

DIGITAL FORENSIC ANALYTICS IN SOCIAL MEDIA ENVIRONMENT USING DNN APPROACH

OHOU ALSHABIB¹, RANDA AHMED JABEUR² and FAEIZ MOHAMMED ALSERHANI³

¹Student, Jouf University, Department of ComputerScience, Al Jouf,KSA

²Assistant professor. Jouf University, Department of ComputerScience, Al Jouf,KSA

³Assistant professor. Jouf University, Department of ComputerScience, Al Jouf,KSA

E-mail: ¹421204013@ju.edu.sa, ²rjabeur@ju.edu.sa, ³fmserhani@ju.edu.sa

ABSTRACT

Cyberbullying has increased due to the digital growth of Social Networking Applications “S.N.Apps”. As a result, criminal activities in cyberspace became a cause of concern, particularly towards the start of 2022. Thus, a digital forensic analyst's task of acquiring digital evidence can be challenging. The purpose of this paper is to design an intelligent digital system for analysis by using an Artificial Neural Network “ANN” as a deep learning approach. Additionally, this study is to provide insight into the practical processes required criminal legislation enacted against cybercrimes. In ANN training process, we obtained detection accuracy rates of 99.63% for Twitter, Facebook, and Instagram applications and tested the proposed system with a 94.02% accuracy. And one of the Kali Linux commands was used to speed up access during the Phase I of digital forensic “Gathering Information” regarding the suspect. With the investigation Crowdsourcing system, it allows for better judgment of cyberbullying crime behaviors.

Keywords: *Cyberbullying, Social Media Application (S.M.Apps), Crowdsourcing, Kali Linux, Intelligent-system.*

1. INTRODUCTION

Digital forensic science focuses primarily on gathering, retrieving and seizing digital data[1]. For a statement stating that the information will be presented in court and to ensure that correct results are acquired for detecting cybercrimes, all disciplines related to digital forensics rely heavily on the tools they use during their examinations[2].

The study focuses specifically on cyberbullying. There are concerns regarding privacy in cyberspace that may be adversely affected by the digital world. Due to the challenges of digital forensics, it is challenging for investigators to strike a balance between the results of digital investigations and the protection of privacy. These challenges indicate the future of digital forensics generally[3].

Accordingly, cyberbullying is a phenomenon that has been studied extensively in arabic and other languages. By detecting cyberbullying, can learn more about the adverse psychological effects and impacts on the victims.

Therefore, the recent efforts by researchers to detect cyberbullying are encouraging[4].In addition, digital criminals can spread this crime faster and in shorter amounts of time using the

social media environment. There is a need to collect social media evidence in a legal and respectful Manner privacy rights. As a result, obtaining valid evidence and conducting effective investigations are challenging tasks for legal investigators.

In near real-time, social media profiles of potential suspects and victims can be mined for valuable information. Contacts, messages, geolocation data, photos, and general information appear chronologically in the activity list. The analysis of social media metadata and network data can assist criminal investigations and validate social media evidence. since any-user can create a fake profile on social media and hide under someone else's false identity[5], social media presents exceptional challenges to the authenticity of cyber evidence.

In addition, during COVID-19, there was an increase in the use of electronic learning and longer periods of sitting on devices. Also, cyberbullying and its impact on social media users were mentioned. Indeed, there are significant rates of cyberbullying and the resulting fears and risks, such as stress and depression.

Due to advances in research and awareness of cyberbullying, researchers have been developing new methods to detect cyberbullying by combining deep learning methods with artificial intelligence to Simulate systems that are competitive with humans in detection[4].

Thus, the work will detect cyberbullying across three social media applications “Twitter, Facebook, and Instagram” in arabic language.

1.1. The problem statement:

- How to reduce cyberbullying crime?
- In the social media environment, how does the forensic investigator gain access to the suspect?
- Is deep learning a reliable choice to detect cyberbullying words in arabic text?

1.2. The research questions:

Q1:How to extract cyberbullying arabic words from social media applications?

Q2:How can ANN deep learning algorithm be used to build an intelligent system along with Kali Linux to detect the suspect at a crime scene?

