Fake Accounts Identification in Mobile Communication Networks Based on Machine Learning

https://doi.org/10.3991/ijim.v17i04.37645

Ahdi Hassan[1](✉), Abdalilah. G. I. Alhalangy[2], Fahad Alzahrani[3]
[1] Global Institute for Research Education & Scholarship, Amsterdam, Netherlands
[2] Department of Computer Science, Qassim University, Buraydah, Saudi Arabia
[3] Department of Languages and Translation, University of Tabuk, Tabuk, Saudi Arabia
ahdihassan441@gmail.com

Abstract—Fake accounts on online social networks are increasing today with an increase in the number of active social network users. Social media websites allow users to share thoughts, facts, views and re-sharing these into various networks. Social media platforms provide users with enormous valuable information but with this great amount of information in social media, many issues like fake profile, online hacking have also grown. The fake profiles in online social sites create fake news and share unwanted material which contains spam links that affect natural users. The massive issue in social media communication networks is spam and it is necessary to identify fake profiles to stop spam. In this paper, a supervised machine learning algorithm called support vector machine (SVM) is used to identify fake accounts on social media effectively. In order to automatically identify fake online profiles, Random Forest classifier is used with SVM. With this concept, it can be applied online easily to identify millions of accounts that cannot be examined manually. The result of this model is compared with other identification techniques and the results show that the proposed algorithm performs better with high precision and recall. This method efficiently safeguards social media networks from online threats and attacks.

Keywords—social media, SVM, random forest classifier, fake profile, mobile network

1 Introduction

Online social networks like Facebook, Instagram, Twitter, and Linkedin are becoming widely prevalent in the last few years. It is becoming more and more challenging for social media to identify fake accounts through manual inspection as the volume of content on the platform is expanding so quickly. The fake accounts on various social media may mislead the readers, provoke social panic, and even create violence, which could be prevented by using early identification methodology to timely detect fake accounts on social media. The fake profile people purposely design content to gather attention of other users. Due to the social network's qualities of

quick distribution and low cost, a significant amount of news content is disseminated there quickly [1].

It has been a concern for users that as the usage of social networks grows, bad users would try to violate other users' privacy and create phony profiles using their names and login information. Therefore, in order to remove malicious individuals and fake accounts from social networking environments, social network service providers are attempting to identify them. More crime is done by making fake accounts on social networks than by any other type of cybercrime [2]. Through the online social networks, people can communicate, share information, plan organization, and even run their online businesses.

Online social networks have an impact on various domains like science, education, business, employment, etc. The common technique followed by the fake accounts on social media to easily attract users is catchy headlines. The preprocessing of the dataset has been done to determine false profiles on mobile social communication sites. Random classifiers with support vector machine classification results are utilized to find fake accounts. Using a machine learning system to compare the precision rates of phony accounts, the method with the highest accuracy is suggested [3].

1.1 Threats

Due to widespread use of online social networks, many users are vulnerable to both privacy and security issues. These threats can be divided into four main categories: Conventional threats, Modern threats, Combination threats, and Targeted threats as shown in Figure 1.

