International Journal for Multidisciplinary Research (IJFMR)
E-ISSN: 2582-2160 ● Website: www.ijfmr.com ● Email: editor@ijfmr.com

Social Engineering 2.0 Deepfake and Deep Learning-Based Cyber-Attacks (Phishing)

Siva Krishna Jampani

Software Engineer

ABSTRACT

This research paper examines the emerging risk of deepfake phishing, or modern social engineering, where deep learning techniques are used to produce realistic fake media for cyberattacks. The paper investigates the topic of deepfakes: voice and video manipulation which have become widespread tools to mislead the target audiences and bring severe financial and organizational losses. The study relies on secondary sources and provides trend analysis, case and quantitative data to demonstrate increase in deepfake phishing attacks in the finance, healthcare and technology sectors. By considering sectoral data, the paper determines the biggest focused areas, which stated, finance and healthcare have the highest proportion of the financial damages. The present research also explores the development in the technology of deepfake and explains how the technology has improved the authenticity of fake videos and audio by providing examples of how these technologies have advanced from simple imitations to near-photorealistic replicas of real voices and faces. What consequences do have this progression for the rate of successful phishing attacks? The answer is quite serious – with this progression, the attackers can act more freely as they do not need to overcome conventional security barriers. Finally, the paper urges people to fight these risks using AI tools for time-independent deep fake detection, stricter guidelines and specific informative training for employees for such shiny attacks. Through exploring the current state of deepfake phishing as a threat, the work underscores the importance of continuing caution and development in cyber defence. The research indicates that to address the threats posed by deepfake phishing one needs to employ both, developed technological solutions and actively maintained organizational shields.

Keywords: Deepfake phishing, Social engineering, Cyberattacks, Voice and video manipulation, Finance sector cybersecurity, Healthcare sector cybersecurity, Deep learning, AI-based detection, Cyber defense strategies, Organizational training, Financial losses, Phishing trends, Technological advancements in deepfakes, Cybersecurity awareness

1. INTRODUCTION

In the world of cyber criminals, social engineering has always been identified as one of prevailing factors which use people's psychology to convince them to release secret data or do something catastrophic. Historically, there was a clear and strong focus on phishing which is a particular type of social engineering attack. Phishing works as a trickery of different identities to make the users release the right passwords, financial details among other things [1]. However, in the current world, where new threats are being developed, and older ones are expanding variations of phishing have become much more sophisticated and dangerous.

This new wave relies on deepfake and deep learning techniques that enhanced the efficiency and dangerous impact of phishing to the unknown level. The combination of these progressive technologies with commonplace social engineering strategies has clearly evolved the threat paradigm, raising new problems for concerned parties, including individuals, organizations, and those in cybersecurity. Deepfake technology when enhanced by deep learning mechanisms such as GANs fosters enhanced and convincing audio, video and image fakes [2]. These synthetic media artifacts have quickly migrated from something that was only seen in theoretical work to a current risk. The deepfake attack has been extremely effective especially in trust manipulation, perception manipulation and evasion of traditional detection [3]. When it comes to phishing, deepfake helps the attackers mimic the genuine identity of people or firms that can be trusted. For example, adversaries can create botnet videos, calls, or audio messages from the company's executive leaders and compel the staff to release their monetary funds or organization secrets [4]. This sort of attacks is as realistic and unsuspected as the audiovisual communication is trusted to be without vulnerabilities. The psychological effects of deepfake phishing are staggering, thereby failing the traditional approach of security awareness training for humans scroll roller. New to the list is the combination of deepfake technology and deep learning in sharpening phishing attacks [5]. This type of neural networks, using big data sets, can identify users' behaviours patterns, preferences and weaknesses. This enables the perpetrator to make the best and well-deserving crafted phishing messages in respect to an individual or a certain group of people. A remarkable feature expanse the possibility of generating local and grammatical correct phishing emails due to the integrated natural language processing options. Moreover, deep learning algorithms can learn from big data and decided which of them are more valuable and how they will react to the certain stimulation.

This level of sophistication has turned phishing from what was just an impersonal and haphazard practice to a precise and very efficient type of cyber-crime [6]. The integration of deepfake-generated creative content with unique phishing content generates a credible and diverse threat that is difficult to deal with using typical approaches. The emergence of Social Engineering 2.0 also implies a serious shift in the whole spectrum of threats. Deepfake-based phishing is resistant to conventional prevention measures such as email filter and blacklist. This is because these technologies work through rules and heuristic models that cannot adapt quickly enough to the ever-evolving threat landscape defined by deep learning. Furthermore, it raises a problem of C2: Reliance with Synthetic Media in Phishing Campaigns where the difficulty of detecting the authenticity of produced AV content can be perceived as hard. Biometric data to be precise, voice and face identification as a protection criterion is also vulnerable to deepfake technology [7]. Thus, the efficiency of measures, which are currently employed to protect a company's infrastructures, is becoming questionable, which has led to a shift in paradigm regarding the problem of cyber security. It is quite significant to understand that the financial and image loss due to deepfake phishing attack are beyond imagination. Such attacks cost organisations not only money, but also the reliability of the customers, the partners, and the stakeholders. Prominent examples are deepfake-generated voice scams that targeted people and organizations and that have cost all the victims tens of millions of dollars, at the very least. Besides, the material loss, the mental trauma inflicted on victims should not be omitted. In deepfake phishing, employees targeted by the criminals go through guilt and fear, and companies are left with reputational losses and legal problems [8]. Such attacks have implications which transcend the singularity of an organization since they cast doubt on the efficiency and even credibility of whole sectors.