Q3:How can we understand cyberbullying behaviors using crowdsourcing?

1.3. The purpose of the paper:

- Determine the outcomes of cyberbullying using an Artificial Neural Network approach.
- A suspect may be able to be identified by the use of phishing commands in Kali Linux.
- Examining the possible use of crowdsourcing “CRWD” for digital forensic crimes such as cyberbullying schemes in the applications of this study .
- Designing a smart framework that may highly accurate in its detection.

Contribution is the design of smart solution to identify cyberbullying. By using deep learning neural network, may can detect cyberbullying crimes with high accuracy. And will incorporate the proposed approach by creating a cyberbullying scenario and using the CRWD tool to reach the cybercriminal through a Linux phishing command. Assuming the digital investigation is conducted in a forensic workstation OR in local lab legally.

In this study, an identification of cyberbullying words and behaviors for the suspect contributed to the resolution of an analytical difficult for forensic investigators.

This paper is organized as follows:

In section 2, will discuss related works. Then develop an artificial neural network model for the study in section 3. In general, few digital investigations involve finding at least one of the detection methods. Moreover, since most of them

may also include multiple forms and methods as part of social media analysis, also use CRWD tool to improve the detection accuracy. Therefore, law enforcement agencies must stay on top of the latest social media trends in order to identify and leverage the most effective strategies. Furthermore ; in section 4 discuss the solutions result in detail and the approach performance. As a conclusion, limitation and future plans for the investigation of digital forensic evidence in the final section.

2. LITERATURE REVIEW

Researchers daniel.w et.al. provide a “Tracking criminals on facebook: a case study from a digital forensics REU program” [6]. An analysis of a study is conducted by students at the university of Alabama, Birmingham (UAB). In 2014, Facebook worked with us federal law enforcement agencies to create a database of 400 records that contained criminal activity, leading to an investigation. Authors show that the methods, used while investigating the study, may help the examiner and The Facebook application to detect illegal activity between users. Also, the study enabled law enforcement to discover the activities that led to each crime being prosecuted and punished, as well as encouraging students to develop secure social media applications.

Researchers Ankita.R. J , and Avinash.J. A provide a “A digital forensics investigation model for social networking site”[7]. The purpose of this paper is to examine the rapid growth of social media, its impact on users of different ages, and the activities of social media criminals. According to their results, by employing Naive and Bayes methods, the researchers were able to detect malware and determine the concept behind social media messages. The K-rule algorithm was used to track down messages and report suspected cases. A secure and efficient framework was proposed to expose as many criminals as possible.

Researchers Yeboah-Ofori.A , and Brimicombe.A provide a “Cyber intelligence and OSINT: developing mitigation techniques against cybercrime threats on social media”[8]. They demonstrate how open-source intelligence can be used to collect data from Facebook, Emails, and Twitter. Within a social media environment, participants attempted to identify threats and causes of weaknesses. Authors focused on combining simulations of anti-cybercrime measures and finding out which ones were most effective. The study investigated the prevalence of threats within the social media context.

Researchers Farid.D ,and El-Tazi.N focusing on the, “Detection of cyberbullying in tweets in egyptian dialect”[9].The purpose is to build an analytical model to detect cyberbullying, whether it is written in the Egyptian dialect or in Arabic. According to the study, the sentences are classified and trained to detect insults and racism that may offend Twitter users. Based on the classification characteristics of Emojis and Sentiments, this study was conducted. The model was built using a deep learning algorithm to improve performance. Also, authors added improvement factors to achieve an accuracy of 73%.

Researchers Abdulrahman.A , and Baykara.M presented a “fake news detection using Machine learning and deep learning algorithms”[10]. Based on deep learning and machine learning algorithms, The study examined social media applications in a text to detect fake and real news. This study achieved 81% Accuracy in applying deep learning, by using the CNN algorithm. Authors recommended detecting fake messages in the form Of images, such as promoting fake news and spreading it widely.

