Received: November 4, 2022. Revised: December 13, 2022.

International Journal of
Intelligent Engineering & Systems

http://www.inass.org/

The Security System in Information Technology Detects Fake Accounts on The Youtube Platform

Alaa Thamer Mahmood[1] **Raed Kamil Naser**[2]* **Sura Khalil Abd**[3]

[1]*Technical Instructors Training Institute, Middle Technical University, Baghdad, 00964, Iraq*
[2]*Military Training Directorate, Ministry of Defense, Baghdad, 00964, Iraq*
[3]*Department of Computer Techniques Engineering, Dijlah University College, Baghdad, 00964, Iraq*
* Corresponding author's Email: ayhar_2013@yahoo.com

Abstract: YouTube platform provides paid/ unpaid video services for different users for promotions, education, learning, information sharing, etc. Accessing this platform's content/ channel-based advancements is based on user identification through registration/ credential sharing. In recent years, accessing information through fake accounts has increased other users' illegitimacy and security demands. The article introduces a temporal behavioral method (TBM) using the classified learning (CL) technique to address the impact of fake users on YouTube-like platforms. The proposed method eyes on user credentials, interactions, and access information under different usage series. The data used for the classification learning identifies the active, inactive, and false searches/ information accessed by the user. With the recent meetings removed, the static and inaccurate searches are segregated using the learning process from the last known access session. Rather than applying new rules to the identified account/user, false user flagging reported data sharing and illegitimacy is all avoided. Based on the different session logins and the recommendations by the learning process, the fraudulent users are restricted from communicating with other users and legitimate information. Maximizing data sharing and user detection is facilitated by the better recognizability afforded by the prolonged inactive session categorization following in succession. The suggested TBM-CL strategy addresses the impact of fake users on YouTube-like platforms and increases the maximums for user detection (14.93%), information sharing (11.3%), and classifications (15.93%) across all sessions. It decreases the erroneous ratio by 8.03% and the delay by 9.07% based on the dataset.

Keywords: Behavioral model, Classification learning, Fake account detection, YouTube.

1. Introduction

Fake account detection is a critical and complicated task to perform on YouTube. Fake account causes various user problems that reduce the quality of service (QoS) rate on YouTube. The detection process needs an appropriate data set to detect the fake account among the other charges [1]. Various methods and techniques are used in the YouTube artificial account detection process that improves the security system's performance and feasibility. Counterfeit account detection in online social networks (OSN) needs several techniques and methods [2]. A genetic algorithm is mainly used in the detection process, reducing the computation cost and time in both the classification and identification processes. The genetic algorithm improves the accuracy rate of fake account detection, providing various security policies to other users [3]. Fake accounts are identified based on particular features and patterns publicly available to all users. IP address identification process plays a significant role in the detection process that enhances the robustness of the management system. The victim prediction method is used in the fake account detection process that predicts the patterns of fake accounts by comparing them with original versions. A robust detection mechanism is used in the prediction process that indicates the features and designs of accounts presented in a dataset [4, 5].

The user behavior model is a process that

identifies the customer's behaviors under similar circumstances. The user behavior model observes the common behaviors of users and provides an appropriate set of data for the different analysis processes [6]. YouTube uses various user behavior models to find out the behaviors and interests of users that provide necessary information regarding services. On YouTube, customers search for various topics and things that are related to claims and preferences [7]. The user behavior model finds out the likelihood and preferences of users and provides information that helps to improve the overall quality of service (QoS). Key benefits and qualitative data are first identified by the user behavior model that reduces the latency rate in providing user services [8]. The identifier is widely used in the user behavior model to determine the exact meaning of a user's search and behaviors. Identifier improves the accuracy rate in the identification process, which increases the reliability and effectiveness of YouTube [9]. The user behavior model first discovers the habitual patterns of users on YouTube and produces an optimal set of data for the detection and prediction process. The user behavior model provides necessary information for a YouTube session that helps recommend related topics to the users [10].

The machine learning (ML) approach is primarily used in various fields for detection, recognition, and analysis. ML approach increases the accuracy rate in the detection process, improving the system's performance and feasibility [11]. ML is commonly used in the YouTube fake account detection process that increases the security level of other users. The advanced machine learning (ML) approach is widely used for the artificial YouTube account detection process. The progressive ML approach uses classification and identification processes to find the essential features presented in an account [12]. The classification process plays a vital role in improving the accuracy rate of fake account detection. The k-nearest neighbor (KNN) algorithm uses various classification techniques in the fake account detection process [13]. The classification technique identifies the critical content and details about an account and finds actual details about the user. The classification process identifies details such as profile picture, name, IP address, and other account-related information. The convolutional neural network (CNN) algorithm is also used in fake account detection to improve the accuracy rate in the identification process. CNN detects the details and finds out the meaning of the account's details that reducing unwanted threats to the users [14, 15].

The paper presented a temporal behavioural method (TBM) based on the classified learning (CL) approach for mitigating the effect of false subscribers on YouTube-like sites. The proposed methodology tracks user IDs, activity, and access files over several sessions. Categorization learning uses this data to determine if a user is performing an actual search or just browsing for information. Knowledge from the last recorded login session distinguishes between active sessions and dormant or incorrect searches. CL is necessary because different restrictions must be placed on the recognized login to avoid incorrect user tagging, suspected data exchange, and unconstitutionality.

The rest of the article is organized as follows: section 2 discusses the related works, section 3 proposes the TBM-CL approach, section 4 deliberates the experimental outcomes, and section 5 concludes the research paper.

2. Related works

H. Oh. [16] introduced a new YouTube spam detection method using an ensemble machine learning model. Support vector machine (SVM) and decision tree algorithm are used here to perform the classification process. The classification process finds the spam on YouTube and provides appropriate information for the detection process. The proposed method increases the accuracy rate in the spam detection process, enhancing YouTube's security and performance. Cognitive and psychological indicators, like user satisfaction, are limited in their mathematical equations.

K. Yousaf and T. Nawaz. [17] proposed a deep learning-based approach for YouTube videos' classification and detection process. The bidirectional long short-term memory (BiLSTM) model is used in the deep learning approach that provides a proper classification process. Classifier plays a vital role in identifying videos that provide necessary information for detection. A convolutional neural network (CNN) is used here for the detection process that detects the abnormal videos that are presented on YouTube. Compared with other methods, the proposed method increases the accuracy rate in the classification process. Considering the video classification issue, one potential drawback of LSTM was that it captures the past context only.

J. V. de Souza, J. Gomes, F. M. de Souza Filho, A. M. de Oliveira Julio, and J. F. de Souza. [18] introduced a systematic mapping technique for social media's fake news classification process. Fake news, such as rumors and irrational information, is

spread on social media. Fake news identification is a complicated task to perform on every social media. The systematic mapping technique uses a particular classifier to classify rumors and provide necessary details to social media. The proposed method reduces the fake news spreading rate that improves the feasibility of social media.

H. Jelodar, Y. Wang, M. Rabbani, S.B. Ahmadi, L. Boukela, R. Zhao, and R.S. Larik. [19] proposed a Natural language programming (NLP) based sentiment analysis method for YouTube. The proposed method uses a fuzzy logic algorithm to find the exact meaning of the users' comments. A classification algorithm is also used here to classify users' emotions and provide appropriate information for further analysis. NLP identifies the exact feeling of users that enhance the performance rate in recommending videos.