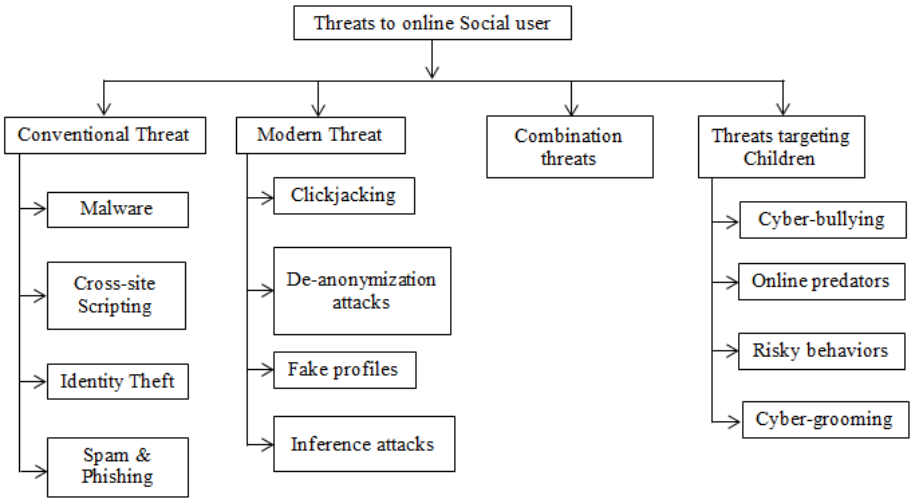

Fig. 1. Various Threats to Users of Online Social Network by Fake Accounts

Conventional threats. The prevalence of the internet has raised concerns about conventional threats. Threats involving malware, spam, phishing, identity theft, and cross-site scripting appear to be a recurring issue [4]. Due to the structure and prevalence of online social networks, these threats have gone viral and could spread quickly among network users.

Malware. Malware is software designed to interrupt a device's operation in order to record user passwords and access your personal information. The online social network framework is used by social network malware to proliferate among users and their network companions.

Cross-site scripting. It is a web based attack. The hacker who utilizes cross-site scripting abuses the trust of the site user and installs spyware on the user's mobile or computer to collect private information.

Identity theft. The attacker steals another person's identity by using their social security number, contact number, and address without their consent.

Spam & Phishing. Unwanted bulk electronic messages are termed as spam. Email is the regular technique to disseminate spam; social networking sites are highly effective in disseminating spam. Phishing is a type of social engineering attack where the attacker can obtain personal and sensitive data like username, password, and credit card details of people via fake accounts on emails and websites which appears to be real [5].

Modern threats. Modern threats are assaults that penetrate user accounts using cutting-edge methods, while targeted attacks are those that are directed at a specific user and can be carried out by any user for a variety of personal reasons.

Clickjacking. The practice of "clickjacking" involves tricking a person into clicking on a page other than the one he intended to.

De-anonymization. In order to reveal a user's true identity, deanonymization attacks use methods such as cookie monitoring, network design, and community affiliation.

Fake profiles. OSNs replicate human behavior using active or semi-automatic profiles, also referred to as styles or social bots. The use of bogus accounts is another way that users of social networking sites may be approached for personal information.

Inference attacks. The inference attack uses other statistics that the user posts on a social networking site to infer confidential information of users.

Combination Threats. The combination of conventional and modern threats to create a more complex threat is known as combination threats.

Targeted Threats. Children and teens using online social networks are affected through these targeted threats.

Cyberbullying. It is the utilization of e-media such as chats, emails, mobile chats, and online social networks to bully a user. Unlike classic bullying, it is continuously maintained through social media.

Online predators. Cyberpredators, often known as child predators online, are the biggest threat to the privacy of children's private information.

Risky behaviors. Children may engage in risky behaviors such as Internet interaction with foreigners, the use of chat boards for foreigner meetings, chats with

foreigners that are sexually suggestive, and sharing private information and pictures with foreigners.

Cyber grooming. The primary goal of cyber grooming is to win the belief of the child so that the child will provide sensitive and personal information.

1.2 Objective

Number of mobile users is increasing currently and they can easily access social media with mobile in their hand. Fake accounts are created by some person to target this large group of users with an intention of causing harm. The main objective of this paper is to develop a framework that automatically finds the online fake accounts or profiles. Support vector machine learning technique is combined with random forest for automatic detection of fake profiles. It can secure the users social life and with this automatic detection method, the websites can manage the enormous number of profiles, which is not manually possible.

2 Literature review

Numerous studies have focused on removing fake accounts, hence in-depth studies on identifying false accounts in social networks have been conducted.

[6] Proposed a spam recognition AI technique for Twitter sites. In this work, the author employed an ANN, vector support machine and a random forest method to create a technique. The outcomes are compared with RF and ANN techniques; the proposed SVM algorithm has the high precision, recall, and F-measure. This result is used in managing and tracking social media public images for the detection of offensive material and fake photos, as well as to protect social media from online threats and attacks.