Finally, the researchers Tahboub.K et.al. proposed “An intelligent CRWD system for forensic analysis of surveillance video”[11]. CRWD was employed to detect fraud using different camera devices and the effects of fraud on detection e.g., blurring of light, decreasing video quality. As a primary objective of the research and investigation, video was recorded. Thus, the adopted Approach proved highly effective in improving performance. Also, authors were suggested to expand their research to include machine learning algorithms in detection.

2.1. Comparison study

The method used in this analysis paper will be compared with that used in the previous analyses discussed in “section 2”, as in table1.

Table1: Comparative analysis of the proposed model

References	Solution used
[6]	Development of an application that simulates the criminal activities of the Federal Law Enforcement Agency.
[7]	Use one of the machine learning algorithms to detect suspicious messages under the scope of malicious programs.
[8]	Developing an OSINT solution to simulate cybercrime.

[9]	Using a deep learning algorithm to detect cyberbullying tweets with a 73% accuracy rate.
[10]	In their research, deep and machine learning algorithms are recommended to detect fake news and the truth, also recommend using deep learning with image recognition to increase its effectiveness.
[11]	Using the crowdsourcing tool to detect fraudulent activities with a focus on implementing machine learning algorithms in the future.

The related works provided various solutions and contributions for social media cybercrime with various applications.

This paper is similar to paper No. [9] it involves deep learning and crowdsourcing as discussed in paper No. [11].

Moreover, papers No. [6, 7, 8, and 10] have different objectives, but all strive to detect and reduce cybercrimes. As in this paper attempts to combine several solutions to speed up the detection process and simulate a smart scenario.

3. METHODOLOGY

Neural network approach is based on biological neural networks, brain analysis and how the nervous system works. A neuron has three layers “Input, Hidden, and Output”. A neurode can be connected to all or a specific group of neurons in the following layer until the network hyperparameters is simulated[12].

Hence ; the proposed system is described in figure1. The data were collected from several ready-to-use tools “TRACKMYHASHTAG, EXPORT COMMENTS, EXPORT INSTAGRAM COMMENTS” websites.

The trained model will contain a set of pre-processing: “Cleaning, Tokenization, Normalization , Stemming, and Stop Word Filtering” until the data is accurately predicted and classified.

Consider the table of this study. As an example the tweet “The Ministry of Health warns from the Ministry of Health” was found on the Twitter account of @SaudiMOH...Ministry of Health. As in figure2 it is cyberbullying tweet, due to there are no warnings except that the ministry of health Mentioned in a tweet its content. “If there are any comments or opinions, we receive your opinions and always respond”.

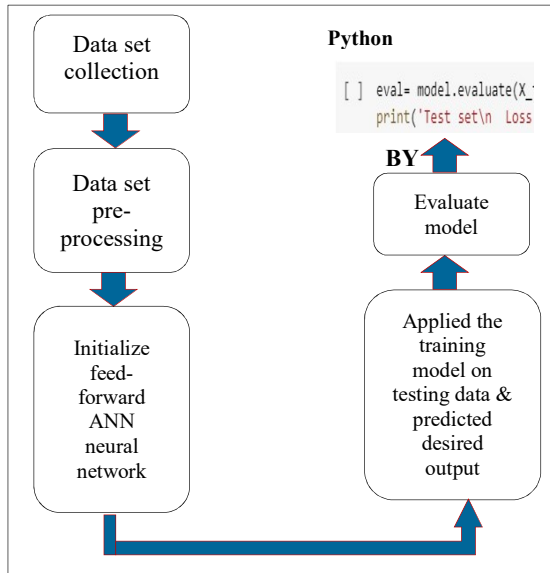


Figure1: Proposed ANN system



Figure2: Cyberbullying study example

The study found there is no warning as the bully words , but rather an awareness from the Saudi Ministry of Health to overcome COVID-19 epidemic.

Trained in hyperparameter , and to improve the quality of the neural network architecture will as the table2 :

Table2: The ANN detail layers

ANN model	
No. of layers	Working on
Layer1	Embedding:: Var Embed size
Layer2	LSTM:: Var Embed size
Layer3	2node, softmax
Layer4	Dropout : 0.2
Layer5	SpatialDropout1D: 0.2

To further clarify, in the ANN approach, the accuracy of the classified system was 94% cyberbullying for the sentence “The Ministry of Health warns from the Ministry of Health”.