B. Jang, S. Jeong, and C. K. Kim. [20] introduced a distance-based customer detection method for YouTube. The proposed method is used to find out the fake YouTube followers that play a significant role in providing services for the actual users. Node geographical locations are first identified and are used in the detection process. Experimental results show that the proposed method increases the overall accuracy rate in the detection process, which improves the feasibility and robustness of YouTube.

A. Mulahuwaish, K. Gyorick, K. Z. Ghafoor, H. S. Maghdid, and D. B. Rawat. [21] proposed an efficient classification model for the web by using a machine learning (ML) algorithm. K-nearest neighbor (KNN) and decision tree (DT) algorithms are used here to find the exact news on the web. Classifiers are implemented here to classify the critical information set and provide the necessary details for the detection process. The proposed method improves the accuracy rate in the classification process, reducing time consumption and energy consumption rate in the computation process. However, more research in web mining is needed considering space and time complexity.

R. M. Ortiz-Gaona, M. Postigo-Boix, and J. L. Melús-Moreno. [22] introduced a mathematical model that predicts the influence propagation in online social networks (OSN). Identifying the exact node is a challenging task in the classification process that contains the actual content regarding the news. The proposed method is mainly used to determine if the virus spreads in a computer network that causes severe damage to the users. Users' behaviors and interests are also predicted by the model that provides the necessary information to improve the performance rate in providing services for the users.

T. Chauhan and H. Palivela. [23] proposed a deep learning-based approach for the fake news detection process in social media. A decision tree (DT) algorithm is used here to find out the nodes containing fake news. Long short-term memory (LSTM) neural network is also used here to classify the actual content of fake news and provide appropriate information for the detection process. LSTM increases the detection process's accuracy rate, reducing the phony news spreading rate on social media. However, the suggested algorithm needs a numerical feature vector of fixed sizes rather than a raw text document of variable length. The raw data, a series of symbols, cannot be given directly to them.

Zappin, H. Malik, E. M. Shakshuki, and D. A. Dampier. [24] introduced a machine learning-based methodology to identify monetization and censorship on YouTube. The proposed method analyses the videos uploaded on YouTube and finds the videos' actual meaning. A random forest (RF) algorithm is used here for the censorship process that provides the necessary function for YouTube. The accuracy rate is increased by using RF, which enhances the performance and feasibility of the system. The paper does not differentiate the frequency of advertisements in a video.

E. Van Der Walt and J. Eloff. [25] proposed fake identity detection using machine learning (FID-ML) approach for the process in social media platforms (SMP). Many users create counterfeit identities to earn more money from YouTube. The proposed method finds a profile's follower count and friend count and then identifies the user's IP address. The ML approach in SMP also identifies certain features and patterns. The proposed method reduces the fake identities rate in SNP, improving the user's security level. The accuracy measure, however, does not account for incorrect predictions and undergoes skewed distribution.

M. Egele, G. Stringhini, C. Kruegel, and G. Vigna. [26] introduced a new method for compromised accounts detection in social networks (CAD-SN). The proposed method detects the fake accounts available in OSN, reducing the attack rate on significant companies and channels. Large-scale compromises are first identified using specific classification and detection processes. The user's behavior dataset provides the necessary information for identifying high-profile accounts. The compromised account detection process improves the accuracy rate in the detection process, which reduces the latency rate in searching.

J. Kang, H. Choi, and H. Lee. [27] proposed a deep recurrent convolutional network user interest prediction in social media. The word-embedding technique is used here to find the essential vectors in social media. Vectors and nodes are used here to form a sentencing matrix. Sentence matric provides an appropriate set of user behaviors that plays a vital role in the prediction process. Experimental results show that the proposed method predicts users' actual interest and preference in social media. The research offered can be comprehensive in several potential ways. In addition to the textual information of social media, social network data like friend relationships can be utilized to understand users' interests more precisely.

B. Bebensee, N. Nazarov, and B. T. Zhang. [28] introduced an ego-graph topology method for fake account detection in social media networks. The proposed method finds spam, fake accounts, and malware distributors in social media. Ego-graph identifies fake account users in social media that improves the security level of other eligible users. The proposed method increases the detection process's accuracy rate, which enhances social media's performance and feasibility rate.

X. Jiang, Q. Li, Z. Ma, M. Dong, J. Wu, and D. Guo. [29] proposed a single-machine graph computing framework for the fake account detection process in social networks called QuickSquad. Graphs are divided into light and heavy datasets that provide appropriate information for the detection process. Important notes and vectors are identified using a random forest algorithm. The proposed method improves the accuracy rate in the fake account detection process, reducing social media problems and threats. Experimental results show that the proposed method increases the performance and effectiveness of social media.

Social information, multimedia, etc., sharing systems like YouTube are impacted by phone identities and bogus users that violate confidentiality and legitimize the platform. Threats must be identified and countered to keep these characteristics intact. The detecting procedure for YouTube fraudulent accounts uses some tools and strategies to boost the effectiveness and usability of network security. Multiple components and strategies are required for detecting fake accounts in OSNs.

Using an ensemble of machine learning models, Oh [16] established a new approach to detecting spam on YouTube, improving the site's safety and performance. Yousaf [17] advocated using deep learning to automate YouTube video recognition and categorization. In this approach, a convolutional neural network is used to detect out-of-the-ordinary videos on YouTube. The suggested strategy improves classification accuracy when compared to existing processes. Bebensee [28] proposed a method for detecting false accounts in social media platforms using ego-graph topology, which has the added benefit of increasing the efficiency and viability of these platforms. Jiang [29] suggested a single-machine graph computing architecture to identify bogus accounts in social networks. The recommended solution enhances the precision of the false account identification procedure, hence decreasing issues and risks in social networks. The proposed method improves the efficiency and efficacy of social media, as evidenced by experiments. The proposed TBM model outperforms the state-of-the-art in terms of identifying users and providing data, and it does so by classifying inactive sessions in sequence. The suggested strategy increases the maximums for user recognition (14.93%), information sharing (11.3%), and categorization (15.93%) across all sessions. It decreases the erroneous ratio by 8.03% and the delay by 9.07%

3. Proposed method

The TBM method is designed to improve the security systems in information technology to detect fake accounts/user identification on the YouTube platform. Based on the user behavior on a social platform is to see false user flagging, reported data sharing, and illegitimacy through different restrictions on the identified user/account based on user identification through registration/credential sharing input user information. The information is required from the users, their habits, their friends and family, etc.; false user detections are observed at any interval. This method aims to reduce the false searchers on the YouTube platform wherein independent user identification. The challenging task in this proposed work is user credentials, interactions, and access information under various usage series instances detected based on the previous session logins and recommendations. The security demands and illegitimacy instances are stored as records for the user identification/verification instances. The aid of users' personal information sharing on social platforms increases new privacy concerns and obtains insights into security system issues. In general, the security problems in YouTube platform applications are doubled with the identifiability and platforms content based on the information available in this social network platform, its possible user registration or credentials sharing, and its potential

Received: November 4, 2022. Revised: December 13, 2022.

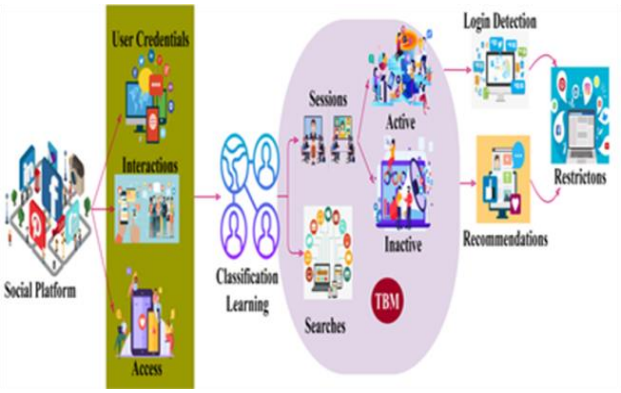

Figure. 1 Proposed TBM-CL process

uses. Malware is not only the threat. Based on the unlimited information access to the personal profiles of users/accounts, false searchers or users cab further gain the user's information relies on commercial secrets and corporation details. The proposed method's process is illustrated in Fig. 1.

