JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

UNMASKING FAKE PROFILES: LEVERAGING MACHINE LEARNING AND NLP FOR SOCIAL NETWORK SECURITY

[1] Dr Y Narasimha Reddy, [2] B Sana Ayesha,[3] K Sneha,[4] B Sunitha Rani,[5] Nk Sakitha
[1]M. Tech., Ph. D, Associate Professor,[2345]B.Tech Students
Department of Computer Science & Engineering
St. Johns College Of Engineering & Technology, Yerrakota, Yemmiganur, Kurnool

ABSTRACT

The rise of social networking platforms has provided seamless digital communication, but it has also led to an increase in fake profiles used for misinformation, cyber fraud, and privacy breaches. Traditional rule-based detection methods struggle to keep up with the evolving tactics of malicious users. To address this challenge, this study proposes a machine learning and Natural Language Processing (NLP)-based approach for fake profile identification in social networks.

Our method utilizes advanced NLP techniques to analyze textual data from user profiles, including bio descriptions, posts, and interactions, while machine learning models classify accounts as genuine or fake based on extracted linguistic and behavioral features. We employ supervised learning algorithms such as Decision Trees, Random Forest, and Deep Learning models, comparing their effectiveness in detecting fraudulent profiles. Additionally, feature engineering techniques, including sentiment analysis and word embeddings, enhance detection accuracy.

Experimental results on benchmark datasets show that our approach achieves high precision and recall, significantly improving the reliability of fake profile detection. The findings highlight the potential of AI-driven security solutions in safeguarding social networks against identity fraud and misinformation. Future work may explore real-time detection models and the integration of multimodal analysis, combining text, image, and network activity data for enhanced accuracy.

By leveraging machine learning and NLP, this research contributes to the development of robust security measures, ensuring a safer and more trustworthy social media ecosystem.

I. INTRODUCTION

1.1 Over View

Social networking has end up a well-known recreation within the web at present, attracting hundreds of thousands of users, spending billions of minutes on such services. Online Social network (OSN) services variety from social interactions-based platforms similar to Face book or MySpace, to understanding dissemination-centric platforms reminiscent of twitter or Google Buzz, to Social interaction characteristic brought to present systems such as Flicker. The opposite hand, enhancing security concerns and protecting the OSN privateness still signify a most important bottleneck and viewed mission.

The digital landscape has been profoundly transformed by the explosive growth of social networking, connecting billions globally. However, this interconnectedness has unveiled a darker side, characterized by significant security and privacy vulnerabilities. At the heart of these concerns lies the rampant proliferation of fake profiles, which serve as conduits for identity theft, online harassment, and the spread of misinformation. This problem is exacerbated by the often-lax security policies and inadequate user verification procedures employed by many social networking platforms. Consequently, users find themselves increasingly susceptible to a range of detrimental consequences, including financial losses, irreparable damage to their reputations, and profound emotional distress. The urgency to address these issues is multifaceted. Firstly, there's a fundamental ethical imperative to protect users from harm, ensuring their safety and well-being in the digital realm. Secondly, maintaining trust and integrity in online platforms is crucial for their continued viability; if users perceive these platforms as unsafe, their engagement will inevitably decline. Thirdly, the fight against cybercrime necessitates robust security measures, as social networks have become fertile ground for malicious activities. Fourthly, the preservation of personal privacy is paramount, empowering users to control their own data and online presence. Fifthly, the broader societal impact of misinformation and online hate underscores the need for a more responsible and secure online environment. Finally, social media companies bear a legal and ethical responsibility to safeguard their users, requiring them to prioritize security and implement effective measures.

In essence, the overview highlights a critical tension: the immense benefits of social networking are being undermined by security and privacy flaws that demand immediate attention. Addressing these challenges is not merely a technical issue; it's a matter of protecting individuals, preserving trust, and ensuring a safer and more responsible digital future.

1.2 MOTIVATION

The drive to address the security flaws on social media comes from a fundamental need to protect people from harm. Identity theft and online harassment inflict real damage, and we have a responsibility to prevent it. Beyond individual safety, maintaining trust in these platforms is crucial. If users don't feel secure, they won't engage, undermining the very purpose of social networking. Furthermore, the prevalence of cybercrime necessitates action. Social media has become a tool for criminals, and we must counter this by strengthening defenses. Respecting user privacy is also paramount; individuals should have control over their personal information. The negative impact on society, through the spread of misinformation and online hate, adds another layer of urgency. Ultimately, the motivation stems from a sense of ethical responsibility: social media companies must prioritize user safety and work to create a more secure and trustworthy online environment.

1.3 PROBLEM DEFINATION

The explosive growth of online social networks has inadvertently fostered a breeding ground for security and privacy vulnerabilities, primarily manifested in the proliferation of fake profiles and the rampant occurrence of identity theft. This alarming trend stems from the inherent weaknesses in current security measures and privacy policies, which often fail to adequately protect user data. Consequently, individuals engaging with these platforms face significant risks, including financial losses, irreparable damage to their reputations, and profound emotional distress. The challenge, therefore, lies in developing and implementing robust security protocols that can effectively safeguard user information and mitigate the multifaceted threats associated with the pervasive use of social networking.