3.1. Cyberbullying crime scene

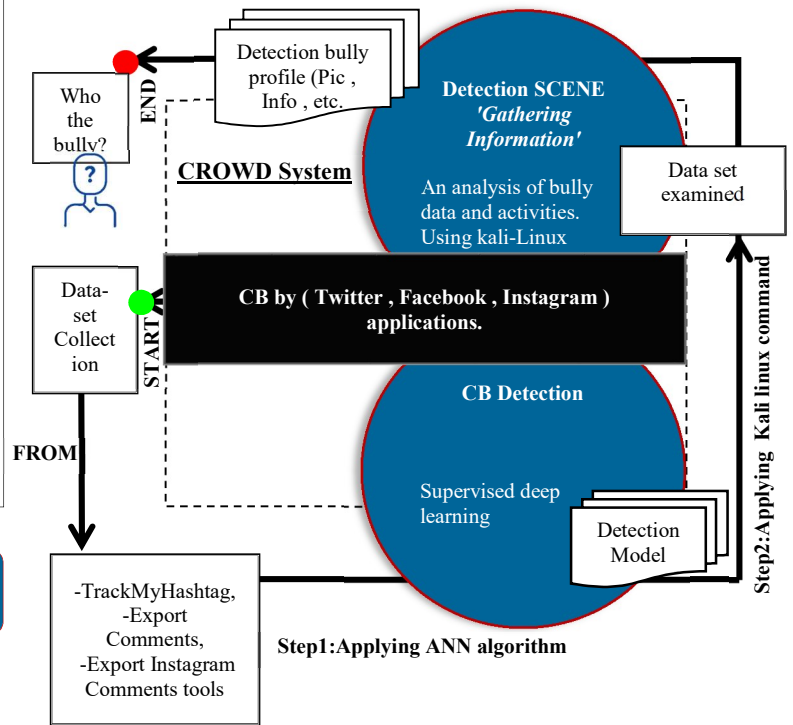


Figure3: Cyberbullying crime scenario

A simulated cyberbullying crime scenario involving social media applications “Twitter, Facebook, Instagram” was shown in figure3. The Crowdsourcing system uses both human and machine intelligence, also trains machine learning models to categorize data with individual feelings [11],[12].

Based on the study, can recommend the following cyberbullying behaviors:

- 4- basic labels; Harassment , Defamation ,Hate , Coprolalia ,
- 1-reserve; Uncyberbullying. “if none of the label choose”.

For detecting bullies and for accessing information about them, we employed an online phishing command to gather information about a bully. A KALI LINUX “Zphisher command”. This command automates phishing and gives us access to the suspect.

The objective case: to gather information about the suspect. As shown in figure4&5, the investigator chooses an application to gain access. Then create a phishing link , send the 'URL' to the

target and monitor the IP address, username, and password of the suspect in the local lab.

4. RESULT AND DISCUSSION

According to the contribution was made to detect cyberbullying crime based on the results of using Artificial Neural Network with the proposed scenario along with a crowdsourcing system. Results as in figure6 reflect the percentages of the last 10 steps of training conducted to detect cyberbullying via social media applications.



Figure4: The number of an applications that is to be phished . An example we choose 08'Twitter'