Some aspects of Social Engineering 2.0 can be best explained by research from the psychological and sociological fields in addition to technical findings. In this paper, deepfake phishing attacks are modelled where the data must leverage cognitive bias, trust aspects, and social prestige.

For example, attackers seek social engineering by employing potent Ards such as aiming at organization's officials and individuals who possess vital information about the organization and who can manipulate other people in the organization [9]. Trust is especially crucial here; in a world where deepfake technology is rapidly blurring the line between the real and the fake, the roles of formal and informal media trust indicators become even more nuanced. This erosion of trust affects society in general, because the overall index of people's trust in personal communication has plummeted, and everyone is paranoid about the implications of digital communication. The fight against Social Engineering 2.0 requires strategies targeting the technological side of the threat and strategies targeting human element. Technologically, therefore, there is a need to improve on the detection mechanism, for instance, deepfakes detection algorithms, and AI-aided phishing filters. These technologies take advantage of artificial learning incorporating elements of machine learning and deep learning to draw out features of malicious activities. For example, deepfake threat-detecting algorithms check fake media for certain failings that laypersons will not even consider, like the sleepy eyes in a person's lid or the noise that is not there [10]. Phishing filters based on AI is also similar in that, they use natural language processing to analyse messages for contextual inconsistency or what is suspicious. However, the previously mentioned measures are slightly effective as the techniques of the cyberattack are constantly developing, and new ones are appearing from time to time.

The human element in minimizing the success of said threat. Current cybersecurity awareness and training solutions must be adapted to overcome new threats represented by deepfake and deep learning. Recurrent training just to name the more well-known indicators of phishing is inadequate to prepare the employees to handle the more-stringent threats. It is for this reason that training programs for employees should point towards; critical thinking, suspicion, and verification of shrink's identity from other means [11]. There is also the need to cultivate an organizational cybersecurity culture that will make employees to volunteer information of any violation deemed provocative without being penalized. Further, active cooperation of the industry, academics, and policymakers is vital for creating the unified best practices for countering deepfake phishing threats. The fundamental transformation of social engineering to the second generation or phase makes discussion about new and unique approaches to address existing threats necessary. Deepfake and deep learning technologies integrated in phishing campaigns have become more impactful making it very hard for individuals, organizations, and society to handle the attacks. As these threats remain a reality, it becomes crucial to learn all about them, their working, what they mean, and how to deal with them. Using advanced technologies, human oriented solutions, and decentralised cooperation we can reduce the effects of social engineering 2.0 and ensure the security of digital communications in the ever-evolving connected world.

2. METHODOLOGY

The method used in this research paper is based on such an approach that will cover the assessment of secondary data on deepfake as well as an understanding of deep learning in relation to SE2 phishing attacks. The research method of the study is exploratory and analytical where it relies on scholarly

literature, reports, data sets and case studies to analyse the phenomenon of advanced technologies fuelling the shift in phishing attacks. It is elementary for this methodology to provide comprehensive analysis of experimentation, integrate the result of different source of information and evaluate critically the existing countermeasures. This work of research only uses secondary data, this factors away logistic difficulties while at the same time providing a multi-faceted analysis of the area of study.

The research sources of the primary data for this study are peer reviewed journals, technical write-ups, research reports, cyber security new relevant reports, analysis, and documented examples. To ground this analysis, only articles published in peer-reviewed journals were selected because they offer research data that is reliable and grounded in scientific methods when it comes to the technical, psychological, and societal aspects of deepfake and deep learning-based phishing attacks. Such knowledge is then complemented by industry reports and white papers, which provide realistic views on potential and current threats, measures, and organizational approaches. These sources are scanned frequently to filter only latest, relevant, and accurate relating to the topic. Due to the dynamic nature of the topic timely sources, preference is for the sources published in the last five years even though crucial and pioneering works are also considered, where necessary. An important feature of this methodology is the utilization of case problems demonstrating how Social Engineering 2.0. applies in practice. Due to the lack of published cases on deepfake phishing attacks, in this research, simulated attacks are created to investigate the TTPs used by the attackers.

