Classification of Instagram fake users using supervised machine learning algorithms

Vertika Singh, Naman Tolasaria,Patel Meet Alpeshkumar, Shreyash Bartwal

SCOPE, VIT University, Chennai Campus, Tamil Nadu, India

Abstract

In the contemporary era, online social networks have become integral to social life, revolutionizing the way individuals manage their social connections. While enhancing accessibility and immediacy, these networks have concurrently given rise to challenges, notably the proliferation of fraudulent profiles and online impersonation. This paper proposes an application designed to detect and neutralize such dishonest entities, with a focus on safeguarding companies from potential fraud. The user-centric design of the application ensures accessibility for investigative agencies, particularly the criminal branch, facilitating navigation of complex social media landscapes and integration with existing investigative procedures.

1. Introduction

1.1 About social media

Social media, as interactive Web 2.0 Internet-based applications, enable the creation and sharing of ideas, content, and connections. User-generated content forms the essence of social media, with individuals creating profiles to connect with others. While facilitating online social networks, social media also serves various purposes, from learning and self-promotion to memory-keeping and idea development.

1.2 Historical Timeline of social media

1) The evolution of social media spans several decades, from early forms of internet communication like email and Bulletin Board Systems (BBS) to the emergence of modern platforms such as Facebook, Twitter, and TikTok. The continuous evolution of social media reflects shifts in user behavior and the introduction of new features and platforms.

2. Fake Social-Media (Instagram) Profiles and Their Growing Prevalence

Instagram, ranking third among social media platforms in terms of active users, faces challenges related to fake profiles. These profiles, often associated with identity theft, social engineering, and the spread of harmful content, pose significant security risks to users. Business owners engaging in influencer marketing also encounter issues with overpaying for endorsements due to the prevalence of fake followers. To address these concerns, Instagram employs measures such as algorithms and reporting mechanisms, urging users to be vigilant.

2.1 Security Threats Posed by Fake Profiles

Fake profiles not only contribute to identity theft but also serve as conduits for harassment, cyberbullying, fraud, scams, and data harvesting. The extensive use of fraudulent accounts undermines user trust in the platform, affecting the overall user experience and jeopardizing the platform's ability to maintain a safe environment. Instagram utilizes various measures to combat fake profiles, including the examination of user metadata and machine learning algorithms to enhance detection capabilities.

3. Proposed Solution: Machine Learning for Profile Verification

To address the challenges posed by fake profiles on social media platforms like Instagram, a potential solution involves examining user metadata and employing machine learning algorithms. These algorithms can analyze posting trends, content engagement, and user behavior to differentiate between genuine and fraudulent profiles. Additionally, the analysis of metadata, including account creation date, IP address history, and device details, provides valuable insights into profile authenticity. Machine learning models, through learning from past data on known fake accounts, offer a robust method for enhancing online social landscape safety and security.

4. Literature Review

A variety of methods were employed to classify profiles according to account activity, the quantity of requests that were fulfilled, the quantity of messages that were sent, among other things. System graphs serve as the foundation for the models. Others have attempted to distinguish between cyborgs and robots utilizing certain methodologies. A summary of a few earlier studies is supplied beneath. If specific terms are found in a message, it's deemed unsolicited. This supposition has been employed to identify phony social media accounts. Such Pattern matching algorithms were used to locate phrases on social media. But this criterion falters substantially from the regular creation and use of new terminologies.

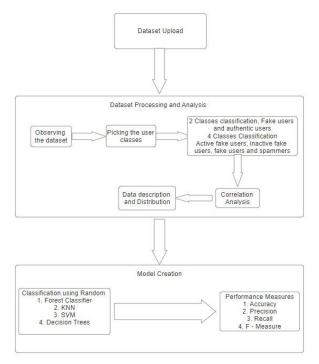
Different methods were used to group profiles according to variables including account activity, the number of requests that were answered, the quantity of messages that were delivered, and other characteristics. A framework based on graphs is employed in the models. Others attempted to distinguish between robots and cyborgs using particular techniques. Below is a list of some previous research. Messages can be categorised as spam by using specific terms. This concept has been applied to social media fraud detection profiles. Using pattern matching tools, these phrases were found on social media. However, one major disadvantage of this criterion is the constant generation and application of new nomenclature. Acronyms such as gbu, gn, and LOL are becoming common on Twitter.