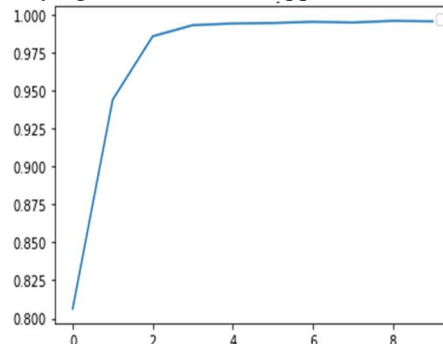


Figure 6 :ANN Trainings 99.63% Accuracy

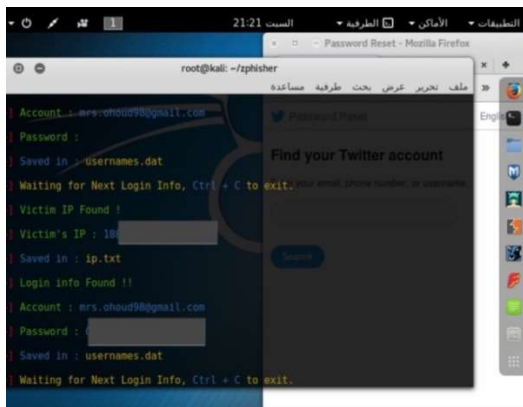


Figure5 : Kali Zphisher tool Version 2.2 A full profile and all data on 08 'Twitter' application number of the bully was accessed in a local lab

The URL formatted as in table3:

<http://get-unlimited-followers-for-Twitter-@drug-refrigerator-championserbia.trycloudflare.com>

Table3: URL phishing format

http://	Protocol identifier
get-unlimited-followers-for-Twitter-@drug-refrigerator-championserbia	Host server URL
trycloudflare.com	Target page or domain name where the : .com : Top level domain trycloudflare : Second level domain

The trained model was contain a set of pre-processing: “Cleaning, Tokenization, Normalization , Stemming, and Stop Word Filtering” until the data is accurately predicted and classified . Hence the validation dataset are 82% as in figure7. These data were used to assess how well hyperparameters fit the training data. Thus a regularization process can be carried out using validation datasets.

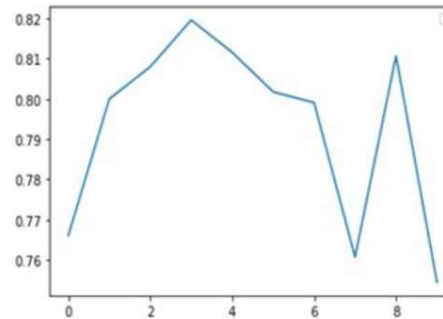


Figure7:Data-set Validation

Moreover, the ANN model does not perform well in certain cases. It was a decent data set, even though the model was properly set up with the loss function and optimizers set correctly. The architecture of the model was also well defined, but it never quite measured up to expectations.

Figure8 , figure9 and table 4 show how the F1 score, random forest, and confusion matrix for study table explained the classification model performance. This provides a holistic evaluation of how well the model performs and the types of errors it makes.

Using the F1 score classifier, the performance of this model was shown to be the highest as it was “80.06 in the training and 78 in the testing process” in comparison with a random forest classifier “71 in the training and 70 in the testing process”.

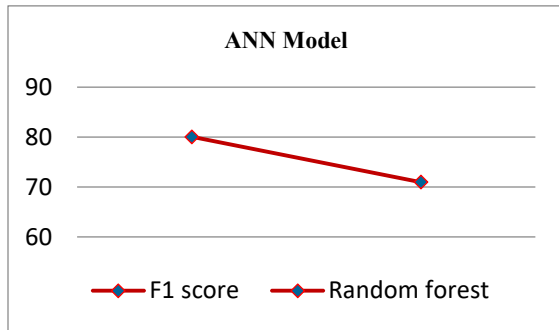


Figure8 :ANN Trainings Performance

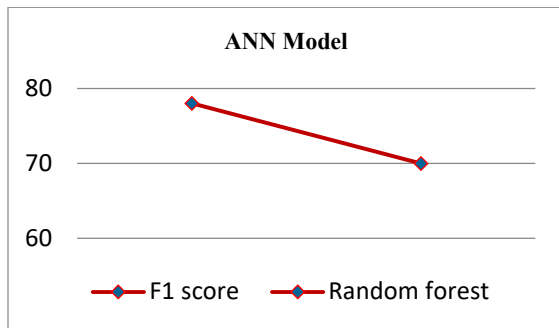


Figure9 :ANN Testing Performance

Table4: ANN model confusion matrix

		Actual values	
Predicted values	Positive	TP=5	FP=49
	Negative	FN=17	TN=38

As shown in the confusion matrix table , True Positive(TP) actually showed a positive value and the model predicted a positive 5 value class only. True Negative(TN) showed a negative value and the model predicted a negative value of 38 class.

