Academy of Child and Adolscent Psychatiry,"Bullying Facts of Families" No. 80(3/11), Washington, retrived from https://www.aacap.org/App_Themes/AACAP/docs/facts_for_families/ 80_bullying.pdf, pp.1, March 2011.

[3] Nizar Habash, Ryan Roth, Owen Rambow, Ramy Eskander, and Nadi Tomeh, "Morphological Analysis and Disambiguation for Dialectal Arabic" NAACL-HLT, Atlanta, Georgia,pp. 426–432, June 2013.

[4] "12 Arabic Swear Words and Their Meanings You Didn't Know," [Online]. Available: http://scoopempire.com/swear-words-meanings-around-middle-east/#.V0fdjPl96M9, August 2014.

[5] Social Media Stats Syrian Arab Republic, retrived form StatCounter Global Stats: https://gs.statcounter.com/social-media-stats. Syria, 2020.

[6] Hariani Imam Riadi," Detection Of Cyberbullying On Social Media Using Data minig technique", International Journal of Computer Science and Information Secutity, pp. 244-250, March 2017.

[7] S. Susanto and D. Suryadi, "Introduction to Data Mining, Digging Up Data Chunks Of Knowledge," C.V. Andi Offset, Yogyakarta, 2010.

[8] N. W. S. Saraswati. Tagawa, "Text Mining with Naïve Bayes Method and Support Vector Machines for Sentiment Analysis, " Thesis, Udayana Univesity Denpasar, 2011.

[9] S. Garner, "Weka: The waikato environment for knowledge analysis," Newzealand, 1995.

[10] Ahmed Serhrouchni, Chamoun Maroun Batoul Haidar, "A Multilingual System for Cyberbullying Detection: Arabic Content Detection using Machine Learning", Advances in Science Technology and Engineering Systems Journal Vol. 2, No. 6, pp. 275-284, December 2017.

[11] N. Potha and M. Maragoudakis, "Cyberbullying Detection using Time Series Modeling," IEEE International Conference on Data Mining Workshop, 978-1-4799-4274-9/14 $31.00 © 2014 IEEE.

[12] B. Nandhinia and J. Sheebab, "Online Social Network Bullying Detection Using Intelligence Techniques," International Conference on

[13] K. Darwish and W. Magdy, "Arabic Information Retrieval," 978-1-60198-777-8, , vol. 7, no. 4, pp. 239–342, 2013.

[14] S. M. Mohammad and F. Bravo-Marquez, "Emotion Intensities in Tweets," in Joint Conference on Lexical and Computational Semantics (*Sem), Van-couver, Canada, pp.65-77, August 2017.

[15] M. B. K. P. G. C. D. &. K. A. Thelwall, "Sentiment strength detection in short informal text ," Journal of the American Society for Information Science and Technology, vol. 61, no. 12, pp. 2544–2558, 2010.

[16] Ahmed Abdelali, Kareem Darwish, Nadir Durrani, and Hamdy Mubarak, "Farasa: A fast and furioussegmenter for arabi," the conference of the North American chapter of the association for computational linguistics, pp.11-16, 2016.

[17] Piotr Bojanowski, Edouard Grave, Armand Joulin, andTomas Mikolov, "Enriching word vectors withsubword information," Transactions of the Association for Computational Linguistics, pp. 135-146, 2017.

[18] Abu Bakr Mohammad, Kareem Eissa, and Samhaa El-Beltagy, "Aravec: A set of arabic word embedding models for use in arabic nlp," Procedia Com-puter Science, 117, pp. 256-265, 2017.

[19] Ibrahim Abu Farha and Walid Magdy, "Mazajak:An online Arabic sentiment analyser, ". InProceedings of the Fourth Arabic Natural Language Processing Workshop, Florence ,Italy, Association for Computational Linguistics, pp. 192-198, 2019.

[20] Bradley, Bryan W, "Detection of Cyberbullying in SMS Messaging," Computer Science Summer Fellows.3. https://digitalcommons.ursinus.edu/comp_sum/3, July 2016.

[21] Bandeh Ali Talpur, Declan O'Sullivan, "Cyberbullying severity detection: A machine learning approach," plos one journal community,https://journals.plos.org/plosone/article?id=10.1371/journal .pone.0240924, October 2020.

[22] Rezvan M, Shekarpour S, Balasuriya L, Thirunarayan K, Shalin VL, Sheth A, "A Quality Type-aware Annotated Corpus and Lexicon for Harrassment Research," proceeding of the 10th ACM Conference on Web Science, Newyork, NY, USA: ACM, pp. 33-36, 2018.