As previously mentioned, the concept of a false user can now include human users as well. Many studies, however, only classify phoney users as bot users. The widely used fake project dataset is exclusively made up of bot users; it was obtained through CAPTCHA validation and the market for bot users. Table 1 presents a collection of false users' classifications. The majority of research used Twitter as the platform and employed supervised, feature-based detection techniques. Using several feature sets, supervised machine learning (ML) approaches were used to detect phoney accounts in the fake project dataset. Facebook has more functions connected to media than Twitter does. Identification can be achieved through features like likes (given and received), shares, tags, and comments. There is a report that classified fake accounts on Instagram. Nevertheless, only metadata features were employed, and human judgement was used to determine whether a user was real or fake rather than purchasing fictitious accounts from bot-selling marketplaces. Instagram is becoming more and more popular in influencer marketing, but it doesn't seem to be a popular platform for research.

5. Methodology

Data collected from both real and fictitious consumers is the first step in this study. Since media data cannot be obtained from private users, only their metadata could be obtained, all of them were eliminated. Instagram provides the following metadata: username, complete name, bio, link, profile photo, number of posts, following, and followers. These attributes will be extracted following data collection, and a correlation analysis will be performed. Following the configuration of the features, the users will be categorized using machine learning algorithms. In this study four user classes real users and three types of fraudulent users (active, inactive, and spammers) will be discovered. The foundation of these classes is manual behavior observation.

Block Diagram



6. DATASET

There are five different sources of the features, i.e. metadata, media info, engagement, media tag, media similarity.

- 1. The Metadata(M) inculcates pos, flg, flr, bl, pic, lin -
 - pos the number of posts that the user has posted in total
 - flg number of accounts followed by the user
 - flr number of followers on the user's account
 - bl length of the user's biography
 - *pic* availability of profile picture on the user's account (Value 1 if picture is available and 0 if picture is not present)
 - *lin* availability of any external links (Value 1 if link is available on the profile and 0 if not)
- 2. Media Info involves the user's media involving engagement info
 - *cl* average length of the captions in the posts or any engagement on the app
 - cz percentage (0.0 0.1) of captions that has a length almost equal to zero or non-significant.
 - ni percentage (0.0 0.1) of media that doesn't contain any image.
 - ENGAGEMENT INVOLVES THE USER'S ACTIVENESS ON THE APP -
 - ERL IT IS THE ENGAGEMENT RATE WHICH IS DEFINED AS NUMBER OF LIKES DIVIDED BY NUMBER OF MEDIA AND NUMBER OF FOLLOWERS
 - ERC IT IS DEFINED AS THE NUMBER OF COMMENTS DIVIDED BY NUMBER OF MEDIA AND NUMBER OF FOLLOWERS
 - Media tags involves the statistics for tag usage by the user –
 - lt percentage (0.0 0.1) of posts tagged with location
 - hc the average count of hashtags used in a post
 - 5. Media similarity –
 - pr average count of promotional keywords (i.e. contest, repost, mention) used in hashtags by the user
 - fo average count of follower hunter keywords (i.e. follow, like, follow for follow) used in hashtags by the user
 - cs average cosine similarity between all pair of two posts a user has
 - interval info –
 - pi average interval between posts (in hours)

7.Data Collected

We have used the dataset provided in the paper: Linqing Liu, Yao Lu, Ye Luo, Renxian Zhang, Laurent Itti and Jianwei Lu. "Detecting "Smart" Spammers on Social Network: A Topic Model Approach." Proceedings of NAACL-HLT. 2016.

8.Correlation Analysis

To obtain the correlation table by using the Pearson Correlation values. Upon evaluation we see that there is no strong correlation between any of the variables but two correlation values with the greatest values, bl-lin (0.47) and erc-erl (0.44), are regarded as satisfactory and moderate. The relationship between the length of the biography and the availability of links indicates that users will probably include a link if the biography is lengthy. The comments and the likes correlation demonstrate that there is a linear relationship between the quantity of likes and remarks.