This proved to be a misclassified classifier for this dataset, since we had a relatively larger number of true positive and true negative values. False Positive (FP) were negative value 49, but models had a positive value predicted .

Moreover; models showed a positive value, but False Negative (FN) were negative value equal to 17 class by the model.

Regarding the validation of the crowdsourcing tool:

In round one: judges were initially asked to give a label based on the desired classification in the three applications used in the study.

In case of an incorrect choice, a new panel presented and ratings will be reviewed until match the study results. In the final step, judges could select more than one sub-label if needed.

During round two: the figure10 to figure12 shows that the main reason for moving to the second round is that we did not find enough annotation on twitter, facebook, and instagram applications.

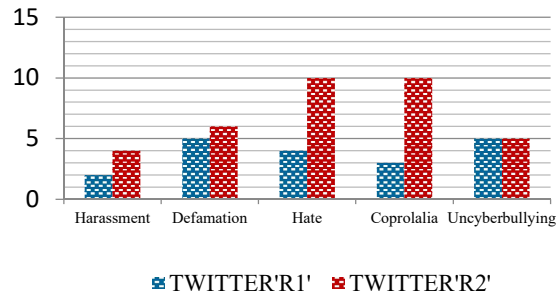


Figure10: Judgment distributions by Cyberbullying behaviors during the two rounds. 'TWITTER DATA'

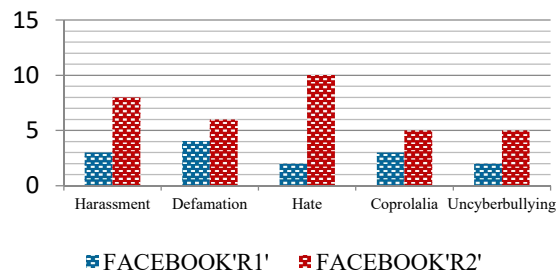


Figure11: Judgment distributions by Cyberbullying behaviors during the two rounds. 'FACEBOOK DATA'

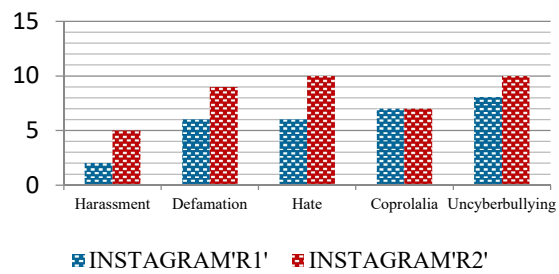


Figure12: Judgment distributions by Cyberbullying behaviors during the two rounds. 'INSTAGRAM DATA'

As a consequence of using the phishing command in the social media applications "Twitter, Facebook, Instagram", we are able to determine the suspect's identity as in table5.

Table4: Suspect gathering information

Name & username	Name: O**** Username: @_M*****	Name & Username : L*** A*****	Name: O**** Username: @_m*****
Email Address	mrs.*****@gmail.com	No confirmed email	mrs.*****@gmail.com
Phone Number	+96650*****	No confirmed phone	+96650*****
Bio & Profile info	Postgraduate researcher at @ju_csi Honor MSc #cybersecurity degree Free soul Life-long learner Ambitious-yet Fullfed by healthy routine Peace🌿.	There are No Bio or any info about the suspect	"عينها ، عينا طفلة" H ealthy lifestyle M sc cybersec🌿 juF-Ksa 🌿🌿
A snapshot of evidence	573tweet- replies – retweet and 119 Photos - videos with 846 likes. Hence; there is NO evidence that the suspect is a bully according to the above data	There are No posts , replies or any likes There is NO evidence that the suspect is a bully according to the account	There are 312 posts, 0 followers and 566 following AND There is no interaction on other pages There is NO evidence that the suspect is a bully according to the interaction
S.M.Apps Name	Twitter	Facebook	Instagram

4.1. Threats of validity approach

There are several threats to overcome, including:

1.Blast of complexity

Expanding the presence of evidence from social media, capacity units, and networks made it necessary to have sufficient expertise, time, and many tools. The speed of the investigation, digital analysis, and the observed pace of progress may be affected in a number of ways, increasing the complexity for forensic specialists.

