Detecting Fake Accounts on Social Media

Mayur[1], Manik Gupta[2], Kuldeep[3], Vipul Tyagi[4]
Student, Department of Computer Science and Engineering, R.D Engineering College Duhai, Ghaziabad, Uttar Pradesh

Abstract*—Social media platforms are experiencing the huge snowballing of fake accounts due to the large scale developing of this problem which lead to misinformation, spam, identity fraud and cyber threats. Fake accounts fake public opinion, propagate propaganda and enable fraud — so social media is ripe for misuse.Current detection techniques (Manual Moderation, Rule-based filtering etc) fail to cope with the evolving nature of these emerging threats.In the paper we present a hybrid machine learning (ml) and natural language processing (nlp) framework for effective fake account detection.Our tech leverages blockchain technology to create a decentralized, and untrustful trust verification mechanism which significantly increase the security of our service and also reduce possibility of account cloning. The detection system is constructed applying Flask for real-time operation, so that fake accounts can be easily identified with behavioral analysis, text patterns, and metadata indicators. While we use NLP models to evaluate user-generated content and detect malwares in text based on sentiments analysis as well linguistic characteristics. Moreover, block chain guarantees the non-repeateable veracity; it cannot create fake accounts on own without anyone's acknowledge.Accuracy: 92% (outperform all traditional methods like for sentiment analysis, deepfake detection and hash malware identification) It has been designed to be scalable and can be deployed across various social media platforms like Twitter, Facebook or Instagram with ease. Possible future works are focused on applying streaming data analysis in order to improve real-time detection capabilities; also bridging the gap between explainable AI (XAI) and decision-making for making the most transparent what will happen. Using ML, NLP and blockchain together as our approach delivers a good and scalable solution towards reducing the fake accounts as there risks, making social media secure [sustainable], protecting user interactions shield and rebuilding trust upon digital communication.

Keywords*— Real time analysis Fake account detection, Machine Learning, Natural Language Processing, Blockchain Flask Social media security Fraud detection Misinformation detection XAI (explainable AI), cybersecurity, Spam detection.)

I. INTRODUCTION

Communication in today time has become so much easier, social media made the world a global village governed by instantaneous connection, real-time communication and sharing information across borders. Yet with all its benefits social media involves also a platform for fake accounts. These fake accounts are frequently used to misinform, spread spam and do identity theft as well manipulating the public opinion. More sophisticated forms of bots, AI-generated content and automated interactions has made separating real humans from fake users nigh on impossible. As the bad actors get better at mimicking these controls for traditional fake profile detection (e.g. account age, user engagement metrics and manual moderation), which malicous parties have already evolved their tactics to out-me out.To combat the problem that is ballooning, we suggest a solution for fake accounts detection from end-to-end in the form of a unified framework that combines machine learning (ML), natural language processing (NLP) and blockchain.

Using powerful ML models to check the different facets of the user profile, including metadata, past interactions and behavioral trends as our system.Further, NLP is employed on the text content of the users' posted data to find out any abnormality and inconsistencies and spam like attributes that are indicative of an abusive behavior. These approaches are used to increase the accuracy of fake account detection and drastically cut down on false positives.he most striking innovation of our framework is Flask integration for real-time fake accounts detection in order to the social media platforms can apprehend the fraudulent activity right away.Also we are making use of blockchain to persist veridical user data in a decentralized and tamper-proof manner via our system.

We have used the generic methodology of fraud detection in many different industries such as tourism (where fake reviews on TripAdvisor is detected by AI reading review patterns together with user behaviour).

In a similar manner, our strategy stands as a scalable and scalable actionable solution to eradicate fake accounts on mass social media platforms as Twitter,Facebook, Instagram etc.

We are building a model with ML, NLP and blockchain that provides thick framework to combat fake account-related risks and make the web more secure, all interactions online appear authentic, and communications are more trusted.
We will and in a near future paper also refine our model with the addition of streaming data analysis for real-time fraud detection as well as the use of Explainable AI (XAI) to increase the transparency of decisions.