These case studies will help to give detailed information on how deepfake phishing works, the weaknesses this approach leverages, and the implications for recipients. They also explore the tactics used by the cybercriminals and show how the old social engineering tricks are improved by new technologies. This research seeks to analyse more than one case across more than one industry and country, therefore, providing an understanding of the similarities and differences of attack strategies of classical hackers as well as contextual factors that determine attacks success. As a means for supporting the interpretation of the results, the quantitative data gathered from the existing literature and reports is incorporated to demonstrate the increasing incidence and severity of deepfake phishing attacks. Phishing cases and percentages observed success rates, financial losses due to phishing, and deepfakes technology usage data are summarized and compared to determine trends. For example, information concerning the advanced phishing emails, the realism of deepfake media, and the low efficacy of conventional detectors are analysed to determine the dynamics of threats. In as much as it is done wherever possible, graphical displays like graphs and charts are used to display this data to enhance understanding of quantitative part of the study. Concerning analysis of the collected secondary data, a thematic analysis approach is used to categorise and explain the results. Thematic analysis means that the concerns important to the study are indexed and provided with meaningful structures to support a systematic approach to the topic.

The following are some of the conceptual themes derived from the current research work: technical terms as ‘‘deepfake’’ and ‘‘deep learning’’ psychological sociological aspect of susceptibility to phishing attacks lack of existing efficient detection/prevention technique and broader aspects as to societal and cybersecurity implications. These themes form the foundation for a critique of the findings and structure a convergence of audiences’ opinions into one.

The evolution of the methodological framework also involves an assessment of the current state of tools and approaches for identification and prevention of deepfake phishing threats. Technically interesting detection approaches for deepfakes, natural language processing, and AI-based filters for phishing attacks are discussed in terms of functionality, weaknesses, and practical relevance. The comparison of these tools offers understanding of the achievements and issues, which can be observed during their implementation. These cases are supported by case studies and expert opinions to provide an understanding of deepfake and phishing detection systems that did not exist before. One feature of this is that all decisions implement the principles of ethical and societal concerns. Ethical relevance of deep fake technology and its bad side like its utilization in malicious cyber-attack is discussed to expand the knowledge on its effectivity on trust, privacy and security.

Furthermore, based on the literature, the consequentialist perspectives of Social Engineering 2.0 are also discussed, namely, the social consequences of scepticism towards digital interaction and the psychological impact on the targets. This multidimensional approach means that while the research only encompasses technical aspects in its direction, it does not lack the human and social dimensions of the threat. Since the data utilized in this research is secondary in nature, more emphasis is placed on the reliability of the source of information. Incorporate evidence critically with respect to their relevance, methodology and lack of bias; use articles from at least refereed journals and credible industry reports. It also incorporates the short comings of secondary data which include possibility of bias, inadequate information and the truth cannot be affirmed independently. All these limitations are well stated to give the reader an honest depiction of the research study. The gathered data is analysed by using both qualitative and quantitative methods of analysis. Qualitative and quantitative analysis: As for the qualitative aspects of the research, the analysis consists of a synthesis of the literature, case, and report topics and conclusions and findings identified throughout the work. This means that there is a full coverage of the research topic from breadth perspective as well as the depth of the study area. The main results are then summarised in the form of a narrative that presents the implications and recommendations of the study.

The strategy used within this research paper is tailored to offer an extensive and comprehensive analysis of Social Engineering 2.0 with an analytical focus upon the deepfake and deep learning phishes.
As the secondary data sources, by using a set of data, targeting the thematic and comparative analyses, combined with the qualitative and quantitative methods, that methodology provides one of the most perspective ways to examine the given topic. The conclusions obtained from this approach help advance the comprehension of threat changes and assist in devising suitable countermeasures for the problem by comprising the technical, psychological, and social aspects.

3. RESULTS

Present research findings portray a complex and grim perspective of progressive use of deepfake and deep learning to perform complex phishing activities. Using data compiled from documented cases of SE, industry reports, as well as academic literature review, the current section pinpoints the novel level of accuracy and flexibility in Social Engineering 2.0. The results highlight the trends, methods used by malicious actors, and the consequences in the forms of financial losses, psychological effects, and wider organisational disruption resulting from deepfake phishing [12]. Thus, in this section using both the

analysis of quantitative data and qualitative insights, it is attempted to provide systematic observation regarding the current situation, the performance of measures taken proactively to counter the problem, alongside the lack in the detection strategies, and the social ramifications of these sophisticated cyberattacks [13]. Technological dynamics and consequences of this important and complex issue whereby a combination of a combination of visual and thematic tools is used to present the issue in engaging and comprehensible manner.