[23] Gimpel K, Schneider N, O'Connor B, Das D, Mills D, Eisenstein J, "Part of Speech Tagging for Twitter : Annotation, Features, and Experiments," proceeding of the 49th Annual Meeting of Association for Computational Lingusitics: Human Language Technoogies, Portland, Oregon, USA, pp. 42-47, 2011.

[24] Lovins JB, "Development of a stemming Algorithm," Mech Transl Comp Linguist, pp. 22-31, 1968.

[25] Turney.P, "Semantic Orientation Applied to Unsupervised Classification of Reviews," proceeding to the 40th Annual Meeting of the Association for Computational Linguistics, Philadelphia, Pennsylvania, USA, pp.417-424, 2002.

[26] Einarsen S, Hoel H, Cooper C, "Bullying and Emotional Abuse in the workplace," International Perspective in Reasearch and Practice, CRC Press, 2002.

[27] Dadvar M, de Jong F, "Cyberbullying Detection : a step toward a safer Internet yard, " proceeding of the 21st International Conference companion on World Wide Web-WWW' 12 Companion, Lyon, France, pp.121, 2012.

[28] Greenwood, Perrin, A., and Duggan, "Social Media Update," Pew Research Center, 11, 2016.

[29] Jünger, Jakob, Keyling, Till, "Facepager, an application for automated data retrieval on the web," Source code and releases available at https://github.com/strohne/Facepager/, 2020.

[30] Kaufer, D, "Flaming: A White Paper, " . Carnegie Mellon, Department of English, June 2000.

[31] Matt Harrison, "Learning Pandas, Python Tools for Data Munging, Data Analysis, and Visualization, " 2016.

[32] Pedregosa et al, "Scikit-learn: Machine Learning in Python," JMLR 12, pp. 2825-2830, 2011.

[33] Joachims, "Text categorization with Support Vector Machines: Learning with many relevant features, " In the Computer Vision—ECCV; Springer Science and Business Media LLC, Berlin, Germany, pp. 137-42, 2018.

[34] WittenIan H, Frank Eibe, Hall Mark A, "Data Mining, Practical Machine Learning Tools and Techniques, " Elsevier, 2011.

[35] Al-garadi MA,Varathan KD,Ravana SD,"Cybercrime Detection in Online Communications, " Comput Hum Behav, pp. 433-443, https://doi.org/10.1016/j.chb.2016.05.051, 2016.

[36] Li YH, Jain AK," Classification of Text Documents, "Comput J, 1989.

Advanced Computing Technologies and Applications (ICACTA), 1877-0509 , pp. 485 – 492 ,© 2015.

[37] Awad M, Khanna R," Efficient Learning Machines: Theories, Concepts, andApplications for Engineers and System Designers**,**" Apress, 2015.

[38] Rounak Ghosh , Siddhartha Nowal , Dr. G. Manju, "Social Media Cyberbullying Detection using Machine Learning in Bengali Language," Internationl Journal of Engineering Research and Technology (IJERT) Volume 10, Issue 05, May 2021.

[39] Nonauharc element.Education.gouv.fr. Non Au Harcèlement, Appelez Le 3020, Available Online: https://www.nonauharcelement.education.gouv.fr/, 2020.

[40] D. H. Patil , Gautami Kharul , Pranjali Gaikwad , Vaishali Khawse, , "Cyber Bullying Detection using SGD Classifier," Internationl Journal of Engineering Research and Technology (IJERT) Volume 10, Issue 05, pp. 587-588, May 2021.

[41] Kowsari K,Meimandi KJ, Heidarysafa M, Mendu S, Barnes LE, Brown DE, "Text Classification Algorithms: A survey. Information," 10:150, https://doi.org/10.3390/info10040150, 2019.

[42] N. Chinchor**,** "Muc-4 Evaluation Metrics, " in Fourth Message Understanding Conference, 1992.

[43] Y. Sasaki, " The truth of the F-measure," University of Manchester, October 2007.

Reem Thabt Ali got her bachelor degree in Informatics Engineering, from Tishreen University, Syria. Currently, she is pursuing her Master degree in Web science program at Syrian Virtual University, and her current work is teaching the practical section in Informatics Engineering College at Syrian Private University and International University for Science and Technology.

Mohamad-Bassam Kurdy is a professor of AI, currently works at the Program in Web Science, Syrian Virtual University, Syria, and RSB Rennes and BSB Dijon France. Mohamad-Bassam does research in Data Mining, Algorithms and Artificial Intelligence.