Table present below:

	pos	fl w	fl g	bl	pi c	li n	cl	c z	ni	e rl	e r	lt	h c	p r	fo	es	pi
p o s	1.000	0. 13 50 49	0. 06 12 12	0. 16 01 35	0. 05 19 38	0. 16 95 30	0. 18 48 61	0 0 7 9 3 8	0. 0 7 7 7 8 6	0. 0 3 0 3 8 5	0. 0. 3 8 6 8 8	0. 0 2 9 8 5	0. 0 1 4 9 6 8	0. 0 2 0 7 6	0. 01 06 36	0. 01 59 24	0.0 856 87
f I W	0.135 049	1. 00 00 00	0. 00 72 97	0. 04 05 20	0. 01 00 87	0. 05 10 48	0. 03 34 44	0 0 2 2 6 3 0	0. 0 2 5 2 3 5	0. 0 0 6 1 7 9	0. 0 0 8 2 6 5	0. 0 1 2 0 7 6	0. 0 0 5 4 0 7	0. 0 0 1 5 5	0. 00 14 94	0. 01 20 64	0.0 131 27
f l g	0.061 212	0. 00 72 97	1. 00 00 00	0. 00 93 65	0. 12 94 41	0. 03 34 37	0. 06 31 18	0 1 6 6 9 8 7	0. 0 7 7 3 9	0. 0 2 3 7 3 3	0. 0 2 4 2 6 7	0. 1 1 5 4 3 7	0. 0 4 2 0 8 7	0. 0 4 6 6 4 2	0. 02 29 82	0. 22 90 80	0.0 847 63
b I	0.160 135	0. 04 05 20	0. 00 93 65	1. 00 00 00	0. 16 62 68	0. 47 17 50	0. 35 01 13	7 2 7 2 3 4 4	0. 1 4 1 1 1 5	0. 0 3 9 5 0 9	0. 0 6 0 4 3 8	0. 2 1 8 2 4 6	0. 1 6 0 8 8	0. 0 2 7 1 9	0. 01 77 02	0. 13 59 14	0.1 115 49
P i c	0.051 938	0. 01 00 87	0. 12 94 41	0. 16 62 68	1. 00 00 00	0. 12 42 27	0. 11 95 68	0 0 6 2 6 2 7	0. 1 2 7 7 8 7	0. 0 1 9 8 0 4	0. 0 2 1 2 3 1	0. 1 2 7 4 6	0. 0 6 2 9 0 6	0. 0 2 2 6 0 8	0. 00 12 48	0. 26 45 26	0.0 844 05
l i n	0.169 530	0. 05 10 48	0. 03 34 37	0. 47 17 50	0. 12 42 27	1. 00 00 00	0. 30 07 96	0 2 3 9 5 4 8	0. 1 3 4 8 7	0. 0 4 5 4 3 9	0. 0 6 7 3 7	0. 1 9 6 8 6 7	0. 0 9 3 8 2 0	0. 0 3 9 4 5	0. 00 95 35	0. 09 61 47	0.0 967 17
c l	0.184 861	0. 03 34 44	0. 06 31 18	0. 35 01 13	0. 11 95 68	0. 30 07 96	1. 00 00 00	0 3 5 3 7 2 2	0. 1 0 9 3 1	0. 0 3 9 9 3 8	0. 0 5 0 6 1 7	0. 0 8 4 0 1	0. 1 8 5 3 8 3	0. 2 0 7 8 1 8	0. 05 55 34	0. 08 98 11	0.1 127 08
c z	0.079 381	0. 02 26 30	0. 16 69 87	0. 27 23 44	0. 06 26 27	0. 23 95 48	0. 35 37 22	1 0 0 0 0 0	0. 1 2 5 5 2 7	0. 0 7 8 0 8 4	0. 0 8 5 5 0	0. 2 0 5 2 4 3	0. 2 1 8 1 6	0. 0 6 0 8 7 4	0. 04 57 88	0. 32 02 40	0.0 633 16
n i	0.077 786	0. 02 52 35	0. 07 73 94	0. 14 11 15	0. 12 77 87	0. 13 48 77	0. 10 93 10	0 1 2 5 5 2 7	1. 0 0 0 0 0	0. 0 2 4 2 0 1	0. 0 3 3 4 4 2	0. 2 1 9 1 2 8	0. 0 4 8 0 6 6	0. 0 2 3 0 1 5	0. 00 78 50	0. 24 06 86	0.0 005 21
e r l	0.030 385	0. 00 61 79	0. 02 37 33	0. 03 95 09	0. 01 98 04	0. 04 54 39	0. 03 99 38	0 0 7 8 0 8 4	0. 0 2 4 2 0 1	1. 0 0 0 0 0	0. 4 4 3 5 6 7	0. 0 1 4 6 7 5	0. 0 2 3 9 9	0. 0 0 4 5 3 8	0. 03 40 61	0. 03 18 88	0.0 007 53