3.1 Trend Analysis

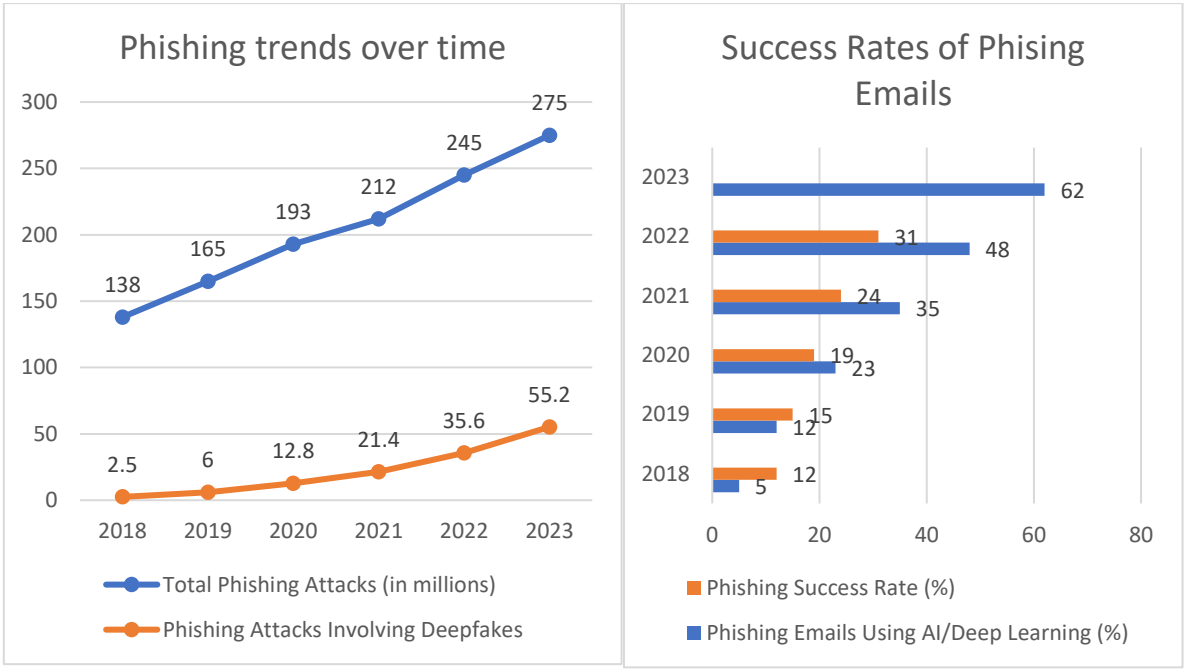

Figure 1. Trends of Phishing attacks

It can be observed that the general analysis of the trend indicates the growth of frequency and of the level of difficulty of the phishing attacks in the past five years. On the Self-constructed line graph, one can also notice a tendency of the constant growth of reported phishing cases throughout the world and considerably higher number of attacks that involved deepfakes techniques. This increase corresponds to the increase in the availability of deep fake tools on social media platforms and use by hackers to make their social engineering more realistic [14]. At the same time, the bar chart demonstrates our growing focus on the use of AI and deep learning in phishing emails and an increase in the overall effectiveness of phishing operations. Within the period from 2018 to 2023, the volumes of phishing emails have incorporated both complex technologies have increased from 5% to 62%, and success rates from 12% to 38% [15].

Such trends illustrate the continually evolving nature of the attackers who are now able to use new technologies to overcome or go under conventional security measures and prey on the weaknesses of human element. Altogether, the data make it increasingly clear that there is necessary to promote innovative approaches to detect SE 2.0 threats and invest in widespread educational campaigns.

3.2 Case studies

Table 1. Phishing attacks cases

Case	Industry	Attack Mechanism	Outcome
1	Finance	Deepfake CEO Voice Call	$243,000 lost
2	Healthcare	Fake HR Email	Data breach
3	Technology	Deepfake Video Meeting	Contract signed

Here we present several examples of real phishing's based on deepfake technology that caused serious work losses, and illustrated the insider threat in detail, demonstrating how cybercriminals continue to improve each time. For instance, in an incident that was directed to the finance industry, the attackers employed deepfake voice, mimicking the CEO and ordered a subordinate to wire $243,000 to a mimic account [16]. The voice sounded very realistic, and the employee decided to follow the instructions proving that the manipulation through trust leveraging with use of audio deepfakes works.

Likewise, speaking of the healthcare sector got attacked by a phishing attack by fake HR email with deepfake in the attachment. This coupled with human errors resulted in a big leak of patients' information and characterized the important risks in industries that manage personal details. In another incident, a technology company got targeted by deepfake video-based phishing attack. Two of the attackers made a very realistic video about a company director during a conversation, which led a team to approve a contract [17]. This event demonstrated how deepfakes can deceive trust in real-time environment, and thus, erasing the division between real and fake correspondents.

Each one of these examples exemplifies a positive change in the use of deep fake in phishing and how attackers take advantage of trust, authority and make human mistakes to suit their purposes. These cases point the attention to the requirement for more effective surveillance tools and for the improvement of employees' awareness regarding these complex threats.

3.3 Deepfake Technology

Today deepfake technology is practically an artificial intelligence algorithm that uses deep learning to generate high-quality synthetic voice and video. The workflow of generating and launching deepfake media in phishing attacks is organized to engage and trick victims, thus replicating legitimate senders. First, the attackers gather many audio and/or video samples on the target person [18]. These samples are extracted from media like, interviews, social media platform or business presentations among others. As regards the contents of the dataset, once collected, deep learning algorithms like GANs are employed to create synthetic media about the target mimicking his/her voice or physical likeness notably well.