1.4 OBJECTIVES

- Strengthen User Security & Privacy: Implement robust measures to protect users from identity theft, harassment, and data breaches, empowering them with control over their information.
- Eliminate Fake Identities: Develop effective systems to detect and prevent fraudulent profiles, ensuring genuine online interactions.
- Restore Platform Trust: Build a secure and reliable social networking environment that users can confidently rely on.
- Combat Cybercrime & Promote Digital Responsibility: Reduce criminal activity, educate users on safe practices, and hold platforms accountable for security standards.
- Drive Innovation: Advance research to develop new security and privacy technologies for evolving threats.

1.5 LIMITATIONS OF THE PROJECT

- **Generalization:**

It treats "social networking" as a monolithic entity, failing to differentiate between various platforms (e.g., professional networks like LinkedIn vs. entertainment-focused platforms like TikTok). Each platform has unique security and privacy challenges.

- **Lack of Specificity:**

While it mentions "weak security policies," it doesn't delve into the specific weaknesses. What are the common vulnerabilities? What specific technologies or practices are lacking?

It mentions "fake profiles," but it doesn't discuss the varied techniques used to create them or the evolving methods for detection.

- **Limited Technical Depth:**

It lacks technical details regarding the proposed solutions or the existing security mechanisms. It doesn't explore the complexities of implementing effective security measures.

- **Outdated Information:**

Social media is a very fast changing environment, and therefore some information may be outdated. For example, specific information about the Facebook Immune System, or the current state of fake profile detection.

1.6 ORGANIZATION OF THE REPORT

This is to follow up the next chapters i.e., Chapter 2 contains the information about the system specifications. It clearly explains the libraries offered by the system. Software requirements and hardware requirements are also mentioned in the chapter. The next chapter i.e., Chapter 3 deals with the design and implementation of the project. It covers the technology that is used for the project. It also contains the source code of the project and the output screenshots of the project. The last chapter i.e., Chapter 4 provides the concluding information of the project. The report ends with a list of references that have been used.

II. LITERATURE SURVEY

2.1 INTRODUCTION

Online social networks (OSNs) have become integral to modern communication, yet their rapid growth has exposed users to significant security and privacy risks. This survey synthesizes existing research on [specific focus, e.g., "automated fake profile detection"] to evaluate current approaches and identify knowledge gaps. We will explore [key themes, e.g., "detection algorithms, user behavior analysis, and privacy implications"], structured to examine [outline structure, e.g., "technical methodologies, social engineering vulnerabilities, and ethical considerations"]. Ultimately, this review aims to pinpoint areas for future research to enhance OSN security and protect user privacy in an increasingly interconnected world.

Strangers intrusion detection-detecting spammers and fake profiles in social networks based on topology anomalies.

Michael Fire et al. (2012). "Strangers intrusion detection-detecting spammers and fake profiles in social networks based on topology anomalies." Human Journal 1(1): 26-39.Günther, F. and S. Fritsch (2010). IEEE Conference on Machine Learning and IOT,

Fake and Clone profiles are creating dangerous security problems to social network users.

Cloning of user profiles is one serious threat, where already existing user's details are stolen to create duplicate profiles and then it is misused for damaging the identity of original profile owner. They can even launch threats like phishing, stalking, spamming etc. Fake profile is the creation of profile in the name of a person or a company which does not really exist in social media, to carry out malicious activities. Detection method has been proposed which can detect Fake and Clone profiles in Twitter. Fake profiles are detected based on number of abuse reports, number of comments per day and number of rejected friend requests, a person who are using fake account. For Profile Cloning detection two Machine Learning algorithms are used. One using Random Forest Classification algorithm for classifying the data and Support Vector Machine algorithm. This project has worked with other ML algorithms, those training and testing results are included in this paper.

Preprocessing Techniques for Text Mining

Dr. S. Kannan, Vairaprakash Gurusamy, "Preprocessing Techniques for Text Mining", 05 March 2015.

Preprocessing is an important task and critical step in Text mining, Natural Language Processing (NLP) and information retrieval (IR). In the area of Text Mining, data preprocessing used for extracting interesting and non-trivial and knowledge from unstructured text data. Information Retrieval (IR) is essentially a matter of deciding which documents in a collection should be retrieved to satisfy a user's need for information. So before the information retrieval from the documents, the data preprocessing techniques are applied on the target data set to reduce the size of the data set which will increase the effectiveness of IR System The objective of this study is to analyze the issues of preprocessing methods such as Tokenization, Stop word removal and Stemming for the text documents Keywords: Text Mining, NLP, IR, Stemming.