_								_					_	_	_		
e	0.020	-	-	-	-	-	-	0	-	0.	1.	-	0.	0.	0.	-	0.0
r	0.038 688	0. 00	0. 02	0. 06	0. 02	0. 06	0. 05	0	0.	4	0	0.	0	0	03 13	0. 06	151 97
С	000	82	42	04	12	73	06	8	3	3	0	1	0	5	12	59	91
		65	67	38	31	79	17	5	3	5	0	6	9	4	1.2	78	
								5	4	6	0	2	4	8			
								0	4	7	0	9	2	0			
								0	2			7					
1	0.029	0.	-	0.	0.	0.	0.	-	0.	-	-	1.	0.	-	0.	-	0.0
t	853	01	0.	21	12 74	19	08	0	2	0.	0.	0	1	0.	00	0.	675 49
		20 76	11 54	82 46	61	68 67	40 10	2	1	0	0	0	2	0 6	08 64	27 26	49
		,,,	37		0.	0,	10	0	1	4	6	0	4	2	٥.	32	
								5	2	6	2	0	3	4			
								2	8	7	9	0	5	5			
								4		5	7			4			
1.	0.014	0		0	0	0	0	3	0	0	0	0	1	0	0		0.0
h c	0.014 968	0. 00	0.	0. 16	0. 06	0. 09	0. 18	0	0.	0.	0.	0. 1	1. 0	0. 1	0. 33	0.	0.0 149
۱	700	54	04	08	29	38	53		4	2	3	2	0	1	69	14	38
		07	20	80	06	20	83	2	8	3	0	9	0	9	90	91	50
			87					1	0	9	9	4	0	5		87	
								8	6	9	4	3	0	8			
								1	6	6	2	5	0	3			
								6									
р	0.020	-	_	_	0.	-	0.		-	-	0.	_	0.	1.	0.	-	_
r	762	0.	0.	0.	02	0.	20	0	0.	0.	0.	0.	1	0	08	0.	0.0
		00	04	02	26	03	78		0	0	1	0	1	0	01	05	183
		15	66	71	08	94	18	0	2	0	5	6	9	0	95	36	84
		51	42	92		53		6	3	4	4	2	5	0		94	
								0	0	5	8	4	8	0			
								8	1 5	3 8	0	5 4	3	0			
								4	3	٥		4					
f	-	-	0.	0.	0.	-	0.	-	-	0.	0.	0.	0.	0.	1.	-	-
o	0.010	0.	02	01	00	0.	05	0	0.	0	0	0	3	0	00	0.	0.0
	636	00	29	77	12	00	55		0	3	3	0	3	8	00	02	122
		14	82	02	48	95	34	0	0	4	1	0	6	0	00	16	04
		94				35		5	7 8	6	3	8 6	9	1		08	
								7	5	1	2	4	0	5			
								8	0	•	_	ı.					
								8									
c	-	-	0.	-	-	-	-	0	-	-	-	-	-	-	-	1.	-
s	0.015	0.	22	0.	0.	0.	0.		0.	0.	0.	0.	0.	0.	0.	00	0.1
	924	01	90	13	26	09	08	3	2	0	0	2	1	0	02	00	438
		20	80	59	45	61	98	2	4	3	6	7	4	5	16	00	05
		64		14	26	47	11	0	0	1	5	2	9	3	08		
								2	6	8	9	6	1	6			
								4	8	8	7	3	8	9			
								0	6	8	8	2	7	4			
р	-	-	-	-	0.	-	-	0	0.	0.	0.	0.	0.	-	-	-	1.0
i	0.085	0.	0.	0.	08	0.	0.		0	0	0	0	0	0.	0.	0.	000
	687	01	08	11	44	09	11	0	0	0	1	6	1	0	01	14	00
		31	47	15	05	67	27	6	0	0	5	7	4	1	22	38	
		27	63	49		17	08	3	5	7	1	5	9	8	04	05	
								3	2	5	9	4	3	3			
								1	1	3	7	9	8	8			
								6	-		•		-	4			
1								-									