In This proposed method, fake account detection control is designed to address the fake profile and users on social platforms through YouTube access/content. In user behavior monitoring based on counterfeit account detection on the YouTube platform, gathering a lack of information, such as interactions, access information, user credential, etc., is monitored that is maintained through classification learning. Based on user behavior on the YouTube platform, the classification learning is performed based on three factors: active sessions, inactive sessions, and false searchers are segregated to analyze and process using the learning process from the last known user access session. The classification learning identifies the active and inactive sessions and then false searchers/information accessed by the user. The gathered information such as user credentials, interactions on social networks, information access, etc. This information is utilized for false user flagging and illegitimacy through accessing information based on the social platform. In the proposed TBM method, reliable, legitimate information and inactive session control are designed using the learning process's different session logins and recommendations. However, to retain the security demands of the other users, the proposed method provides reliable fake account detection through accessing information. The function of TBM is to identify fake users on YouTube-like platforms based on content and identifiability. The different instructions on the identified fake accounts/users and restricted from the community are performed and are credentials shared between platforms to other applications. The security system in information technology detects fake accounts/users on YouTube platforms. The contents of channel-based promotions are based on user identification and are administered to prevent a phony account from accessing users on the YouTube platform.

Problem Definition: Let $\{1,2,\ldots,y_u\}\in Y_U$ Represent the set of YouTube platform users accessing information at different time intervals $Y_P\in(C_A\text{-}C_P)$ where C_A and C_P is addressing the fake profiles in accessing the content and channel-based promotions through user identification, respectively. Based on the different intervals of user connectivity c_v, the fake account detection F_d is to be less under falsely controlled users based on the identification process that is given by Eqs. (1) and (2)

$$\forall (C_A\text{-}C_P)= \quad argmin \sum_{t=1}^{Y_P}(c_v\text{-}F_{Ac})_t,\ Y_P\in[C_A+1,\ C_P] \tag{1}$$

Such that,

$$\forall (C_A\text{-}C_P)= \quad argmin \sum_{t=1}^{Y_P} Y_P,\ \forall\{1,2,\ldots,y_u\}\in Y_U,\ C_A\leq Y_P\leq C_P \tag{2}$$

From the above Eqs. (1) and (2), the variable F_{Ac} denotes the false accounts and $(c_v\text{-}F_{Ac})$ in YouTube accessing any social platform $\in[C_A+1,\ C_P]$. The increase in the illegitimacy and security demands for the other users at different time intervals reduces the false user flagging. The above fake account detection problems are addressed using the identifiability and connectivity of the YouTube applications based on user behavior through classification learning. However, there are some conditions based on YouTube content accessing is computed as

(i) $\forall y_u\in Y_U$ in$[C_A,\ C_P]$, if the condition (y_u+1) is true in $Y_P\in[C_A+1,\ C_P]$, then (y_u+1) YouTube platform provides paid or un-paid video services for different users based on promotions, education, learning, and information sharing for (Y_P+C_A+1) where Y_P denotes the fake account detection instances.

(ii) $\forall y_u\in Y_U$ In YouTube content access, if $Y_P=C_P$ or $Y_P>C_P$, then $F_s=(c_v\text{-}F_{Ac})$ increases and $\frac{F_{Ac}}{Id_f}\rightarrow 0$.

Condition (ii), the variable F_s denotes the false user's identification. The first condition represents accessing information through false accounts based on the condition $(C_A\text{-}C_P)$ where the second

 DOI: 10.22266/ijies2023.0228.45

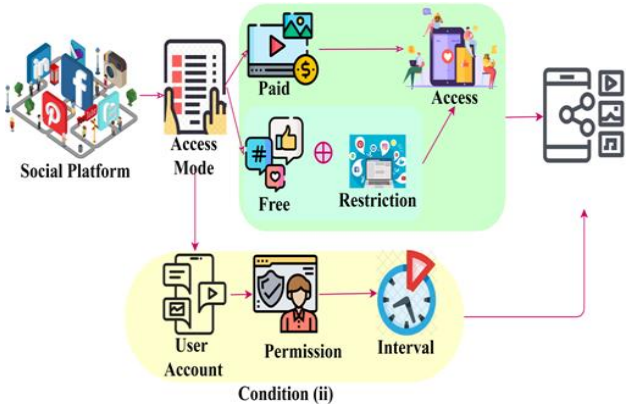

Figure. 2 Initial condition-based classification

YouTube user y_u (i.e.) (y_u+1) co-joint with $Y_P \in [C_A+1, C_P]$ depends on y_u. The augments the fake accounts for both y_u and (y_u+1) conditions. Similarly, the second condition identifies the fake accounts and users on the YouTube platform at different intervals and also verifies if it is the same with successive y_u provided $argmin \sum_{t=1}^{Y_P} C_A \; \forall [C_A\text{-}C_P]$ is satisfied. Fig. 2 presents the initial condition-based classification.

Fig. 2 presents the conditional classification $\forall Y_u \in [C_A, C_p]$ and $y_u \in Y_U$ For access and content sharing. The permissions are delegated to the accounts depending on the user access levels. This delegation relies on permissions, active sessions, and the searches presented by the user. The TBM using a classified learning process is administered between YouTube users and other social applications connected to information technology. The registration and credential sharing of the user identification is the real-time application services for securing the information and details of the social media users, community, and corporation. The Identifiability (Id_f) of fake accounts depends on the user behavior on the platform. However, the phony account less YouTube application access is the considered factor for Id_f. Hence, from on the conditions (i) and (ii), the identifiability is computed as

$$Id_f = \frac{(C_A+1\text{-}C_P)}{(C_A\text{-}C_P)} \quad \forall C_P \leq Y_P \leq C_A \tag{3}$$

The identifiability factor must monitor with the false user and fake account based on the conditions such that $Y_P \in [C_A+1, C_P] \; \forall (y_u+1)$ does not provide security demands to the $Y_P \in [C_A, C_P] \; \forall y_u \in c_v$. The above condition for a fake account and false user identifiability does not access the information as it differs due to counterfeit account detection in the YouTube platform. The user identification through registration or credential sharing of the social media applications is reflected over connectivity that defines $(c_v\text{-}F_{Ac})$ and satisfies $\frac{F_{Ac}}{Id_f} \rightarrow 1$. The two information-accessing technique of the YouTube application is increasing the illegitimacy, and security demands rely on Id_f and $[C_A, C_P]$ the fake account detection is computed as F_s based on the previous user identification on YouTube. Condition (i) and (ii) are serially processed for both the identifiability and connectivity of the social application and user access using session logins and recommendations. Classification learning helps identify the active and inactive sessions and false searchers for reliable user registration and credential sharing. The following section discusses the YouTube platform-based condition suppression using classification learning.