[7] Suggested a set of features by using a famous machine learning algorithm called support vector machine and neural networks. The system is built with the aim of identifying fake users of twitter social network. The accuracy is maintained in this work in recognizing fake accounts by various classification algorithms. After using the suggested features with consistent heaviness, the outcome displays the highest accuracy of the two classification algorithms.

[8] Came up with a system that recognizes fake profiles automatically with high efficiency. This system uses random forest classification techniques to separate the profiles into original and fake profiles. This automatic detection method is useful for millions of social media network profiles that cannot be identified manually.

Identifying fake profiles in LinkedIn is expressed in [9]. First the authors of this paper identified a minimal set of data profiles essential for recognizing fake profiles in LinkedIn, and suggest a proper data mining method for duplicate profile identification. Even with the limited data profile, their method identified fake profiles with eighty seven percent accuracy and ninety four percent true negative rates while comparing to larger data profiles. But the result provides only 14% accuracy approximately when compared to methods using same amounts and data types.

[10] Presented a probable approach to alleviate the threat of the fake profile attack, where an attacker tries to copy a victim on an online social network where the user doesn't have a profile in place.

3 Methodology

Collecting the dataset, pre-processing it, choosing features, and applying machine learning techniques to it are the proposed tasks. The suggested system's architecture is depicted in Figure 2.

Selection of Dataset: The main aim of this paper is to use the proposed algorithm SVM-RF to decide if the accounts identify as true or fake. The instagram dataset is considered in this research and collected from kaggle website. The proposed SVM-RF algorithm has the ability to appropriately categorize the accounts of the training dataset. The training dataset contains data pre-processing which includes feature extraction and machine learning technique. After applying the algorithm it identifies whether this model account is true or not.

Data Pre-processing: To confirm that the system identifies the input and generates the finest feasible model, the collected data must be preprocessed via various steps before entering every classifier.

Tokenization: Tokenization is the process of breaking down a text-based system into individual tokens, such as words, phrases, symbols, or other fundamental parts. Exploring sentences as a single phrase is the main goal.

Stop words removing: Stop words are more general than traditional phrases like “are”, “and”, and more. They don't seem to be relevant to the basis of the data gathered. Therefore, they must be removed.

Stemming: In order to achieve this goal more frequently than not correctly, stemming is a simple spontaneous technique that removes the ends of words. It typically involves the elimination of prefixes and suffixes, which happens regularly in English.

Feature Selection: User-based features, or traits that are particular to each user, are used to describe the behavior of Instagram users. Content-based feature characteristics are connected to user accounts. Duplicate content cannot be shared by regular users, however, scammers frequently share duplicate contents.