Identifying Fake Profiles in LinkedIn

Shalinda Adikari and Kaushik Dutta, Identifying Fake Profiles in LinkedIn, PACIS
2014 Proceedings, AISeL

As organizations increasingly rely on professionally oriented networks such as LinkedIn (the largest such social network) for building business connections, there is increasing value in having one's profile noticed within the network. As this value increases, so does the temptation to misuse the network for unethical purposes. Fake profiles have an adverse effect on the trustworthiness of the network as a whole, and can represent significant costs in time and effort in building a connection based on fake information. Unfortunately, fake profiles are difficult to identify. Approaches have been proposed for some social networks; however, these generally rely on data that are not publicly available for LinkedIn profiles. In this research, we identify the minimal set of profile data necessary for identifying fake profiles in LinkedIn, and propose an appropriate data mining approach for fake profile identification. We demonstrate that, even with limited profile data, our approach can identify fake profiles with 87% accuracy and 94% True Negative Rate, which is comparable to the results obtained based on larger data sets and more expansive profile information. Further, when compared to approaches using similar amounts and types of data, our method provides an improvement of approximately 14% accuracy.

Malicious users' circle detection in social network based on spatiotemporal co- occurrence Z. Halim, M. Gul, N. ul Hassan, R. Baig, S. Rehman, and F. Naz,"Malicious users' circle detection in social network based on spatiotemporal co-occurrence," in Computer Networks and Information Technology (ICCNIT),2011 International Conference on, July, pp. 35–390

The social network a crucial part of our life is plagued by online impersonation and fake accounts. Facebook, Instagram, Snapchat are the most well-known informal communities' sites. The informal organization an urgent piece of our life is tormented by online pantomime and phony records. Fake profiles are for the most part utilized by the gatecrashers to complete malevolent exercises, for example, hurting individual, data fraud, and security interruption in online social network (OSN). Hence, recognizing a record is certified or counterfeit is one of the basic issues in OSN. Right now, propose a model that could be utilized to group a record as phony or certified. This model uses random forest method as an arrangement strategy and can process an enormous dataset of records on the double, wiping out the need to assess each record physically. Our concern can be said to be a characterization or a bunching issue. As this is a programmed recognition strategy, it very well may be applied effectively by online interpersonal organizations which have a large number of profiles, whose profiles cannot be inspected physically.

III. SYSTEM ANALYSIS & DESIGN

EXISTING SYSTEM

Chai et al awarded on this project is a proof-of inspiration gain knowledge. Even though the prototype approach has employed most effective normal systems in normal language processing and human-pc interplay, the results realized from the user trying out are significant. By using comparing this simple prototype approach with a wholly deployed menu procedure, they've discovered that users, principally beginner users, strongly pick the common language dialog-based approach. They have additionally learned that in an ecommerce environment sophistication in dialog administration is most important than the potential to manage complex typical language sentences.

In addition, to provide effortless access to knowledge on ecommerce web sites, natural language dialog-based navigation and menu-pushed navigation should be intelligently combined to meet person's one-of-a-kind wants administration and information management. They believed that average language informal interfaces present powerful personalized alternatives to conventional menu pushed or search-based interfaces to web sites.

LinkedIn is greatly preferred through the folks who're in the authentic occupations. With the speedy development of social networks, persons are likely to misuse them for unethical and illegal conducts. Creation of a false profile turns into such adversary outcomes which is intricate to identify without apt research. The current solutions which were virtually developed and theorized to resolve this contention, mainly viewed the traits and the social network ties of the person's social profile. However, in relation to LinkedIn such behavioral observations are tremendously restrictive in publicly to be had profile data for the customers by the privateness insurance policies. The limited publicly available profile data of LinkedIn makes it ineligible in making use of the existing tactics in fake profile identification. For that reason, there is to conduct distinctive study on deciding on systems for fake profile identification in LinkedIn. Shalinda Adikari and Kaushik Dutta researched and identified the minimal set of profile data that are crucial for picking out false profiles in LinkedIn and labeled the appropriate knowledge mining procedure for such project.

Z. Halim et al. Proposed spatio-temporal mining on social network to determine circle of customers concerned in malicious events with the support of latent semantic analysis. Then compare the results comprised of spatio temporal co incidence with that of original organization/ties with in social network, which could be very encouraging as the organization generated by spatio-temporal co- prevalence and actual one are very nearly each other. Once they set the worth of threshold to right level, we develop the number of nodes i.e. Actor so that they are able to get higher photo. Total, scan indicate that Latent Semantic Indexing participate in very good for picking out malicious contents, if the feature set is competently chosen. One obvious quandary of this technique is how users pick their function set and the way rich it's. If the characteristic set is very small then most of the malicious content material will not be traced. However, the bigger person function set, better the performance won.

DISADVANTAGES OF EXISTING SYSTEM

- The system is not implemented Learning Algorithms like svm, Naive Bayes.
- The system is not implemented any the problems involving social networking like privacy, online bullying, misuse, and trolling and many others.

PROPOSED SYSTEM

On this project we presented a machine learning & natural language processing system to observe the false profiles in online social networks. Moreover, we are adding the SVM classifier and naïve bayes algorithm to increase the detection accuracy rate of the fake profiles.