Condition (i) Analysis: This condition deals with the user behavior on YouTube applications and detects fake accounts based on successive session logins. The first output for connectivity and identifiability is splitting into sharing or searching user information sequentially for the proper identification of the application. Initially, the false user identification of the y_u based on behavior monitoring B_M is computed as

$$F_{Id}^{u}(y_u, B_M) = \sum_{t=1} \left(1 \frac{Id_f}{C_A}\right) \frac{F_{Ac}}{C_A} + \frac{(C_A\text{-}C_P)}{(C_A+1)} \left[arg\,min \frac{C_A}{(C_A+1)_{t+1}} \right] \quad \forall t \in F_{Ac} \tag{4}$$

As per the above false user identification $F_{Id}^{u}(y_u, B_M)$, the registration and credential sharing based on the illegitimacy and security demands provides C_P+1 such that

$$F_{Id}^{u}(y_u, B_M+1) = \sum \frac{Id_f}{Y_U\text{-}y_u} + \left(\frac{C_A+1}{C_P}\right) [(C_A+1\text{-}C_P)_t] \; \forall \; t \in F_{Ac} \tag{5}$$

The second user accessing the information based on active time verification helps to build the user identification through behavior and the proof of their credentials for the user on the YouTube platform by suppressing the information accessing on social applications. The behavior-based false user detection process is portrayed in Fig. 3.

The active inactive (sessions) and searches are used as input between the session times in the classification learning. This learning correlates the fixed (or dynamic) sessions with the blocklisted

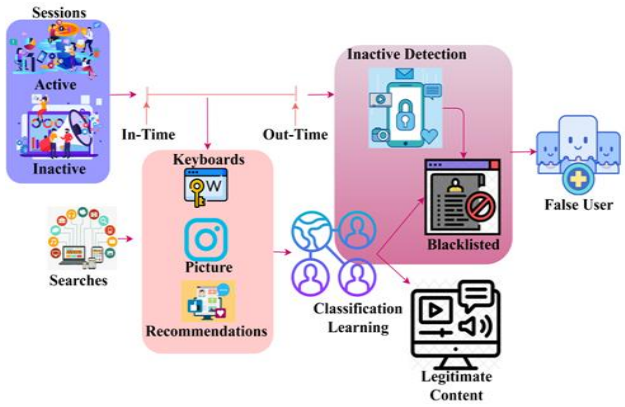

Figure. 3 Behavior-based false user detection

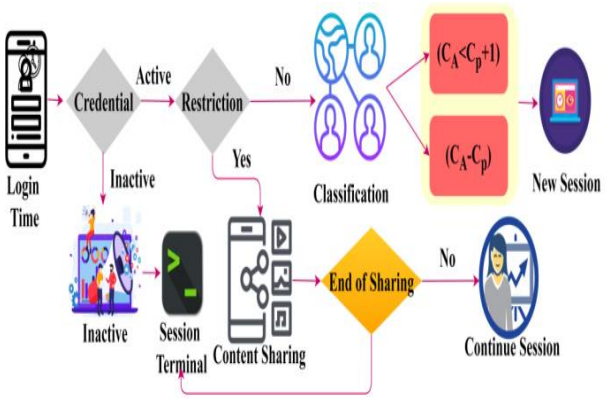

Figure. 4 Active and inactive session classification

contents $\forall B_M$ and B_M+1 . Therefore, the user accounts under this classification are detected as false (Refer to Fig. 3). The reliable condition is the user identification based on registration and credentials shared at different intervals. This consecutive process is performed using classification learning. Classification learning identifies the active, inactive, and false searchers on the YouTube platform $F^u_{Id}(y_u,B_M)$ and $F^u_{Id}(y_u,B_M+1)$, respectively.

Condition (ii) Analysis: In this condition, classification learning is employed for identifying active, inactive, and false searchers based on time verification to prevent idle sessions and false searchers in the YouTube application. From the condition, $F^u_{Id}(y_u,B_M+1)$ is the user behavior and false user identification in the present account; thus, it requires identifiability. However, this Id_f does not identify the available user accessing information, and therefore new time instance is provided. In the recent time sequence admitting to accessing information must satisfy the condition $C_P \leq Y_P$ to prevent additional false searchers. The condition (i) is performed either identifiability or connectivity as in Eq. (3) and is preceded using the active sessions of $F^u_{Id}(y_u,B_M+1)$ until $C_P \leq Y_P$ Condition is satisfied. The active session verification ($Active_{SS}$) is computed for $C_P \leq Y_P$ and $Y_P \leq C_A$ condition based on $F^u_{Id}(y_u,B_M)$ and $F^u_{Id}(y_u,B_M+1)$ as in Eq. (6)

$$\text{Active}_{SS}=\left(1-\frac{Id_{f_t}}{t}\right)+\left(\frac{C_A}{C_P}\right)-\left(\frac{F_{s-1}}{c_v}\right), \text{ for } F^u_{Id}(y_u,B_M) \quad (6)$$

Similarly,

$$\text{In\_Active}_{SS}=\left(1-\frac{F_{s-1}}{t}\right)+\left(\frac{C_A-C_P}{C_A+c_v+1-4}\right) \quad (7)$$

The active verification balances processing time in two sequences, therefore $C_A = C_P+1$ and $(C_A+c_v+1-Y_P)=Y_P$. In the final estimation of In_Active_{SS}, if the condition C_A+c_v+1 is exceeding independent based on the user identification, then the chances of false searchers F_s is observed. This F_s is noticed in computing condition (ii). The active and inactive session classification is presented in Fig. 4.

Based on the login time, the active/ inactive credentials are detected. For the active certificates and restrictions, classification is performed. The conditions $(C_A<C_P+1)$ and (C_A-C_P) are split for a new session. On the other contrary, the inactive sessions are terminated or continued for content sharing (Refer to Fig. 4). In this classification learning, the conditions $Y_P \leq C_A$ and $Y_P \leq C_A+1$ is to be achieved to satisfy precise content access. The false user identification and inactive time verification on the YouTube platform require connectivity and identifiability. Therefore, the static time verification for the above conditions $Y_P \leq C_A$ and $Y_P \leq C_A+1$ are computed in consecutive Eqs. (6) and (7). This straight manner of user inactive time verification is estimated as in Eq. (7)

$$In\_Active_{SS+1}=\left(1-\frac{C_A}{C_P+1}\right)+\left(\frac{C_{a_t}}{F_s-y_u}\right) \quad (8)$$

In the above Eq. (8), the condition inactive time verification for a user accessing content based on $F^u_{Id}(y_u,B_M)$ and $F^u_{Id}(y_u,B_M+1)$ actual behavior monitoring, respectively. In this condition, identifiability and connectivity are provided for either active status or inactive status is verified. Therefore, the number of information, videos, content, etc., accessing the YouTube platform based on the above conditions retards the busy time. Verification for current YouTube application accessing users by augmenting the chances of false user flagging reported data sharing and illegitimacy

by imposing various restrictions on the identified inactive or inaccurate searchers. Therefore, the pursued social media applications are sequentially observed without additional false user identification. The previous fake account detection process of $F_{Id}^{u} \forall c_v\text{-}1(or)\, F_s\text{-}1$ is again verified for inactive time verification and false user identification through the registration and credential-sharing process based on the above condition.

The active or inactive time verification factors based on user accessing the YouTube platform identifies false user identification to analyze the information accessed or false searchers. Therefore, if the user identification is matched with the previous session login, that is the actual user account in social media application instances based on proper user registration and credential sharing. The overall YouTube accessing content here is for monitoring both the analysis $F_{Id}^{u}(y_u, B_M)$ and $F_{Id}^{u}(y_u, B_M+1)$ at the same time interval. The condition for session logins and recommendations is computed with illegitimacy and security analysis based on the user identification process. The proposed method verifies if the user is fake or real until the investigation is processed based on false user restrictions from a community with the others and legitimate information at different intervals is performed for active or inactive time verification on the YouTube provided for use independently.