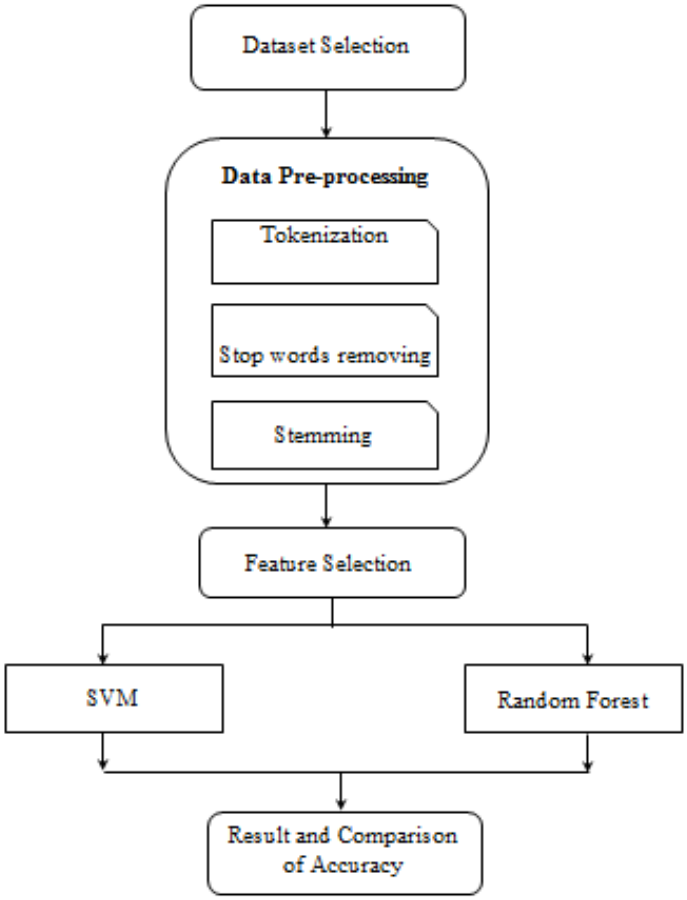

Fig. 2. Proposed System Architecture

Support Vector Machine: SVM is a supervised algorithm of machine learning. The approach to data grouping, training, and issues identification was identified by SVM as one of the most fundamental and practical strategies. It is a simple classification model and computed by using the given equation.

$$s(x) = sign[\sum_{m=1}^{n} \quad \delta_m s_m \varphi(x, x_m) + r] \tag{1}$$

Where δ_m is the "positive real constant" and r is the "real constant".

The SVM increases the margin by locating the ideal feature space. The margin is calculated and then doubled by taking into account a hyperplane and computing its separation between the vectors.

Random Forest Classifier Algorithm: Classification and regression issues can be solved with random forests, commonly referred to as random decision forests. With random forest classifiers, many trees' predictions are combined. Many decision trees are produced using the random forest method. A portion of the features are used to

construct each decision tree. To increase the accuracy of the Random Forest technique, each decision tree establishes a single class and eventually bootstraps the votes. A choice will be made at any node. The steps are as follows:

1. Divide the data so that the total amount of data is gained.
2. The node's state with the highest gain for dividing is picked.
3. Splitting continues until entropy is higher than zero.
4. After repeating the above procedure several times, a class is finally selected for a sample based on significant voting.

4 Result & discussion

The text-based Instagram dataset found from kaggle website. Totally 120 details are there in the testing dataset. The first 10 real and 10 fake profiles are mentioned in Table 1.

Table 1. Instagram Dataset Sample

profile pic	nums/length username	fullname words	nums/length full name	name=user name	description length	external URL	private	No. of posts	No. of followers	No. of follows	Fake or real
1	0.33	1	0.33	1	30	0	1	35	488	604	0
1	0	5	0	0	64	0	1	3	35	6	0
1	0	2	0	0	82	0	1	319	328	668	0
1	0	1	0	0	143	0	1	273	14890	7369	0
1	0.5	1	0	0	76	0	1	6	225	356	0
1	0	1	0	0	0	0	1	6	362	424	0
1	0	1	0	0	132	0	1	9	213	254	0
1	0	2	0	0	0	0	1	19	552	521	0
1	0	2	0	0	96	0	1	17	122	143	0
1	0	1	0	0	78	0	1	9	834	358	0
0	0.05	1	0	0	0	0	0	0	0	2	1
1	0.27	1	0	0	0	0	0	0	45	64	1
0	0.07	1	0	0	0	0	0	0	19	30	1
0	0	1	0	1	0	0	0	0	69	694	1
0	0	2	0	0	0	0	0	0	22	82	1
0	0.22	0	0	0	0	0	0	0	31	124	1
0	0	3	0	0	0	0	0	0	9	25	1
0	0	1	0	1	0	0	0	0	69	694	1
0	0	1	0	0	0	0	0	0	23	33	1
0	0.62	1	0.4	0	0	0	0	1	17	34	1

Evaluation Parameters: The legitimacy of positive (P) and negative (N) data is discussed in this section (N). Hacking is referred to as "hit or positive in reality" (TP), approved in "reality as negative" (TN), and approved fake sites incorrectly as a "false positive" (FP) or "false hit" (FP). The ratio of the classified examples profile over the entire profile number, as indicated in Equation, determines accuracy.