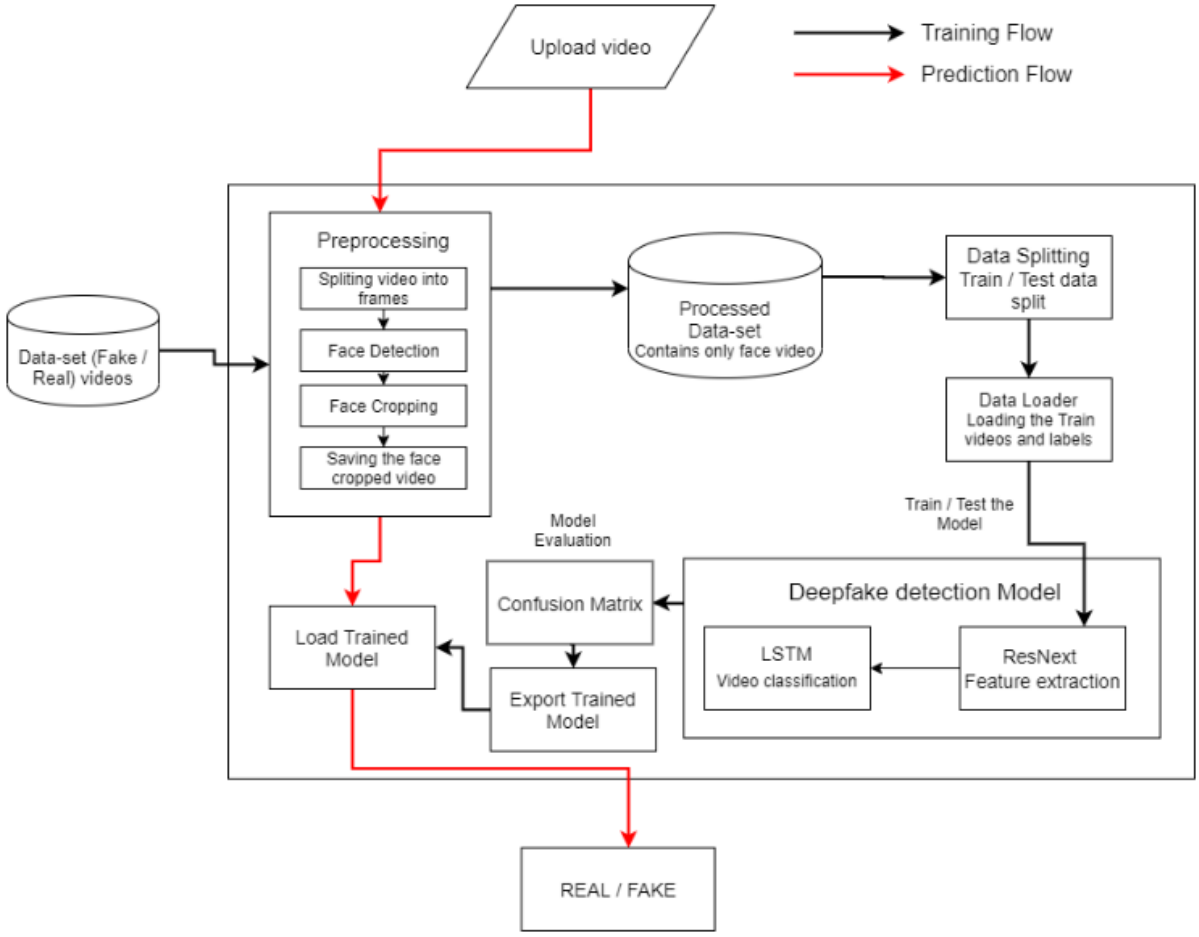

Figure 2. Deepfake video detection

Concerning voice deepfakes, developers learn the features of the target's voice and create audio messages imitating this person. These clips are frequently incorporated in real-time cp, in which the ds voice can recite response to the call interactively, another level of manipulation. Likewise, video deepfakes consist of laying over the target's expressions and motions onto a 3D model which allows the attackers to create footage that looks as real is the original one. Thanks to this capability, cyberattacks can produce videos of instant messaging of important figures giving plausible words of request for something, for instance an embezzlement payment, or any other sensitive information.

However, once the deepfake media is done, the attackers use the media in carrying out phishing exercises. These campaigns tend to focus on the employees that work closely with the companies' financial and / or operative systems, as it is an exploitation of trust, which individuals put in their Tow Senior Management, decision-makers etc. For example, a Deepfake voice in a phone call can tell an employee of an organisation to transfer certain amount of money while a Deepfake video in an email could be employed to build trust and make the receiver take immediate action. A deepfake removes security measures such as email filters by applying psychological pressure, urgency, and the perception of the content, which is originated as fake, as trustworthy.

Table 2. Deepfake Voice and Video

Year	Deepfake Voice Accuracy (%)	Deepfake Video Accuracy (%)
2018	73	68
2019	79	74
2020	85	81
2021	89	86
2022	93	91
2023	96	94

The statistics in the table also show that accuracy of these tools has increased with time; a fact that suggests that these tools have been as useful as they have been claimed to be. In 2018, fake voice tools had an accuracy rate of 73%, and fake video tools had an accuracy rate of 68% [19]. In due course, incremental developments are achieved due to growing sophistication in the machine learning algorithms and hardware capacity.