If the false users are on YouTube with duplicate identification and login session details, the number of fake accounts increases; hence, the minimum false account detection is performed through registration and credential sharing on YouTube. The wrong account user is easily identified based on active time verification. If it once identifies the fake account, then the statement is restricted and blocked, which will reduce the number of false searchers. This classification learning using counterfeit account detection on the YouTube platform under the TBM method is used to reduce false accounts and users. The security demands based on fake account detection and legitimacy mitigation depend on the social media application platform. This learning process from the last known access session using different restrictions on identified fake accounts is used to improve false user identification on social applications.

The suggested TBM technique uses a registration page and information sharing to identify false accounts on the YouTube platform at varying time frames. Accessing user profiles helps reduce modeling latency and verify idle time, but identifying fraudulent profiles is dependent on using

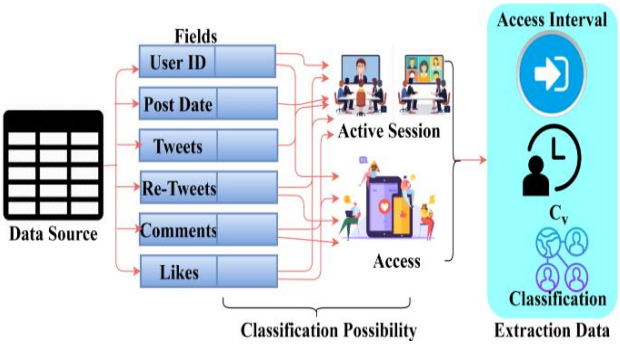

Figure. 5 Data source representation

Table 1. C_v for different features and sessions

Sessions	200	400	600	800	1000	1200	1400
Tweets	455	621	526	758	698	841	930
C_v	0.72	0.62	0.58	0.48	0.42	0.39	0.15
Re-Tweets	36	58	102	159	174	165	212
C_v	0.93	0.81	0.79	0.71	0.64	0.58	0.46
Comments	519	897	714	1247	896	1358	1422
C_v	0.82	0.78	0.72	0.52	0.41	0.39	0.26
Likes	81	96	125	158	163	174	263
C_v	0.92	0.88	0.82	0.78	0.68	0.52	0.43
Collective C_v	3.39	3.09	2.91	2.49	2.15	1.88	1.33

that information to get entry to the product's content and participate in communication marketing. Various transaction logins and suggestions from the learning experience prevent the fake users from accessing other users and accurate data compared with the previous models. The suggested strategy improves performance in user identification by 14.67 percent, information sharing by 11.44 percent, and categorization by 19.36 percent. It decreases the false ratio by 7.73 percent and the delay by 8.85 percent compared to the traditional method.

4. Discussion

The proposed method is analyzed [30] data source. The data stores user index, port, date, re-threat dates, comments, and like attributes. This is split into fake and legitimate users with 8 different fields. C_v is varied between 0.1 and 1, and the sessions run from 100 to `1500. In the classification process, 100 to 800 iterations are varied and used. In

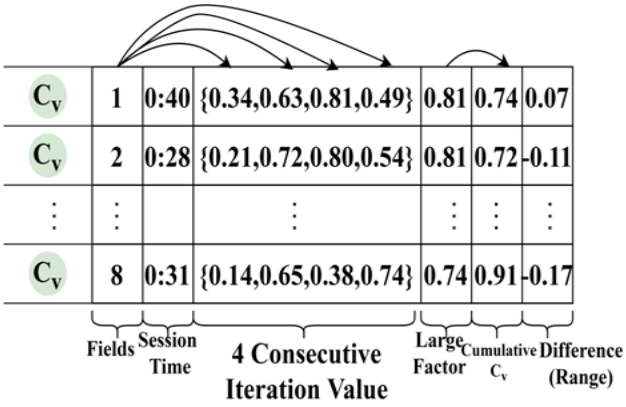

Figure. 6 Active and New Requirement Representation

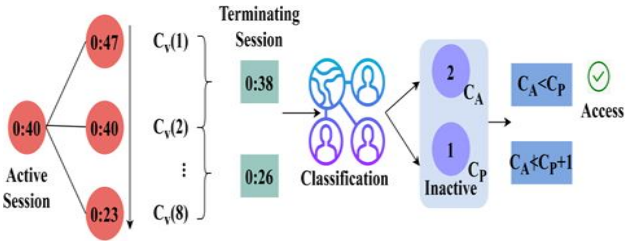

Figure. 7 Active and Access Session

Fig. 5, the data source representation is portrayed.

The data source is extracted for the fields based on classification probability. In the classification, the active section and access-based features are removed. The required data is about access interval, *C_v*, and the variety required for analysis.

Parameters such as user indices, ports, timestamps (including re-threat timestamps), dates, comments, and similar data are all kept in the database. *C_v* represents the number of fields that distinguish between fraudulent and real users. The values of *C_v* are manipulated between 0.1 and 1.

First, the *C_v* is estimated for the threats, re-threats, comments, and likes under different sessions tabulated in Table 1.

In the above table, the C_v is estimated as in Eq. (2) over varying field interconnection. The estimated value is based on mapping classifications for either active session/access. The cumulative C_v is required to detect the low to high values for a minor classification. The classification is performed between two successive access intervals such that F_{ID} are deleted. A single or concurrent iteration is required for further SS estimation in the classification. From the estimate C_v, the active session and new requirements are identified as represented in Fig. 6.

The *-ve* values in the difference require additional sessions, and hence C_v are further increased. The positive value represents the end of a session; the user left the session before termination. Now, the classification is performed for *+ve* and *-ve* values as a mid-range in detecting active and access sessions, as represented below in Fig. 7.

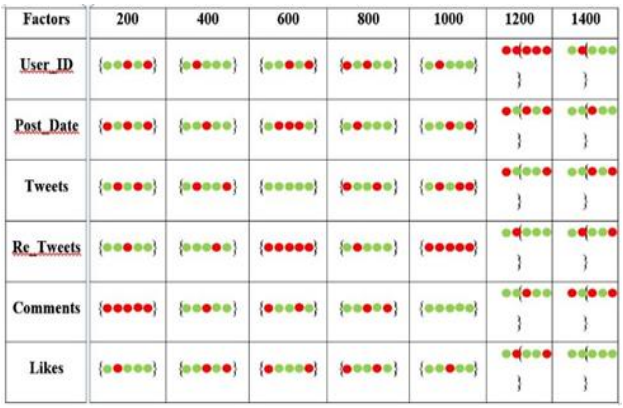

Figure. 8 Maximum Factors Representing False User

The above values are illustrated for a sample of (0:40) session time considered under C_v *(1),* C_v *(2), and* C_v *(8)* . This is used to determine the terminating session as [(0:47+0:29)/2] and [(0:29+0:23)/2 {mid-range}. The parameter is further classified for C_A and C_P as {3(0:47,0:29,0:23)-1} and {2(0:29,0:23)-1} respectively. Now, the condition C_A<C_P is verified by substituting the values. For the above-derived deals, C_A<C_P holds for 2 inactive sessions and, therefore, for any *'n'* sessions, *(n-2)* are the active sessions. From the estimation, the maximum factors representing the false user (detection) are tabulated in Fig. 8.