Data description and distribution

	po s	fl w	flg	bl	pi c	lin	cl	cz	ni	erl	er c	lt	hc	pr	f o	cs	pi
c o u n t	65 32 6. 00 00 00	6. 53 26 00 e+ 04	65 32 6. 00 00 00	65 32 6. 00 00	65 32 6. 00 00 00	65 32 6. 00 00 00	653 26. 000 000	65326 .00000	65 32 6. 00 00 00	65 32 6. 00 00 00							
m e a n	17 6. 57 12 27	1. 18 30 77 e+ 03	23 10 .5 19 21 1	57 .4 97 06 1	0. 95 17 65	0. 28 19 25	13 6. 52 03 44	0. 25 00 12	0. 19 32 29	19 .1 46 64 1	1. 13 94 21	0. 20 88 77	0. 50 77 96	0.0 327 45	0 · 0 5 2 8 3 7	0. 29 90 98	49 6. 47 51 93
s t d	72 3. 47 06 55	2. 17 08 02 e+ 04	25 92 .0 96 10 4	64 .1 29 26 0	0. 21 42 64	0. 44 99 40	21 5. 71 44 86	0. 33 78 91	0. 25 29 39	12 1. 04 75 70	5. 81 06 27	0. 30 03 62	1. 15 69 21	0.2 209 87	0 . 5 1 9 0 5 8	0. 34 96 04	94 4. 90 53 89
m i n	0. 00 00 00	0. 00 00 00 e+ 00	0. 00 00 00	0. 00 00 00	0. 00 00 00	0. 00 00 00	1. 00 00 00	0. 00 00 00	0. 00 00 00	0. 00 00 00	0. 00 00 00	0. 00 00 00	0. 00 00 00	0.0 000 00	0 .000000	0. 00 00 00	0. 00 00 00
2 5 %	6. 00 00 00	1. 23 00 00 e+ 02	39 4. 00 00 00	0. 00 00 00	1. 00 00 00	0. 00 00 00	8. 00 00 00	0. 00 00 00	0. 00 00 00	2. 73 00 00	0. 08 00 00	0. 00 00 00	0. 00 00 00	0.0 000 00	0 .000000	0. 03 33 27	24 .5 71 45 8
5 0 %	30 .0 00 00 0	3. 38 00 00	99 5. 00 00 00	32 .0 00 00 0	1. 00 00 00	0. 00 00 00	46 .0 00 00 0	0. 05 55 56	0. 05 90 00	9. 45 00 00	0. 44 00 00	0. 00 00 00	0. 07 70 00	0.0 000 00	0 0 0	0. 13 69 15	18 3. 22 79 43
7 5 %	12 4. 00 00 00	8. 17 00 00 e+ 02	36 00 .0 00 00	11 0. 00 00 00	1. 00 00 00	1. 00 00 00	17 0. 00 00 00	0. 44 44 44	0. 33 30 00	18 .6 80 00 0	1. 04 00 00	0. 33 30 00	0. 61 10 00	0.0 000 00	0 .00000	0. 45 63 42	58 0. 77 19 27
m a x	76 20 0. 00 00 00	3. 90 00 00 e+ 06	88 00 .0 00 00 0	55 5. 00 00 00	1. 00 00 00	1. 00 00 00	36 44 .0 00 00 0	1. 00 00 00	1. 00 00 00	26 65 0. 00 00	10 09 .0 90 02 7	1. 00 00 00	30 .0 00 00 0	20. 000 000	58.00000	1. 00 00 00	26 78 6. 13 47 66

The Correlation Matrix shows the measure of dependency of variables. As seen in the table no strong correlation can be seen.