2.Improvement of benchmarks

To develop a digital forensic evidence community, investigators must collect information from all sources, including internal and external storage, during investigations of cyberbullying.

3.Legitimacy

Due to the physical infrastructure and the challenges it presents, specialists in digital forensics must conduct investigations in a legal manner without violating any policies.

4. Anti-Forensics Techniques

Due to the remarkable advancements in technology, techniques can be used to hide information. Thus, digital forensic evidence must be equipped with tools that will facilitate its work. It is necessary to use the newest techniques to conduct investigations, especially in social media environment.

4.2. Issues of approach

Due to a paucity of research's in the arabic language, the title only refers to arabic language. In addition, considering the complementary aspect of the proposed scenario, the use of the phishing command may violate local laws without legal authorization.

Moreover, this paper only covered Identification, Gathering information, and Analysis, excluding the acquisition and preservation phases of the investigation.

Hence, some necessary suggestions about how to mitigate the issues as follow:

- Protect perishable data.
- Identify network lines attached to the suspect device.
- Identify all individuals “witnesses” at the scene and file their location on the time of cyberbullying.
- Knowing any passwords required to get admission to the social media apps. “A suspect may additionally have more than one passwords”
- Examining any unique protection schemes.
- Discovering any documentation explaining the hardware or software program established at the device.
- observe jurisdictional policy for securing the crime scene.
- Take a snapshot the whole scene to create a visible file by way of the first responder.
- Seizure of suspect hard drive.
- Departments must have crime scene processing equipment.

5. STUDY LIMITATION

- Crowdsourcing tool workers disagreed in the final judgment about building more detailed arabic data. Additionally ,the “Hate” label in the second round only received the closest rate out of the three applications. Hence, it is an indicative of confusion among the judges regarding the different dialects of arabic and the distinction between the labels.

- The “overfitting” as in figure13 effect of training Artificial Neural Network “ANN” more times “train”, deep learning is currently in this paper unable to predict with new data “test”.

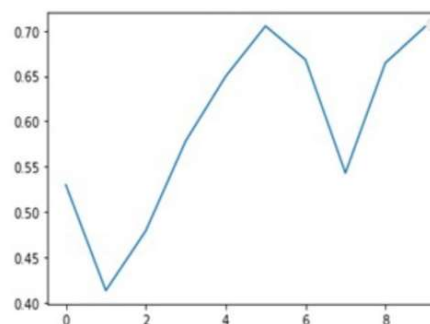


Figure13:..Data-Set overfitting in ANN model

- Only how to track the suspect and what information could gather in tracking was examined in this study. This was achieved by using the Zphisher command on Linux in a local lab. To reach the desired level of security, other digital tools will need to be used in the future.

6. CONCLUSION

The goal of this paper is to investigate how effective ANN algorithm are at detecting cyberbullying through “Twitter, Facebook”, and Instagram applications.

Furthermore, the study included 7k records of cyberbullying and uncyberbullying in arabic language with accuracy of 99.63% and evaluation accuracy of 94.02%.

To clarify the extent to which the words “Coprolalia, Harassment, Defamation, Hate, and Uncyberbullying” are in arabic, we expanded the paper vision using the CRWD tool.

Moreover, to build a logical analysis, the system was developed using phishing command in Kali-linux.

The study suggests that the acquired knowledge should not be used to detect text data in the future, since such an approach will adversely affect any new data tested against other data types. Additionally, the paper recognizes the need to gather information about the network lines connected to the suspect device. Including any documentation explaining the hardware or software installed on the device as evidence of the investigation.

Lastly, we contacted two of the corporations, ¹ and ², as guides for this paper. This approach to gathering information was recommended as valid.

And according to the comments:

First, this study has proved to be successful, and that further digital tools will be needed to enhance the level of investigation through social media.

Second, provide a summary of sponsors comments about how establishing an acceptable methodology in this field is difficult. And the types of evidence and method used in the field are diverse and cannot be limited to different policies and countries.