The green and red dots represent the *C_A<C_P* (Satisfied) and unsatisfied conditions. This intensity varies based on the varying classification; a complete red denotes the fake/ false user account detected. This detection is performed based on 4 consecutive iterations, reducing *C_v*. Therefore, the user_ID exhibits maximum differences, and less *C_v* is considered false. The comparative analysis results are presented in the next section with a discussion. The comparative analysis gives user detection, data sharing, classifications, false ratio, and latency. The methods CAD-SN [26], QuickSquad [29], and FID-ML [25] are considered.

4.1 User detection

This proposed method achieves high user detection on YouTube users at different time intervals. Based on user registration and credential sharing are used for detecting the fake account (Refer to Fig. 9). The model latency and inactive time verification are mitigated based on the visibility for user profile accessing and fake profile identification relies on accessing the platform's

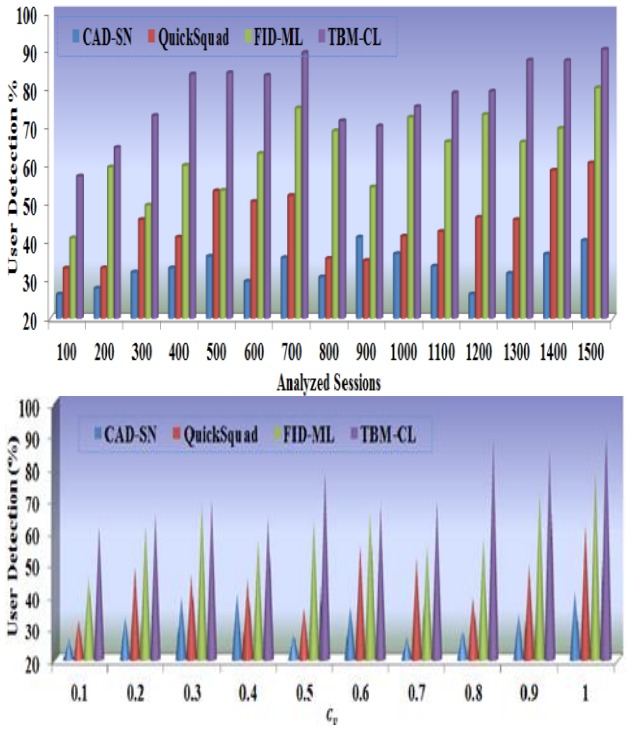

Figure. 9 User detection analysis

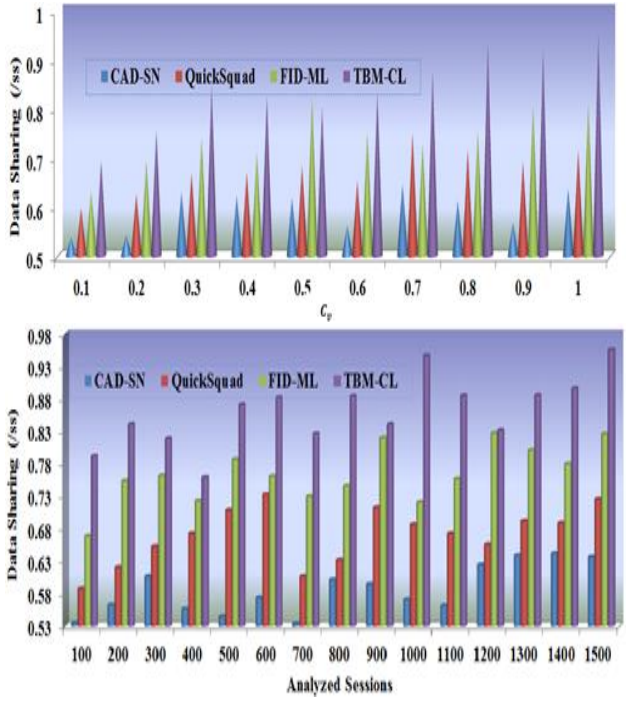

Figure. 10 Data sharing analysis

contents/channel-based promotions based on user identification. The information accessing is found on YouTube content access through a classification learning process. The user registration helps detect fake users and false searches addressing $Y\_P \in [C\_A+1, C\_P\ ]$ based on idle time intervals. The other information access using data sharing and latency detection through illegitimacy and security demands in YouTube accessing any social applications $\in [C\_A+1, C\_P\ ]$ require the user identification process at the entry stage. The user behavior model is used to identify fake users in different instances. Similarly, the false user identification is computed for increasing the security demands and addressing fake profiles on the YouTube platform depending on user credentials and interactions; hence the $F\_s=(c\_v\text{-}F\_Ac)$ is increased and high in user detection.

4.2 Dara sharing

This proposed method achieves high data sharing for user access information and fake accounts detection based on the YouTube platform (Refer to Fig. 10). The false user flagging of reported data sharing and illegitimacy is mitigated based on the $Y_P \in C_A \ \forall (C_A\text{-}C_P)$ condition. The wrong search problem is addressed due to different usage series. The illegitimacy and security demands increase based on accessing information through fake accounts of other users. This problem is managed based on classification learning and information access sessions from the identified account/user's previous login session. Active status is verified at each level of accessing information to reduce fake accounts through the learning process. Therefore, the Id_f is computed for improving user identification along with credential sharing and registration at different time intervals. Therefore, leaving out the active sessions, an inactive and false searcher is to be segregated depending on YouTube access; this fake account detection and monitoring access information have to satisfy two conditions for retaining the current user detection. In the proposed method, user detection is aided by false user identification and increased data sharing.

4.3 Classification

The proposed method for increasing data sharing is high in this classification learning process compared to the other factors in the user information access (Refer to Fig. 11). In this manuscript, fake user identification is used for detecting fake accounts on the YouTube platform through learning to identify $Y\_P \in [C\_A+1, C\_P\ ] \ \forall \ (y\_u+1)$. Based on the condition, the increasing number of fake accounts is identified through illegitimacy and security demands [as in Eq. (4)], then the conditions $(c\_v\text{-}F\_Ac)$ and $F\_Ac/\ Id_f \rightarrow 1$ achieves false user identification is estimated. Based on this method, wrong user flagging and reported data sharing are determined. The maximum fake account