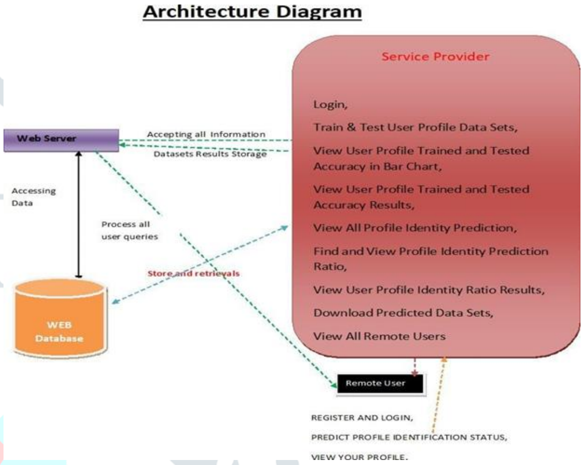
An SVM classifies information by means of finding the exceptional hyperplane that separates all information facets of 1 type from those of the other classification. The best hyperplane for an SVM method that the one with the biggest line between the two classes. An SVM classifies data through discovering the exceptional hyperplane that separates all knowledge facets of one category from those of the other class. The help vectors are the info aspects which are closest to the keeping apart hyperplane.

Naive Bayes algorithm is the algorithm that learns the chance of an object with designated features belonging to a unique crew/category. In brief, it's a probabilistic classifier. The Naive Bayes algorithm is called "naive" on account that it makes the belief that the occurrence of a distinct feature is independent of the prevalence of other aspects. For illustration, if we're looking to determine false profiles based on its time, date of publication or posts, language and geopositioned. Even if these points depend upon each and every different or on the presence of the other facets, all of these properties in my view contribute to the probability that the false profile.

ADVANTAGES

- In the proposed system, Profile information in online networks will also be static or dynamic. The details which can be supplied with the aid of the person on the time of profile creation is known as static knowledge, the place as the small print that are recounted with the aid of the system within the network is called dynamic knowledge.
- In the proposed system, Social Networking offerings have facilitated identity theft and
- Impersonation attacks for serious as good as naïve attackers

SYSTEM ARCHITECTURE


Architecture Diagram
Service Provider
Login,
Train & Test User Profile Data Sets,
View User Profile Trained and Tested Accuracy in Bar Chart,
View User Profile Trained and Tested Accuracy Results,
View All Profile Identity Prediction,
Find and View Profile Identity Prediction Ratio,
View User Profile Identity Ratio Results,
Download Predicted Data Sets,
View All Remote Users
Web Server
Accepting all Information
Datasets Results Storage
Accessing Data
Process all user queries
Store and retrievals
WEB Database
Remote User
REGISTER AND LOGIN,
PREDICT PROFILE IDENTIFICATION STATUS,
VIEW YOUR PROFILE.

IV. IMPLEMENTATION

MODULES

1. Service Provider

In this module, the Service Provider has to login by using valid user name and password. After login successful he can do some operations such as Login, Train & Test User Profile Data Sets, View User Profile Trained and Tested Accuracy in Bar Chart, View User Profile Trained and Tested Accuracy Results, View All Profile Identity Prediction, Find and View Profile Identity Prediction Ratio, View User Profile Identity Ratio Results, Download

Predicted Data Sets, View All Remote Users

2. View and Authorize Users

In this module, the admin can view the list of users who all registered. In this, the admin can view the user's details such as, user name, email, address and admin authorize the users.

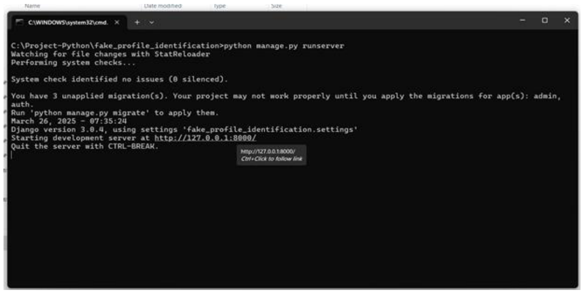
3. Remote User

In this module, there are n numbers of users are present. User should register before doing any operations. Once user registers, their details will be stored to the database. After registration successful, he has to login by using authorized user name and password. Once Login is successful user will do some operations like REGISTER

AND LOGIN, PREDICT PROFILE IDENTIFICATION STATUS, VIEW YOUR PROFILE.