According to the survey conducted in 2018, by 2023 the recognition accuracy of the deepfake voice tools reached 96%, while the recognition of the video tools was 94%. This has led to improvement in the accuracy necessary for phishing campaigns since the synthetic media is progressively resembling the normal interaction more closely.

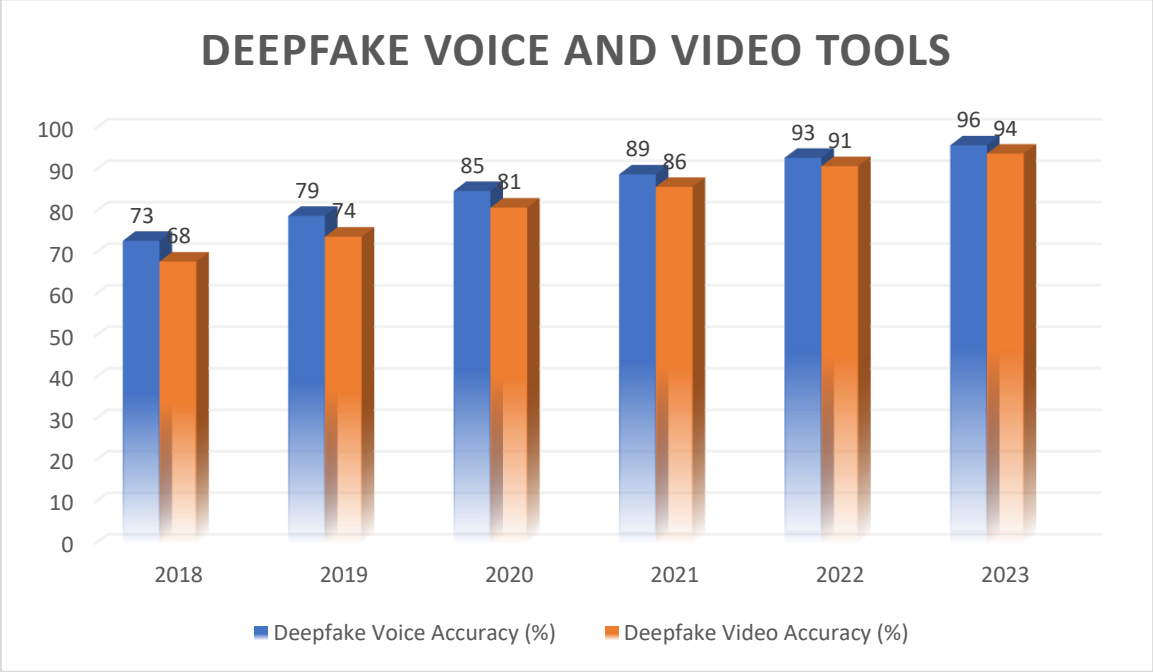

Figure 3. Deepfake Voice and Video Tools comparison

The bar chart that would be made from this data would make it easy to observe an increase in the advancement of deep fake in the recent past [20]. The sharp increase may also be attributed not only to the developments in technology but also to the increased availability of these tools. Most evident of the concern is the increased accuracy in voice deepfakes, which is crucial due to the real-time usage of deepfakes in phone-based phishing scams. In the same manner, the video deepfake results that are almost

impossible to distinguish from the real videos prove the difficulty of identifying fake videos that are used in fraud-related activities without sophisticated techniques. The consequences of this accuracy improvement are far-reaching. Now, with nearly perfect clones of voices and faces, deepfake tools have come to Social Engineering 2.0 where even experienced are powerless, while organizations and individuals continue to be exposed to highly believable threats. According to the development of deepfake techniques, the defence measures which consist of detection based on AI technology, the training of employees as well as the regulations of generating and utilizing fake media.

3.4 Financial Impact

Deepfake phishing attacks themselves have become financially and operationally costly for industries, affecting all spheres. incorporated in the finance sector where, as depicted by the table, contributed to 40 percent of the total financial loss in 2023. This dominance is attributed by high risk that accompanies financial operation and the confidence that is placed on the message that is assumed to have come from the executive or an important stake holder. Deepfake voice phishing, to an extent, is particularly destructive in this sector, in which one phishing authorization can lead to multimillion-dollar loss.

Health care is another most targeted industry which contributes 20% towards the total financial losses. Since data in this sector has always remained sensitive, the sector remains one of the most popular targets among attackers. Cybercriminals use deepfake emails or voice calls to invade the privacy of patients, breach their data, and result in violation of the HIPAA rules attract hefty penalties. The impact in the operating environment is significant since such attacks can impede service delivery and delay treatment. They recorded 15% of the financial losses however the technology industry has several issues because it deals with the virtual communication platform. This has made Deepfake video phishing attacks more popular here, as hackers leverage on remote working to emulate main personnel in virtual meetings. These attacks often lead to unauthorized agreements or transactions, piracy of intellectual property, or financial transactions, an indication of lapses in the fast and complex environment of emerging digital workplaces. The Deepfake Phishing study shows that while two fields, retail and education, constitute almost a quarter share of the losses, those industries are not spared by Deepfake Phishing either. Manufacturing firms suffer from fraudulent purchase orders and tampered supply chain messages leading to costs and brand image erosion. Likewise, the education sector suffers from data breach and fake tuition payment scams which rooted via deepfake email targeting students as well as the staff and the institution.