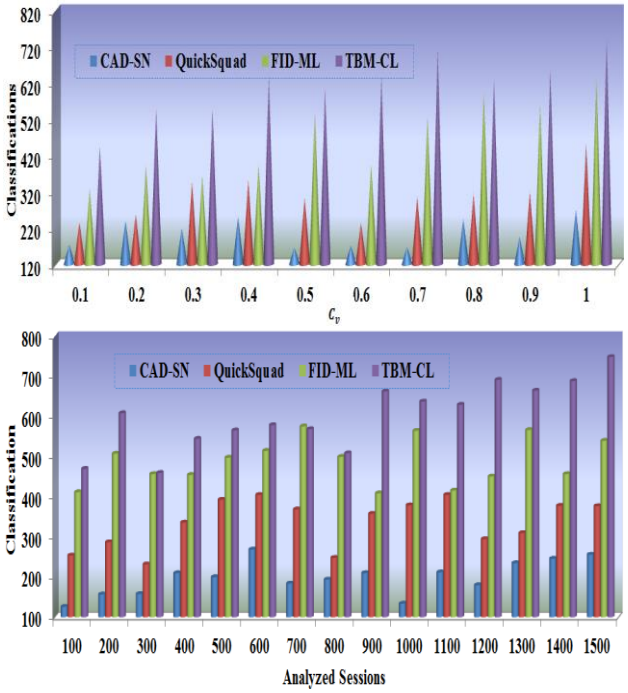

Figure. 11 Classification analysis

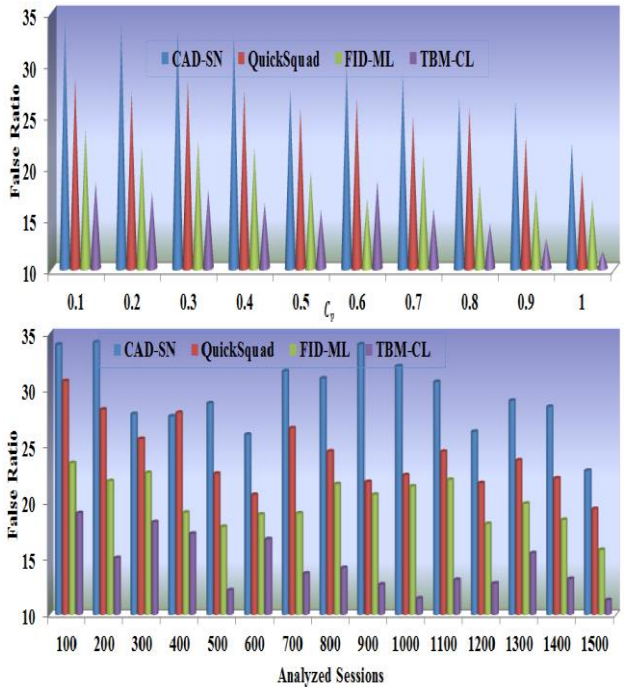

Figure. 12 False ratio analysis

detection due to inactive session time and false searchers is detected. This phony account obtains increasing false searchers and information access, preventing illegitimacy and security demands in a balanced manner. Hence, the access information under different usage series to the connectivity and identifiability is administered as defined in Eqs. (5) and (6) with active time verification. In this proposed method, the monitoring of user behavior depends on legitimate information; hence, the false user is restricted from the community with fewer users.

4.4 False ratio

In this proposed method, latency and false user verification are based on user behavior monitoring as it does not access YouTube content for different users in social applications. The addressing of the fake profile based on the classification learning and user identification is computed from the last known access session for inactive time verification at different time intervals. Counterfeit account detection can be identified for promotions, education, information sharing, etc. Based on this output, the false user on YouTube-like platforms is detected as the sequence of user identification for information accessing through classification learning, preventing fake accounts/users. The session time can be classified into two instances active and inactive are performed without increasing user access session. Instead, the two conditions rely on user identification sequences and provide approximate fake accounts and users on social platforms for each instance based on changes in user registration. In this proposed method, the identifiability function is used for detecting false searchers and achieves a lower false rate, as illustrated in Fig. 12.

4.5 Latency

This proposed method achieves less latency based on accessing the platform's content or

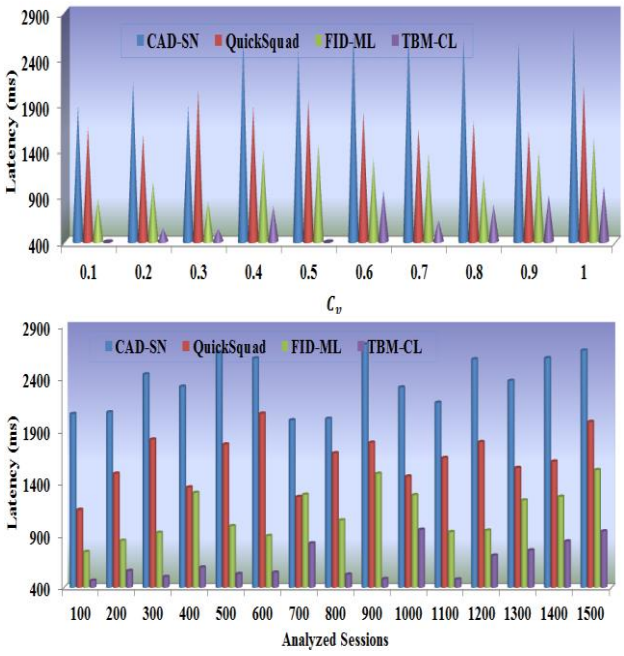

Figure. 13 Latency analysis

Table 2. Comparative analysis result for C_v

Metrics	CAD-SN	Quick Squad	FID-ML	TBM-CL
User Detection %	41.06	62.27	79.42	90.263
Data Sharing (/ss)	0.637	0.716	0.813	0.9508
Classification	270	451	634	737
False Ratio	22.2	19.38	16.75	11.718
Latency (ms)	2736.29	2104.66	1533.1	997.047

Table 3. Comparative analysis result for analyzed sessions

Metrics	CAD-SN	Quick Squad	FID-ML	TBM-CL
User Detection %	40.32	60.57	80.19	90.21
Data Sharing (/ss)	0.637	0.726	0.826	0.9556
Classification	256	377	540	749
False Ratio	22.8	19.4	15.76	11.29
Latency (ms)	2668.2	1985.17	1523.37	938.212

channel-based promotions compared to the other factors, as presented in Fig. 13. The illegitimacy and security demands on YouTube platforms detecting fake accounts/users based on the active, inactive, and false searchers/information accessed by the user are determined. The user registration and credential sharing through fake accounts; this problem is identified and then controlled by the proposed method. It is crucial to prevent data sharing and user detection on the YouTube platform in various instances used for false search reduction. The new user identification through registration is computed for different fake accounts detection for other users, preventing additional false user identification. The data sharing ensures separate session logins and recommendations in social applications and is retained using active time verification as in Eq. (7). Hence, the false users are restricted from the community with legitimate information, and another user at different time intervals through classification learning is used for fake account detection. This data accessing obtained counterfeit account/user detection is processed under additional restrictions. Thus the proposed method verifies the behavioral model for YouTube users, and the latency in detecting fake accounts is reduced. In Tables 2 and 3, the comparative analysis results are tabulated.

Findings: The proposed method maximizes user detection, data sharing, and classification by 14.67%, 11.44%, and 19.36%, respectively. It reduces false ratio and latency by 7.73% and 8.85%, respectively.

Findings: The proposed method maximizes user detection, data sharing, and classification by 14.93%, 11.3%, and 15.93%, respectively. It reduces false ratio and latency by 8.03% and 9.07%, respectively.

5. Conclusion

YouTube's social data, video, etc. sharing platforms are influenced by fake accounts and false users, breaching privacy and legitimacy. Detecting and mitigating the adversaries is mandatory for preserving the features mentioned above. Therefore, this article introduced a temporal behavioral method using a classified learning technique. This method provides legitimate access to data/ content based on user credentials and their behavior. Classification learning identifies active and inactive sessions, and false information leads to active sessions. By classifying the idle sessions, the learning process identifies unknown and incorrect access information. The connectivity between the contents and the user sessions is grouped to prevent exceeding the false ratio. Maximizing data sharing and user detection in the prolonged inactive session classification successively improves identifiability. The proposed method maximizes user detection, data sharing, and variety by 14.93%, 11.3%, and 15.93%, respectively, under varying sessions. It reduces false ratio and latency by 8.03% and 9.07%, respectively.

The suggested strategy increases the maximums for user detection (14.93%), information sharing (11.3%), and classifications (15.93%) across all sessions. It decreases the erroneous ratio by 8.03% and the delay by 9.07%.

Conflicts of interest

The authors declare no conflict of interest.