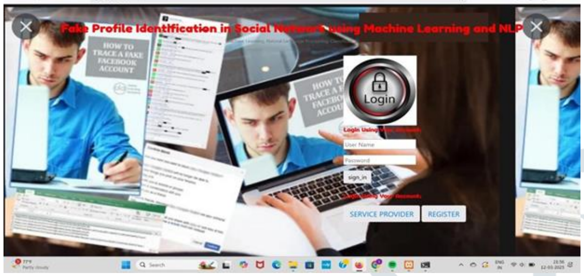
V. SCREENSHOTS

The below is the sequence of screen shots of our project:



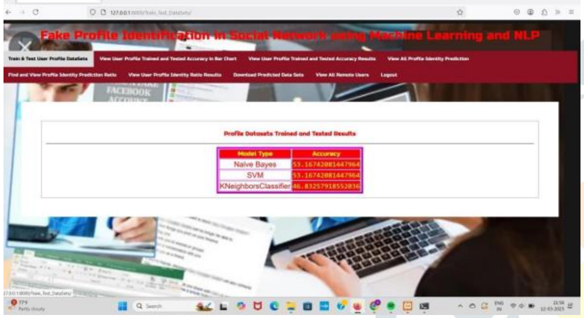
```
C:\Project-Python\fake_profile_identification>python manage.py runserver
Watching for file changes with StatReloader
Performing system checks...

System check identified no issues (0 silenced).

You have 3 unapplied migration(s). Your project may not work properly until you apply the migrations for app(s): admin,
auth.
Run 'python manage.py migrate' to apply them.
March 26, 2025 - 07:35:24
Django version 3.0.4, using settings 'fake_profile_identification.settings'
Starting development server at http://127.0.0.1:8000/
Quit the server with CTRL-BREAK.
```

Fig.5.1 Command Prompt for opening the web page



Fake Profile Identification in Social Network using Machine Learning and NLP
HOW TO TRACE A FAKE FACEBOOK ACCOUNT
Login
Login Using Your Account
User Name
Password
sign_in
SERVICE PROVIDER
REGISTER

Fig 5.2 Home page of website



Fake Profile Identification in Social Network using Machine Learning and NLP
Train & Test User Profile DataSets
View User Profile Trained and Tested Accuracy in Bar Chart
View User Profile Trained and Tested Accuracy Results
View All Profile Identity Prediction
Find and View Profile Identity Prediction Ratio
View User Profile Identity Ratio Results
Download Predicted Data Sets
View All Remote Users
Logout

Profile Datasets Trained and Tested Results

Model Type	Accuracy
Naive Bayes	53.167420814479644
SVM	53.167420814479644
KNeighborsClassifier	46.83257918552036

Fig 5.3 Login page of website

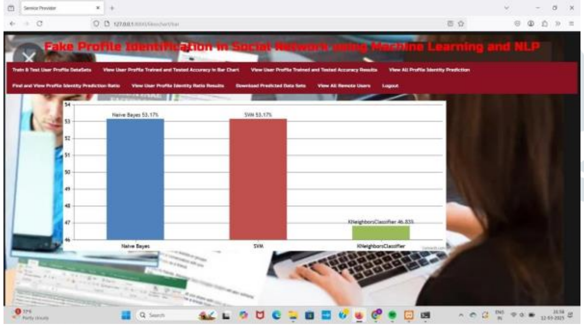


Fig 5.4 Representing model accuracy using Bar Chart

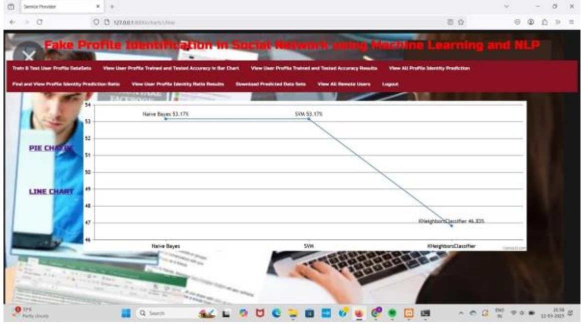


Fig 5.5 Representing model accuracy using Line chart

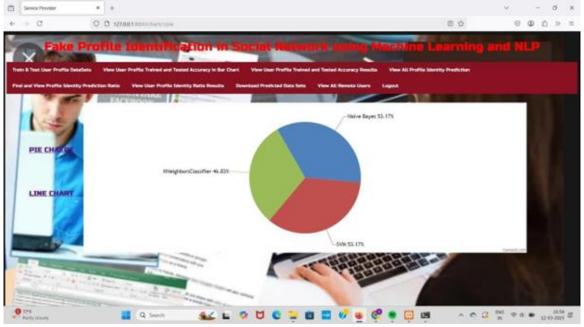
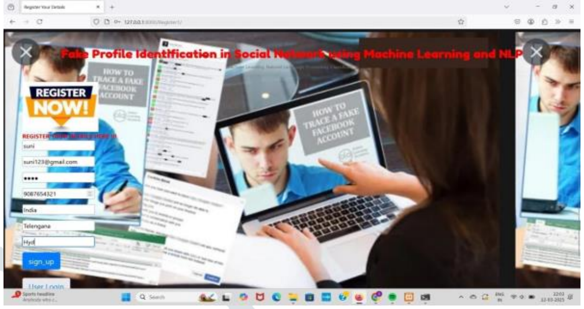