$$Overall\ Accuracy\ (\%) = \frac{TP+TN}{TP+FP+TN+FN} \tag{2}$$

Precision is the percentage of returning hits that were true positive or correct hits.

$$Precision\ (P) = \frac{TP}{TP+FP} \tag{3}$$

True positive recalled amount, i.e. how many accurate hits were also discovered.

$$Recall = \frac{TP}{TP+FN} \tag{4}$$

Table 2. Accuracy Results Obtained Through

S. No	Algorithms	Accuracy	Precision	Recall
1	Logistic Regression	96%	93%	94%
2	Artificial Neural Network	94%	93%	92%
3	Naïve Bayes	90%	88%	89%
4	SVM with Random Forest	98%	97%	98%

When comparing the proposed SVM-RF algorithm to the logistic regression, artificial neural network, and Naïve Bayes algorithm, the proposed algorithm has the highest accuracy, precision and recall.

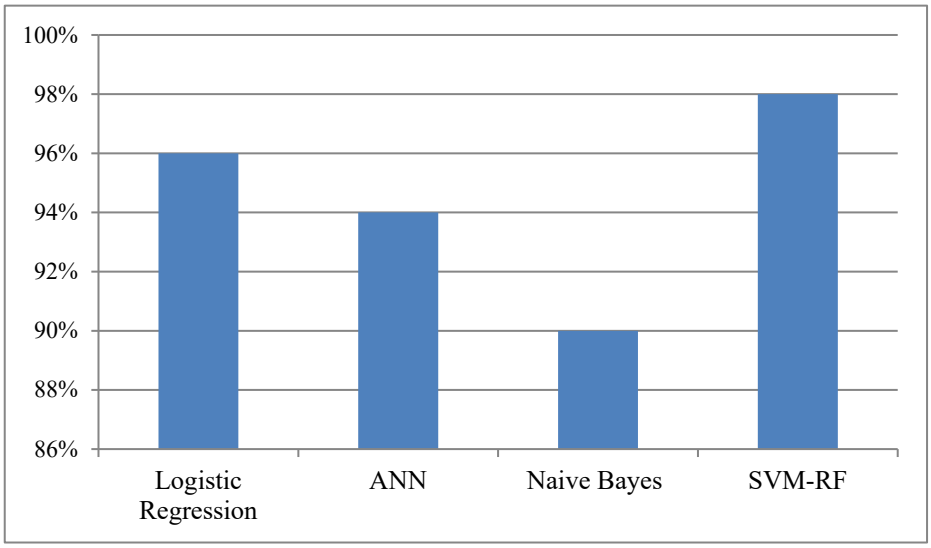

Fig. 3. Accuracy Comparison

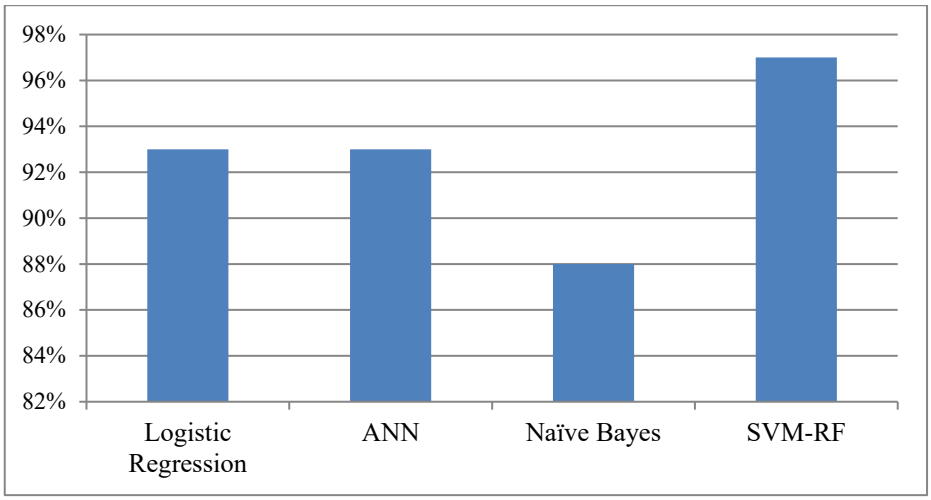

Fig. 4. Precision Graph Comparison

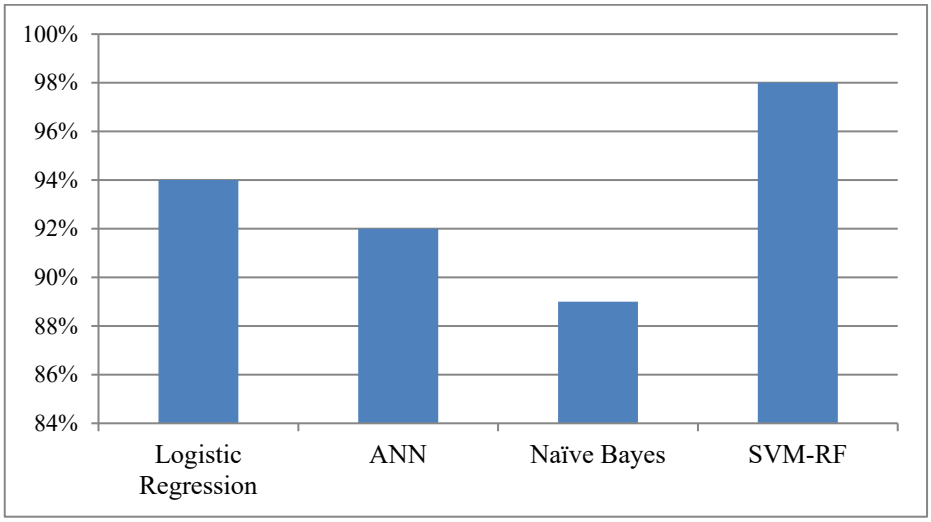

Fig. 5. Recall Graph Comparison

Amongst the four classification methods compared, support vector machine learning algorithm with random forest provides the highest accuracy of 98%, precision of 97%, and recall of 98%. The proposed classification algorithm automatically identifies fake accounts in online social networks.

5 Conclusion

The proposed work gives an analysis of essential approaches for identifying fake accounts on online social network websites. In this paper, the Instagram dataset is considered for instance. Consumers have recently found it more and more challenging to locate reliable information due to the vast amount of information