STUDY ETHICS :

Authors disclaim any responsibility for misuse of the Zphisher phishing tool. Researcher *OHOUD conducted an analytical study in order to maintain users security.

Sponsors : Digital Forensic Corporation¹ and Secure Techware Corp²

7. FUTURE WORK

In the future, looking forward to this approach will be hybridized with optimization and ML algorithms.

Also, testing this approach on real-life datasets is a future contribution.

This paper was success as a first step in a digital forensics investigation by conducting logical analyses, with the need to conduct an analysis and collect a logical image of the perpetrator's hard drive data as a second step.

DEDICATION

The authors would like to thank the Deanship of Graduate Studies at Jouf University for funding and supporting this Research through the initiative of DGS, Graduate Students Research Support "GSR" at Jouf University, Saudi Arabia.

REFERENCES

- [1] Fran Casino, Thomas K. Dasaklis , Georgios P. Spathoulas , Marios Anagnostopoulos, Amrita Ghosal , István Bořočz , Agusti Solanas , Mauro Conti , and Constantinos Patsakis, "Research Trends, Challenges, and Emerging Topics in Digital Forensics: A Review of Reviews", *IEEE Access*, February 24,2022, Vol.10 ,pp. 25464 -25493.
- [2] GraemeHorsman , "Tool testing and reliability issues in the field of digital forensics. Digital Investigation", *Forensic Science International: Digital Investigation*, March ,2019,Vol.28,pp. 163-175.
- [3] Stefania Costantini , Giovanni De Gasperis , and Raffaele Olivieri , "Digital forensics and investigations meet artificial intelligence", *Annals of Mathematics and Artificial Intelligence* ,April,2019,Vol.86,pp.193-229.
- [4] Ibrahim Baggili and Frank Breitingner , "Data Sources for Advancing Cyber Forensics: What the Social World Has to Offer", *AAAI Spring Symposium*,2015, pp.6-9.
- [5] Muhammad Firdaus , "Forensic Analysis of Social Media Data : Research Challenges and Directions", *Journal of Research In Science Teaching* ,December ,2020,pp.1-8.
- [6] Daniel Weiss,and Gary Warner , "Tracking Criminals on Facebook: A Case Study From A Digital Forensics REU Program", *ADFSL Conference on Digital Forensics, Security and Law*, May19, 2015,pp205-214.
- [7] Ankita R. Jadhao , and Avinash J. Agrawal , "A Digital Forensics Investigation Model for Social Networking Site", *ICTCS '16: Proceedings of the Second International Conference on Information and Communication Technology for Competitive Strategies*, March 2016 , No.130,pp.1-4.
- [8] Abel Yeboah-Ofori , and Prof. Allan Brimicombe, "Cyber Intelligence and OSINT: Developing Mitigation Techniques Against Cybercrime Threats on Social Media A Systematic Review", *International Journal of Cyber-Security and Digital Forensics (IJCSDF)*, July, 2017,No.7, Vol.1,pp.87-98.
- [9] Dina Farid, and Neamat El-Tazi , "Detection of Cyberbullying in Tweets in Egyptian Dialects", *International Journal of Computer Science and Information Security* , July, 2020, No.18,Vol.7,pp.34-41.
- [10] Awf Abdulrahman , and Muhammet Baykara , "Fake News Detection Using Machine Learning and Deep Learning Algorithms", *2020 International Conference on Advanced*



-
- Science and Engineering (ICOASE)* , Duhok-Iraq, May 31 ,2021,pp. 13773- 13781.
- [11]Bhumika Bhatia, Anuj Verma, Anjum, and Rahul Katarya, “Analysing Cyberbullying using Natural Language Processing by Understanding Jargon in Social Media”, *Sustainable Advanced Computing - Select Proceedings of ICSAC 2021* , April,23, 2021,pp.1-10.
- [12] Minakshi Sharma ,and Sourabh Mukharjee, “Brain Tumor Segmentation Using Genetic Algorithm and Artificial Neural Network Fuzzy Inference System (ANFIS) ”, *Advances in Computing & Inform. Technology, Part of the Advances in Intelligent Systems and Computing book* , Vol 177, pp. 329–339.