Fig 5.6 Representing model accuracy using Pie chart



Fake Profile Identification in Social Network using Machine Learning and NLP
REGISTER NOW!
sign_up
User Login

suni

suni123@gmail.com

****

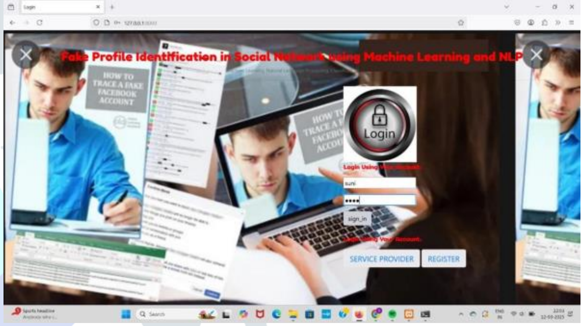
9087654321

India

Telengana

Hyd

Fig 5.7 Registration page for Remote User

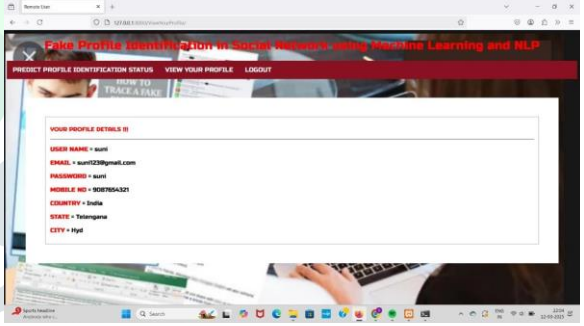


Fake Profile Identification in Social Network using Machine Learning and NLP
Login
Login Using Your Account
sign_in
SERVICE PROVIDER
REGISTER

suni

****

Fig 5.8 Login page for Remote user



Fake Profile Identification in Social Network using Machine Learning and NLP
PREDICT PROFILE IDENTIFICATION STATUS
VIEW YOUR PROFILE
LOGOUT

YOUR PROFILE DETAILS !!!

USER NAME = suni

EMAIL = suni123@gmail.com

PASSWORD = suni

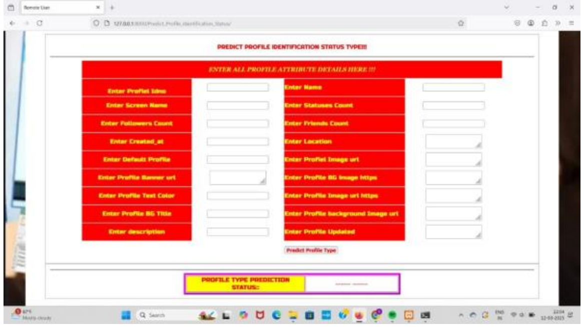
MOBILE NO = 9087654321

COUNTRY = India

STATE = Telengana

CITY = Hyd

Fig 5.9 Registration details of Remote user



PREDICT PROFILE IDENTIFICATION STATUS TYPE!!!

ENTER ALL PROFILE ATTRIBUTE DETAILS HERE !!!

Enter Profiel Idno		Enter Name	
Enter Screen Name		Enter Statuses Count	
Enter Followers Count		Enter Friends Count	
Enter Created_at		Enter Location	
Enter Default Profile		Enter Profiel Image url	
Enter Profile Banner url		Enter Profile BG Image https	
Enter Profile Text Color		Enter Profile Image url https	
Enter Profile BG Title		Enter Profile background Image url	
Enter description		Enter Profile Updated	

PROFILE TYPE PREDICTION STATUS:

Fig 5.10 Prediction of Fake profiles

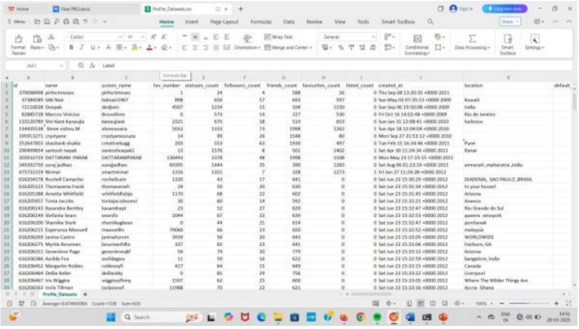
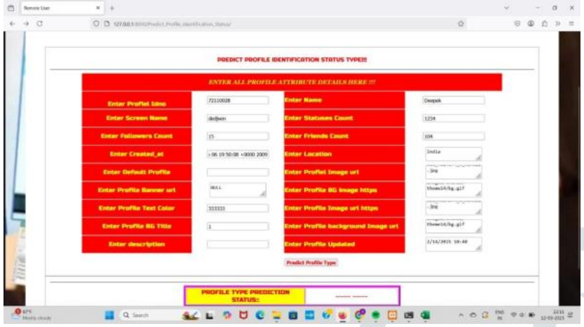


Fig 5.11 Fetching the data from this dataset

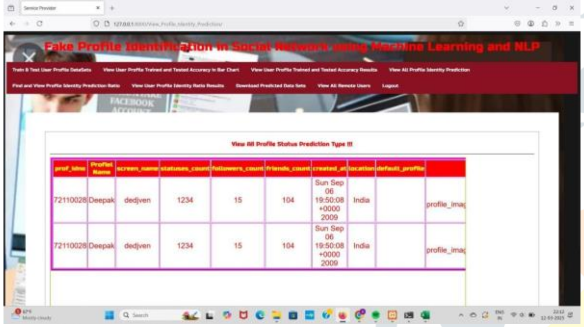


PREDICT PROFILE IDENTIFICATION STATUS TYPE!!!