available on social media sites. The SVM-RF machine learning model developed to separate an input as true or fake in order to combat the growing fraud on the internet. Many social media platforms are working to integrate these technologies into their platforms in order to stop the spread of fraudulent activities. The accuracy, precision, and recall values of the proposed model is compared with logistic regression, artificial neural network and Naïve Bayes and shown that the proposed algorithm automatically identifies fake accounts better than all three.

6 References

[1] Song, C., Ning, N., Zhang, Y., & Wu, B. (2021). Knowledge augmented transformer for adversarial multidomain multiclassification multimodal fake news detection. *Neurocomputing*, *462*, 88-100. https://doi.org/10.1016/j.neucom.2021.07.077

[2] Mohammadrezaei, M., Shiri, M. E., & Rahmani, A. M. (2018). Identifying fake accounts on social networks based on graph analysis and classification algorithms. *Security and Communication Networks*, *2018*. https://doi.org/10.1155/2018/5923156

[3] Hemeida, A. M., Alkhalaf, S., Mady, A., Mahmoud, E. A., Hussein, M. E., & Eldin, A. M. B. (2020). Implementation of nature-inspired optimization algorithms in some data mining tasks. *Ain Shams Engineering Journal*, *11*(2), 309-318. https://doi.org/10.1016/j.asej.2019.10.003

[4] Kumari, S., & Singh, S. (2015, April). A critical analysis of privacy and security on social media. In *2015 Fifth International Conference on Communication Systems and Network Technologies* (pp. 602-608). IEEE. https://doi.org/10.1109/CSNT.2015.21