Correlation between lin(external link) and bl(bio length) is moderately high(0.471750) this tells us that the users with long bio will most likely put a link in their profile. Erc and erl also show moderate correlation(0.443567) this means that the number of likes is linearly related to number of comments.other variables with high correlation which is on the weaker side include:

Cz and cl with a correlation of -0.353722, the purpose of the cz is to enhance the cl as they are inversely dependent. Typically, fake users will upload media with almost no caption> in the dataset the percentage of fake users with almost zero caption is high.

Correlation between bl and cl is 0.350113, users with lengthy captions appear to have a higher association between bl (biography length) and cl (caption length) as they tend to have longer biographies help in identifying authentic accounts.

Fo and hc also have good correlation as fo is a subset of hc.

Statistics of features available in the two datasets-

Metadata (pos, bl, image, lin): The quantity of posts made by fake users is nearly equal to that of real users. The spammers, however, have a notably larger number of posts. The longest biography text is seen among authentic users, and many of them include a link. Of inactive users, 76% of them just have a profile while other classes all nearly have profile pictures.

Follow info (flw, flg): Authentic users have the highest followers count, but lowest following count. In contrast, fake users have a lower followers count, but a higher following count if compared to the authentic users. This indicates that fake users like to follow others to increase their presence.

Engagement (erl, erc): When compared to real users, fake users will get more likes. On the other hand, genuine users get more comments. This suggests that getting comments is more difficult, thus helping in finding the real users. Fake users typically follow other fake users in terms of likes, in order for them to get automatic likes.

Media information (cl, cz, ni): Real users have more detailed captions and less zero captions than fraudulent users.

Media tags (lt, hc): When compared to fraudulent users, real users utilize location tags and hashtags more frequently. To draw users, spammers utilize more hashtags.

- Media similarity (pr, fo, cs, pi): A lower cs (cosine similarity) value is associated with authentic users. Thus, the majority of their posts are distinct from their earlier postings, in contrast to phony users. Those who spam have the highest FO(follow keywords) and PR(promotional keywords).

	po	flw	flg	bl	pi	li	cl	cz	ni	er	er	lt	h	р	fo	cs	pi
	s				С	n				1	С		С	r			
c																	
1																	
a																	
S																	
S	10	00	21		0	0	1.1	0	0		0	0	0	0		0	20
a	18 6.	82	31 42.	51 .1	0. 9	0. 1	11 7.	0.	0.	9. 85	0.	0. 1	0.	0.	0.	0.	39
	02	9.5 37	12	24	8	4	73	9	2	72	4	7	4	1	0	2	8. 96
	97	41	78	19	9	9	93	8	4	96	3	0	8	4	8	2	48
	83	5	41	19	8	4	40	2	0	96	6	6	7	2	5	8	86
	0.5	3	41	1	7	1	40	0	2		3	9	0	9	7	2	00
					9	1		0	7		4	9	3	7	9	5	
i	1.	21	38	12	0.	0.	11	0.	0.	38	1.	0.	0.	0.	0.	0.	31
•	80	5.7	16.	.4	7	0.	.7	3	0.	.2	8	0.	0.	0.	0.	6	5.
	80	14	52	75	6	3	05	4	5	05	5	4	9	o	0	4	99
	39	28	20	87	5	0	75	1	6	26	9	7	4	0	o	2	67
	37	6	40	4	4	6	4	5	8	5	6	1	7	5	4	7	23
					7	1		6	3		2	1	0	3	5	4	
					5	9		8	9		3	8	3	8	5	5	
r	12	11	12	61	0.	0.	12	0.	0.	20	1.	0.	0.	0.	0.	0.	73
	3.	93.	04.	.8	9	2	0.	2	2	.5	5	3	6	0	0	1	1.
	94	70	07	82	8	2	67	0	8	70	3	4	0	0	3	6	10
	52	80	97	57	6	1	85	5	4	57	3	4	7	7	8	6	13
	16	74	82	8	2	6	75	8	4	5	4	7	4	6	3	5	68
					0	2		6	1		3	1	6	1	6	4	
					8	6		2	5		8	2	9	8	8	3	
s	29	10	30	59	0.	0.	23	0.	0.	14	0.	0.	0.	0.	0.	0.	32
	7.	81.	52.	.3	9	1	6.	3	1	.3	7	1	7	1	2	3	7.
	27	28	56	92	8	6	93	1	6	28	2	1	7	6	2	7	96
	90	57	94	77	2	8	91	8	7	24	2	3	1	0	2	6	84
	61	84	24	0	7	6	99	8	2	0	2	9	6	7	4	9	66
					5	6		5	7		5	0	2	5	8	7	
					4	4		8	4		2	4	8	7	9	6	

9. Machine Learning Models used

We have used five machine learning algorithms for the classification in our research. We have basically done two types of classification, the first one being a 2-class classification for detecting the real and fake users and the second one being a 4-class classification for segregating the users into 4 classes namely active fake users, inactive fake users, spammers and authentic users.