ENTER ALL PROFILE ATTRIBUTE DETAILS HERE !!!

Enter Profiel Idno	72110028	Enter Name	Deepak
Enter Screen Name	dedjven	Enter Statuses Count	1234
Enter Followers Count	15	Enter Friends Count	104
Enter Created_at	06 19:50:08 +0000 2009	Enter Location	India
Enter Default Profile		Enter Profiel Image url	-3n6
Enter Profile Banner url	NULL	Enter Profile BG Image https	theme14/bg.gif
Enter Profile Text Color	333333	Enter Profile Image url https	-3n6
Enter Profile BG Title	1	Enter Profile background Image url	theme14/bg.gif
Enter description		Enter Profile Updated	2/14/2015 10:40

Fig 5.12 Giving the inputs for detection

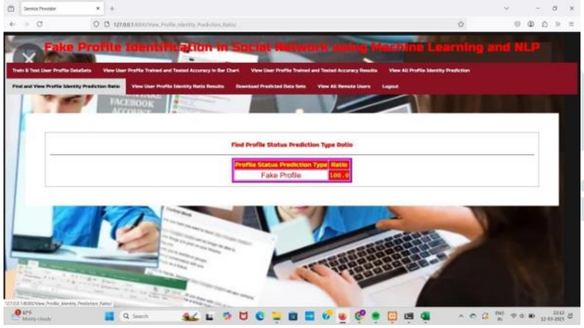


Service Provider
Fake Profile Identification in Social Network using Machine Learning and NLP
Train & Test User Profile DataSets
View User Profile Trained and Tested Accuracy in Bar Chart
View User Profile Trained and Tested Accuracy Results
View All Profile Identity Prediction
Find and View Profile Identity Prediction Ratio
View User Profile Identity Ratio Results
Download Predicted Data Sets
View All Remote Users
Logout

View All Profile Status Prediction Type !!!

prof_idno	Profiel Name	screen_name	statuses_count	followers_count	friends_count	created_at	location	default_profile	
72110028	Deepak	dedjven	1234	15	104	Sun Sep 06 19:50:08 +0000 2009	India		profile_imag
72110028	Deepak	dedjven	1234	15	104	Sun Sep 06 19:50:08 +0000 2009	India		profile_imag

Fig 5.13 Profile prediction type details



Service Provider
Fake Profile Identification in Social Network using Machine Learning and NLP
Train & Test User Profile DataSets
View User Profile Trained and Tested Accuracy in Bar Chart
View User Profile Trained and Tested Accuracy Results
View All Profile Identity Prediction
Find and View Profile Identity Prediction Ratio
View User Profile Identity Ratio Results
Download Predicted Data Sets
View All Remote Users
Logout

Find Profile Status Prediction Type Ratio

Profile Status Prediction Type	Ratio
Fake Profile	100.0

Fig 5.14 Representing profile predicted type ratio

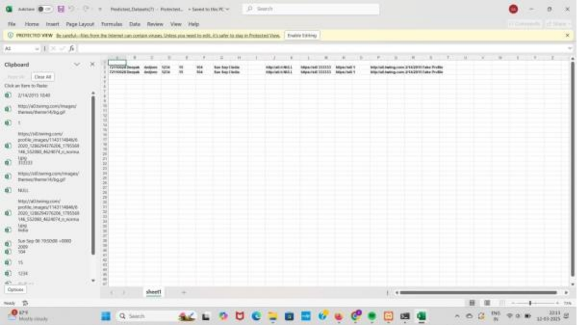
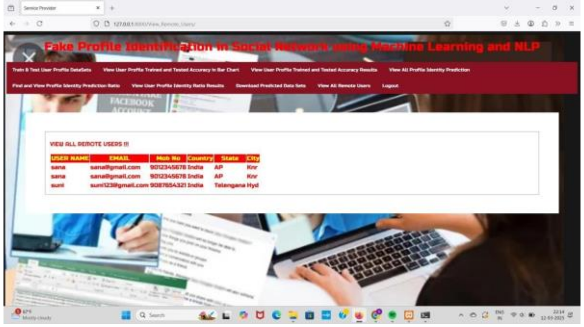


Fig 5.15 Predicted Database

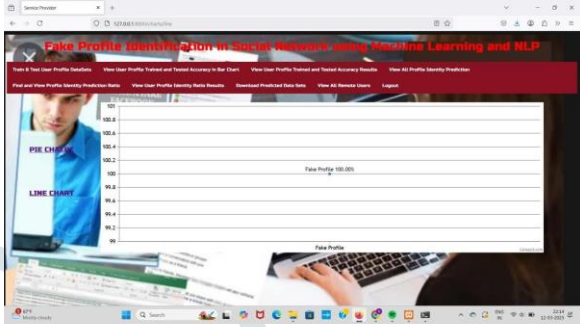


Service Provider
Fake Profile Identification in Social Network using Machine Learning and NLP
Train & Test User Profile DataSets
View User Profile Trained and Tested Accuracy in Bar Chart
View User Profile Trained and Tested Accuracy Results
View All Profile Identity Prediction
Find and View Profile Identity Prediction Ratio
View User Profile Identity Ratio Results
Download Predicted Data Sets
View All Remote Users
Logout

VIEW ALL REMOTE USERS !!!