[5] Fire M, Goldschmidt R, Elovici Y (2014). Online social networks: threats and solutions. IEEE Commun Surv Tutorials 16(4):2019–2036. https://doi.org/10.1109/COMST.2014.2321628

[6] Prabhu Kavin, B., Karki, S., Hemalatha, S., Singh, D., Vijayalakshmi, R., Thangamani, M., ... & Adigo, A. G. (2022). Machine learning-based secure data acquisition for fake accounts detection in future mobile communication networks. *Wireless Communications and Mobile Computing*, *2022*. https://doi.org/10.1155/2022/6356152

[7] Kasliwal, N., Bachhav, T., Sonavane, D., Shinde, S., & Nivangune, M. (2019). Detection of fake accounts of Twitter using SVM and NN algorithms. *IEEE Transactions on dependable and secure computing*, *5*(1), 37-48.

[8] Reddy, S. D. P. (2019). Fake profile identification using machine learning. *International Research Journal of Engineering and Technology (IRJET)*, *6*(12), 1145-1150.

[9] Adikari, S., & Dutta, K. (2020). Identifying fake profiles in linkedin. *arXiv preprint arXiv:2006.01381*.

[10] Conti, M., Poovendran, R., & Secchiero, M. (2012, August). Fakebook: Detecting fake profiles in on-line social networks. In *2012 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining* (pp. 1071-1078). IEEE. https://doi.org/10.1109/ASONAM.2012.185

7 Authors

Ahdi Hassan, he is serving as a Researcher at Global Institute for Research Education & Scholarship, Amsterdam, Netherlands, Commissioning Editor, IGI Global publisher publishes more than 170 journals quarterly and semi-annually, Researcher, "Vanishing Languages and Cultural Heritage", Austrian Academy of Sciences, and representative of Imperial English UK-A Trusted British Brand in English Language, Independent Research International [IRI] and Advisor Scholarly Journal Management. He has been Associate or Consulting Editor of numerous journals and also served the editorial review board from 2013- to till now. He has a number of publications and research papers published in various domains. Founder of Pakistani languages corpora and has earned his master's degree in Linguistics from Quaid-i- Azam University, Islamabad in 2013. He has given contribution with the major roles such as using modern and scientific techniques to work with sounds and meanings of words, studying the relationship between the written and spoken formats of various Asian/European languages, developing the artificial languages in coherence with modern English language, and scientifically approaching the various ancient written material to trace its origin. He teaches topics connected but not limited to communication such as English for Young Learners, English for Academic Purposes, English for Science, Technology and Engineering, English for Business and Entrepreneurship, Business Intensive Course, Applied Linguistics, interpersonal communication, verbal and nonverbal communication, cross cultural competence, language and humor, intercultural communication, culture and humor, language acquisition and language in use (email: ahdihassan.41@gmail.com, Orcid: https://orcid.org/my-orcid?orcid=0000-0003-1734-3168).

Dr. Abdalilah Alhalangy is an assistant professor of information systems. He got a bachelor's degree in information technology from Al-Sharq Private College, a master's in information technology, and a PhD in information systems from Al-Neelain University in Sudan. He has taught at the level of higher education in Sudan (University of Kassala, Faculty of Computer Science and Information Technology) and Saudi Arabia (Qassim University) since 2006. At the University of Kassala, Sudan, he held several positions. He taught courses in the departments of computer science, information technology, and information systems (email: a.alhalangy@qu.edu.sa, ORCID ID: https://orcid.org/0000-0003-2735-8208).

Dr. Fahad Alzahrani is a university professor of applied linguistics with educational backgrounds in linguistics and computer science. His major research interests fall under discourse analysis, computational linguistics, natural language processing, and corpus linguistics (email: fahad_alzahrani7@hotmail.com, ORCID: https://orcid.org/0000-0002-4270-8598).

Article submitted 2022-11-24. Resubmitted 2023-01-04. Final acceptance 2023-01-06. Final version published as submitted by the authors.