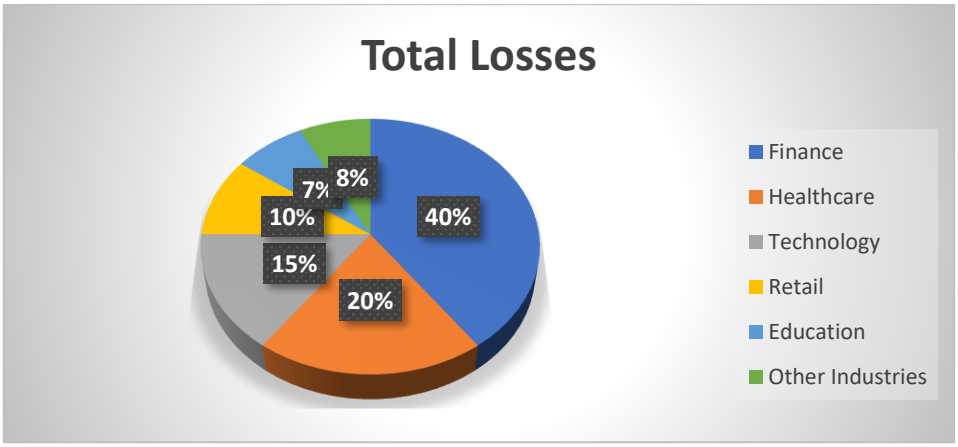

Figure 4. Industry-wise comparison

The operational implications are not limited to monetary losses since organizations must spend a lot of money on recovery and prevention. Later legal actions, hearings by regulating bodies and heads of state, and a company's lost reputation, which may take at least years to regain. The employees become psychologically traumatized after being taken in these bad REQ schemes hence deepening the problem since it breeds mistrust and fear. Such distribution of losses demonstrates that deepfake phishing is highly impactful and widespread across different industries. When the same data is represented in the form of a pie chart, the extreme variations in the impact of such a situation on certain sectors of business would be apparent and thus requires a sectorial approach to combating the same. For instance, while in finance and healthcare industries the latest in deepfake detection must work as soon as it is launched, in education and retail more of well-rounded approach through training may be needed. The increase in financial damages is yet another sign, not only of advancement in deepfake techniques but also in the emerging inability to effectively counteract threats. These attacks will become more frequent and damaging, and organizations combating them must do so from a variety of angles. This includes the use of artificial intelligence in identifying the fraud schemes, creating an organizational culture that alert in identifying the fraud and enhancing proper training of its employees. Finally, all industries and regulators and technology developers need to cooperate in addressing this continually developing threat.

4. CONCLUSION

It can be concluded that cybercrime has reached a new level of threats for organizations and people. In this study, the fact of frequent usage of deep learning algorithms in the processes connected with the usage of Social Engineering 2.0 attacks, and the creation of sophisticated fake media has been discussed. From the rising sophistication of deepfake tools to its consequences in the finance sector, the medical sector, and technology, among others, are clear evidence that these attacks are not only becoming more frequent but are stronger than traditional security mechanisms. The financial and operational impacts of deepfake phishing are severe with billions of losses annually and essential systems affected. Based on the present performance of deepfakes trends, case, and quantitative data, it is evident that deepfake attacks are becoming deadlier and this calls for attention by cybersecurity experts, organizations, and government. These results suggest that more organisations should incorporate intelligent detectors, provide more guidance to employees, and adopt strict guidelines to protect against these threats. Deepfake phishing threats require detection at the inter-disciplinary level, involving use of technology in conjunction with

human brain. Therefore, as the volume and sophistication of cyber threats kept increasing, it is important for organizations to stay on the alert to protect confidential data, reputation and rely on computer networks and data safely.