Author contributions

Conceptualization, Alaa Thamer Mahmood; methodology, Alaa Thamer Mahmood , Raed Kamil Naser; writing—original draft preparation, Alaa Thamer Mahmood, Sura Khalil Abd; writing—review and editing, Alaa Thamer Mahmood; supervision, Raed Kamil Naser; project administration, Alaa Thamer Mahmood.

References

[1] M. Masood, M. Nawaz, K. M. Malik, A. Javed, A. Irtaza, and H. Malik, "Deepfakes generation and detection: state-of-the-art, open challenges, countermeasures, and way forward", *Applied Intelligence*, pp. 1-53, 2022.

[2] M. Choraś, K. Demestichas, A. Giełczyk, Á. Herrero, P. Ksieniewicz, K. Remoundou, D. Urda, and M. Woźniak, "Advanced Machine Learning techniques for fake news (online disinformation) detection: A systematic mapping study", *Applied Soft Computing*, Vol. 101, No. 107050, p. 107050, 2021.

[3] Z. Shahbazi and Y. C. Byun, "Fake media detection based on natural language processing and blockchain approaches", *IEEE Access*, Vol. 9, pp. 128442–128453, 2021.

[4] D. Plotkina, A. Munzel, and J. Pallud, "Illusions of truth—Experimental insights into human and algorithmic detections of fake online reviews", *Journal of Business Research*, Vol. 109, pp. 511–523, 2020.

[5] D. Zimmermann, C. Noll, L. Gräßer, K. U. Hugger, L. M. Braun, T. Nowak, and K. Kaspar, "Influencers on YouTube: a quantitative study on young people's use and perception of videos about political and societal topics", *Current Psychology*, Vol. 41, pp. 6808-6824, 2020.

[6] S. Kang, S. Dove, H. Ebright, S. Morales, and H. Kim, "Does virtual reality affect behavioral intention? Testing engagement processes in a K-Pop video on YouTube", *Computers in Human Behavior*, Vol. 123, No. 106875, p. 106875, 2021.

[7] X. Chen, L. Niu, A. Veeraraghavan, and A. Sabharwal, "FaceEngage: Robust estimation of gameplay engagement from user-contributed (YouTube) videos", *IEEE Transactions on Affective Computing*, Vol. 13, No. 2, pp. 651–665, 2022.

[8] T. Ahmad, K. Sattar, and A. Akram, "Medical professionalism videos on YouTube: Content exploration and appraisal of user engagement", *Saudi Journal of Biological Sciences*, Vol. 27, No. 9, pp. 2287–2292, 2020.

[9] Tyagi and D. Yadav, "A detailed analysis of image and video forgery detection techniques", *The Visual Computer*, pp. 1-22, 2022.

[10] Y. Li, Z. Fan, X. Yuan, and X. Zhang, "Recognizing fake information through a developed feature scheme: A user study of health misinformation on social media in China", *Information Processing & Management*, Vol. 59, No. 1, p. 102769, 2022.

[11] J. Brandt, K. Buckingham, C. Buntain, W. Anderson, S. Ray, J. R. Pool, and N. Ferrari, "Identifying social media user demographics and topic diversity with computational social science: a case study of a major international policy forum", *Journal of Computational Social Science*, Vol. 3, No. 1, pp. 167–188, 2020.

[12] B. Cannelli and M. Musso, "Social media as part of personal digital archives: exploring users' practices and service providers' policies regarding the preservation of digital memories", *Archival Science*, Vol. 22, No. 2, pp. 259–283, 2022.

[13] A. M. Abbas, "Social network analysis using deep learning: applications and schemes", *Social Network Analysis and Mining*, Vol. 11, pp. 1–21, 2021.

[14] Y. Ryoo, H. Yu, and E. Han, "Political YouTube Channel Reputation (PYCR): Development and validation of a multidimensional scale", *Telematics and Informatics*, Vol. 61, No. 101606, p. 101606, 2021.

[15] T. Acosta, P. A. Vargas, J. Z. Miranda, and S. L. Mora, "Web accessibility evaluation of videos published on YouTube by worldwide top-ranking universities", *IEEE Access*, Vol. 8, pp. 110994–111011, 2020.

[16] H. Oh, "A YouTube spam comments detection scheme using cascaded ensemble machine learning model", *IEEE Access*, Vol. 9, pp. 144121–144128, 2021.

[17] K. Yousaf and T. Nawaz, "A deep learning-based approach for inappropriate content detection and classification of YouTube videos", *IEEE Access*, Vol. 10, pp. 16283–16298, 2022.

[18] J. V. D. Souza, J. Gomes, F. M. D. S. Filho, A. M. D. O. Julio, and J. F. D. Souza, "A systematic mapping on automatic classification of fake news in social media", *Social Network Analysis and Mining*, Vol. 10, No. 1, 2020.

[19] H. Jelodar, Y. Wang, M. Rabbani, S. B. Ahmadi, L. Boukela, R. Zhao, and R. S. Larik, "A NLP framework based on meaningful latent-topic detection and sentiment analysis via fuzzy lattice reasoning on youtube comments", *Multimedia Tools and Applications*, Vol. 80, No. 3, pp. 4155–4181, 2021.

[20] B. Jang, S. Jeong, and C. K. Kim, "Distance-based customer detection in fake follower markets", *Information System*, Vol. 81, pp. 104–116, 2019.

[21] A. Mulahuwaish, K. Gyorick, K. Z. Ghafoor, H. S. Maghdid, and D. B. Rawat, "Efficient

classification model of web news documents using machine learning algorithms for accurate information", *Computers and Security*, Vol. 98, No. 102006, p. 102006, 2020.

[22] R. M. O. Gaona, M. P. Boix, and J. L. M. Moreno, "Extent prediction of the information and influence propagation in online social networks", *Computational and Mathematical Organization Theory*, Vol. 27, No. 2, pp. 195–230, 2021.

[23] T. Chauhan and H. Palivela, "Optimization and improvement of fake news detection using deep learning approaches for societal benefit", *International Journal of Information Management Data Insights*, Vol. 1, No. 2, p. 100051, 2021.

[24] Zappin, H. Malik, E. M. Shakshuki, and D. A. Dampier, "YouTube monetization and censorship by proxy: A machine learning prospective", *Procedia Computer Science*, Vol. 198, pp. 23–32, 2022.

[25] E. V. D. Walt and J. Eloff, "Using machine learning to detect fake identities: Bots vs humans", *IEEE Access*, Vol. 6, pp. 6540–6549, 2018.

[26] M. Egele, G. Stringhini, C. Kruegel, and G. Vigna, "Towards detecting compromised accounts on social networks", *IEEE Transactions on Dependable and Secure Computing*, Vol. 14, No. 4, pp. 447–460, 2017.

[27] J. Kang, H. Choi, and H. Lee, "Deep recurrent convolutional networks for inferring user interests from social media", *Journal of Intelligent Information Systems*, Vol. 52, No. 1, pp. 191–209, 2019.

[28] B. Bebensee, N. Nazarov, and B. T. Zhang, "Leveraging node neighborhoods and egograph topology for better bot detection in social graphs", *Social Network Analysis and Mining*, Vol. 11, No. 1, 2021.

[29] X. Jiang, Q. Li, Z. Ma, M. Dong, J. Wu, and D. Guo, "QuickSquad: A new single-machine graph computing framework for detecting fake accounts in large-scale social networks", *Peer-to- Peer Networking and Applications*, Vol. 12, No. 5, pp. 1385–1402, 2019.

[30] Y. Lu, "Social Network fake account dataset." 18-Nov-2019.