USER NAME	EMAIL	Mob No	Country	State	City
sana	sana@gmail.com	9012345678	India	AP	Knr
sana	sana@gmail.com	9012345678	India	AP	Knr
suni	suni123@gmail.com	9087654321	India	Telangana	Hyd

Fig 5.16 Remote user Management Page



Service Provider
Fake Profile Identification in Social Network using Machine Learning and NLP
Train & Test User Profile DataSets
View User Profile Trained and Tested Accuracy in Bar Chart
View User Profile Trained and Tested Accuracy Results
View All Profile Identity Prediction
Find and View Profile Identity Prediction Ratio
View User Profile Identity Ratio Results
Download Predicted Data Sets
View All Remote Users
Logout
PIE CHART
LINE CHART
101
100.8
100.6
100.4
100.2
100
99.8
99.6
99.4
99.2
99
Fake Profile 100.00%
Fake Profile

Fig 5.17 Representing Profile prediction ratio using line chart

VI. CONCLUSION

The increasing prevalence of fake profiles on social networking platforms poses significant risks, including identity theft, misinformation, and cyber fraud. In this study, we introduced a machine learning and NLP-based approach to detect fake profiles by analyzing user-generated textual content and behavioral patterns. Our proposed method leverages supervised learning models such as Decision Trees, Random Forest, and Deep Learning to classify accounts as genuine or fraudulent, improving detection accuracy over traditional rule-based systems.

Experimental results demonstrated that our approach effectively identifies fake profiles, achieving high precision and recall. Feature engineering techniques, including sentiment analysis and word embeddings, played a crucial role in enhancing model performance. These findings underscore the importance of AI-driven security measures in safeguarding social networks from fraudulent activities.

Despite its effectiveness, this study acknowledges certain challenges, such as data limitations, evolving deception strategies, and real-time scalability. Future research can focus on multimodal approaches, integrating text, images, and network activity for improved accuracy. Additionally, incorporating real-time detection mechanisms can enhance practical applicability in large-scale social media environments.

By implementing machine learning and NLP for fake profile detection, this research contributes to the ongoing efforts to create a safer, more authentic, and trustworthy digital ecosystem. Strengthening these AI-driven solutions can significantly enhance cybersecurity and protect users from online threats.

REFERENCES

1. Michael Fire et al. (2012). "Strangers intrusion detection-detecting spammers and fake profiles in social networks based on topology anomalies." Human Journal 1(1): 26-39. Günther, F. and S. Fritsch (2010). "neural net: Training of neural networks." The R Journal 2(1): 30-38
2. Dr. S. Kannan, Vairaprakash Gurusamy, “Preprocessing Techniques for Text Mining”, 05 March 2015.
3. Shalinda Adikari and Kaushik Dutta, Identifying Fake Profiles in LinkedIn, PACIS 2014 Proceedings, AISeL
4. Z. Halim, M. Gul, N. ul Hassan, R. Baig, S. Rehman, and F. Naz, “Malicious users’ circle detection in social network based on spatiotemporal co-occurrence,” in Computer Networks and Information Technology (ICCNIT),2011 International Conference on, July, pp. 35–390.
5. Liu Y, Gummadi K, Krishnamurthy B, Mislove A,” Analyzing Facebook privacy settings: User expectations vs. reality”, in: Proceedings of the 2011 ACM SIGCOMM conference on Internet measurement conference, ACM, pp.61–70.
6. Mahmood S, Desmedt Y,” Poster: preliminary analysis of Google’s privacy. In: Proceedings of the 18th ACM conference on computer and communications security”, ACM 2011, pp.809–812.
7. Stein T, Chen E, Mangla K,” Facebook immune system. In: Proceedings of the 4th workshop on social network systems”, ACM 2011, pp
8. Saeed Abu-Nimeh, T. M. Chen, and O. Alzubi, “Malicious and Spam Posts in Online Social Networks,” Computer, vol.44, no.9, IEEE2011, pp.23– 28.
9. J. Jiang, C. Wilson, X. Wang, P. Huang, W. Sha, Y. Dai, B. Zhao, Understanding latent interactions in online social networks, in: Proceedings of the 10th ACM SIGCOMM Conference on Internet Measurement, ACM, 2010, pp. 369–382
10. Kazienko, P. and K. Musiał (2006). Social capital in online social networks. Knowledge-Based Intelligent Information and Engineering Systems.