The algorithms used by us are Random Forest Classifier, KNN, SVM and Decision Trees.

10.Model Analysis

s.no	Algo	2 clas	s class	ificatio	on	4 class classification					
•		acc	prec	reca ll	F1	acc	Prec	reca 11	F1		
1	Random forest	89.6 3	91	90	90	89.1 5	89	89	89		
2	Knn	74.0 5	74	74	74	55.1 0	56	55	55		
3	Svm(polyn omial)	63.9 5	-	-	60.0 9	36.1 8	-	-	60.0 9		
	Svm(rbf)	53.8 3	-	-	41.5 9	29.6 4	-	-	41.5 9		
4	Decision trees	87.2 4	89	87	87	72.2 2	72	72	71		

Random forest has given better results in every aspect so we have used that in our project. Random Forest even performs well in the 4-class classification while other algorithms find it difficult. The top four predictors in the 4-class categorization are pos, flw, bl, and flg. The top five predictors in the two-class categorization are pos, flw, lin, flg, and bl. All of these metadata values are easily obtainable, even for private users hance making our model successful even without all the values.

We achieve high accuracy in 2 class classification unlike in the 4 class classification. The 4-class classification's lower accuracy result can be attributed to the significant dependence on media data for the distinguishing of fake user types.

11.Results with Discussion

Hence upon training and testing various ML models we can conclude that using Random Forest model we can achieve the highest accuracy and is the effective model that can be used to distinguish between authentic and fake users.

The reasons why Random Forest model stood out are as follows:

 Ensemble Method: It is a member of the ensemble method family, which builds on the strengths and accuracy of individual models to produce a more robust model. By combining the predictions from

- several decision trees, Random Forest lowers the chance of overfitting and boosts accuracy.
- Reduction of Overfitting: The Random Forest method uses bootstrapping and feature randomness to train each decision tree on a distinct subset of data and features. Because of the diversity and randomness among the trees, overfitting is less common, and the model performs better on unknown data.
- 3. Managing Non-linear Relationships: Non-linear relationships between features and the target variable can be captured by Random Forests. The system is capable of managing intricate relationships found in the data by utilizing multiple trees and taking into account distinct feature subsets for every tree.

12.Conculsion

At the end of the research, we can say that to ensure authenticity of users on Instagram one should take text and image analysis into consideration, the texts users use in their captions and comments for posts might not be relevant at times which may help in fake user identification. Image analysis and interval between posts analysis is also very useful in identifying spammers. Hence to make Instagram a better platform the above steps can be taken into consideration.

REFERENCES

- [1] Kristo Radion Purba, David Asirvatham, Raja Kumar Murugesan, "Classification of Instagram Fake Users Using Supervised Machine Learning Algorithms", International Journal of Electrical and Computer Engineering (IJECE)., Vol. 10, No. 3, Juni 2020, pp. 2763~2772, ISSN: 2088-8708, DOI: 10.11591/ijece.v10i3.pp2763-2772
- [2] K. Harish, R.Naveen Kumar, Dr. J. Briso Becky Bell, "Fake Profile Detection Using Machine Learning", International Journal of Scientific Research in Science, Engineering and Technology, Print ISSN: 2395-1990|Online ISSN:2394-4099 (www.ijsrset.com) doi: https://doi.org/10.32628/IJSRSET2310264
- [3] Partha Chakraborty*, Mahim Musharof Shazan, Mahamudul Nahid, Md. Kaysar Ahmed, Prince Chandra Talukder, "Fake Profile Detection Using Machine Learning Techniques", Journal of Computer and Communications > Vol.10 No.10, October 2022, DOI: 10.4236/jcc.2022.1010006