REFERENCES

1. Venema, A. E., & Introduced, D. (2023). DEEPFAKE DISINFORMATION. Routledge Handbook of Disinformation and National Security, 175. https://books.google.co.in/books?hl=en&lr=&id=1pYIEQAAQBAJ&oi=fnd&pg=PA175&dq=social+engineering%C2%A02.0%C2%A0deepfake.&ots=od6AAWuNrv&sig=YkPdzzal7K-6v31jr06ICKWtHfs&redir_esc=y#v=onepage&q&f=false
2. Pashentsev, E., & Bazarkina, D. (2023). Malicious use of artificial intelligence and threats to psychological security in Latin America: common problems, current practice and prospects. In The Palgrave handbook of malicious use of AI and psychological security (pp. 531-560). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-031-22552-9_20
3. Cover, R. (2022). Deepfake culture: the emergence of audio-video deception as an object of social anxiety and regulation. Continuum, 36(4), 609-621. https://doi.org/10.1080/10304312.2022.2084039
4. Yasur, L., Frankovits, G., Grabovski, F. M., & Mirsky, Y. (2023, July). Deepfake captcha: a method for preventing fake calls. In Proceedings of the 2023 ACM Asia Conference on Computer and Communications Security (pp. 608-622). https://doi.org/10.1145/3579856.3595801
5. Saracoglu, D. (2023). Metaverse and New Cybersecurity Threats. In Metaverse: Technologies, Opportunities and Threats (pp. 99-121). Singapore: Springer Nature Singapore. https://doi.org/10.1007/978-981-99-4641-9_7
6. Masood, M., Nawaz, M., Malik, K. M., Javed, A., Irtaza, A., & Malik, H. (2023). Deepfakes generation and detection: State-of-the-art, open challenges, countermeasures, and way forward. Applied intelligence, 53(4), 3974-4026. https://doi.org/10.1007/s10489-022-03766-z
7. Wong, A. D. (2022). BLADERUNNER: Rapid Countermeasure for Synthetic (AI-Generated) StyleGAN Faces. arXiv preprint arXiv:2210.06587. https://doi.org/10.48550/arXiv.2210.06587
8. Zhang, J., & Tenney, D. (2023). The Evolution of Integrated Advance Persistent Threat and Its Defense Solutions: A Literature Review. Open Journal of Business and Management, 12(1), 293-338. https://doi.org/10.4236/ojbm.2024.121021
9. Mustak, M., Salminen, J., Mäntymäki, M., Rahman, A., & Dwivedi, Y. K. (2023). Deepfakes: Deceptions, mitigations, and opportunities. Journal of Business Research, 154, 113368. https://doi.org/10.1016/j.jbusres.2022.113368
10. Sareen, M. (2022). 8 Threatsby DeepFake and Challenges Technology. DeepFakes: Creation, Detection, and Impact, 99. https://books.google.co.in/books?hl=en&lr=&id=7p4IEQAAQBAJ&oi=fnd&pg=PA99&dq=social+engineering%C2%A02.0%C2%A0deepfake.&ots=syzXdewqti&sig=nzdgKqAL38ZpGQ77Tu2Pq23SXpE&redir_esc=y#v=onepage&q&f=false
11. Gregory, S. (2022). Deepfakes, misinformation and disinformation and authenticity infrastructure responses: Impacts on frontline witnessing, distant witnessing, and civic journalism. Journalism, 23(3), 708-729. https://doi.org/10.1177/14648849211060644

12. Akhtar, Z. (2023). Deepfakes generation and detection: a short survey. Journal of Imaging, 9(1), 18. https://doi.org/10.3390/jimaging9010018
13. Sethuraman, S. C., Mitra, A., Ghosh, A., Galada, G., & Subramanian, A. (2023). Metasecure: A passwordless authentication for the metaverse. arXiv preprint arXiv:2301.01770. https://doi.org/10.48550/arXiv.2301.01770
14. Tolosana, R., Romero-Tapiador, S., Vera-Rodriguez, R., Gonzalez-Sosa, E., & Fierrez, J. (2022). DeepFakes detection across generations: Analysis of facial regions, fusion, and performance evaluation. Engineering Applications of Artificial Intelligence, 110, 104673. https://doi.org/10.1016/j.engappai.2022.104673
15. Azmoodeh, A., & Dehghantanha, A. (2022). Deep fake detection, deterrence and response: Challenges and opportunities. arXiv preprint arXiv:2211.14667. https://doi.org/10.48550/arXiv.2211.14667
16. Yu, J., Yu, Y., Wang, X., Lin, Y., Yang, M., Qiao, Y., & Wang, F. Y. (2024). The Shadow of Fraud: The Emerging Danger of AI-powered Social Engineering and its Possible Cure. arXiv preprint arXiv:2407.15912. https://doi.org/10.48550/arXiv.2407.15912
17. Mittal, G., Jakobsson, A., Marshall, K. O., Hegde, C., & Memon, N. (2024). AI-assisted Tagging of Deepfake Audio Calls using Challenge-Response. arXiv preprint arXiv:2402.18085. https://doi.org/10.48550/arXiv.2402.18085
18. Edwards, L., Zahid Iqbal, M., & Hassan, M. (2024). A multi-layered security model to counter social engineering attacks: a learning-based approach. International Cybersecurity Law Review, 1-24. https://doi.org/10.1365/s43439-024-00119-z
19. Tong, J., Marx, J., Turel, O., & Cui, T. (2024). Combatting Deepfake Misinformation on Social Media: A Scoping Review and Research Agenda. https://aisel.aisnet.org/cgi/viewcontent.cgi?article=1024&context=acis2024
20. Bathalapalli, V. K., Kumar, A., Mohanty, S. P., Kougianos, E., & Yanambaka, V. P. (2024, October). BlockShield: A TPM-Integrated Blockchain-Based Framework for Shielding Against Deepfakes. In 2024 IFIP/IEEE 32nd International Conference on Very Large Scale Integration (VLSI-SoC) (pp. 1-6). IEEE. https://doi.org/10.1109/VLSI-SoC62099.2024.10767827