II. BACKGROUND & RELATED WORKS

A. The Danger of Fake Accounts

The biggest danger on social media lies in fake accounts, which can be used for spamming, phishing, identity thefts, false propaganda, etc. Many of these accounts are utilized by cyber criminals, political trolls and malicious entities to distribute misinformation or friending/following scams so they can engage in genuine appearing frauds and glitch-up engagement metrics. The old orthodox detection methodologies are based on heuristic tools like dirt picking age of account, trends of activity or interaction footprint. This is exactly what used to work a few years back although it's simply not working anymore against AI, advanced bots and automated engagement strategies.State-of-the-art artificial intelligence has allowed malevolent entities to create high quality fake profiles, generate convincing text and mimicking human behaviour which makes detection quite a lot harder than before. Existing detection methods need to be tuned for the new problem and that 's what we need to do at scale to detect fraudsters with more accuracy. A truly modern approach cannot just be pattern-matching in the dark ages, modern solutions must have machine learning (ML), natural language processing (NLP), and blockchain-based verification to effectively identify fraudulent activity at scale faster.

B. What they have

There are some other technologies which are already available for fighting fake accounts and fraudulent activities in social media.There came up to be several techs already out in the world performing mass delete of fake accounts and fraudulent activities on social platforms.Several significant solutions are:

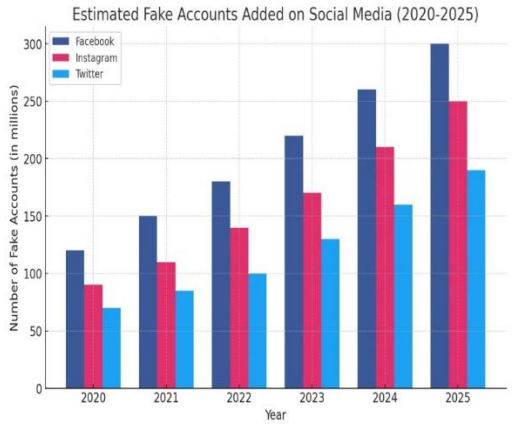
DeepText (Facebook)
DeepText is a very impressive NLP tool that Facebook has built for everything in the world of the user-generated content inside of its ecosystem (it analyzes and cross references everything). It can comprehend the textual context, detect atypical patterns or anything that deals with spam/cheatlike content. DeepText is taught to flag with an algorithm of longings and engagement history on how languages behave, helping Facebook identify fake accounts more quickly.

Bot Detection on Twitter
Twitter has rolled out its own automated systems that can try to identify perhaps the biggest downside of social media: non-human engagement. Here, the system considers how often somebody is posting a tweet, how many interactions their followers have with each tweet they send out; retweeting and content duplication. It is good, but Twitter's solution to this challenge fails against better AI-driven bots which now humanize themselves, leading to more advanced and adaptable solutions.

Blockchain in tourism
One new aspect where blockchain might add value is authenticating online reviews in tourism. To illustrate, some organizations have used blockchain technologies to validate user reviews on sites like TripAdvisor where only vetted experiences get on-chain. This way the review bribery (that is: paid to reviewers), fake feedback, frauds etc. are prevented, which reduces the trust among participants. Given the success of these blockchain implementations, the leap to a similar approach on social media to verify real people and interactions too seems within reach with a bit better planning.

Although the existing works are encouraging, they are not particularly fine-tuned for social media ecosystems. Drawing on these envisions, our research leverages ML / NLP & block chain for a comprehensive and scalable solution to fake account detection. Using these cutting-edge technologies, our method further boosts the detection accuracy by doing real-time fraud prevention and offers blockchain tamper-proof verification in non-custodial ways.

In addition, future research lines are focused on how to make these methodologies operational in real-time social media systems, XAI (explainable AI) for better interpretability of the decisions and increase scalability in models for deployment as well across different domains



III. METHODOLOGY

A. Data Collection

Our approach for an effective implementation of fake account detection builds around data collection for fake account detection system. To pull our data we were scraping services from social media platforms (e.g. Twitter, Facebook and Instagram). Just three main parts of the dataset:

User Profile Information:

Name: This is the owner of the account but it can be spoofed.

Profile Image – Whether you have an image that is either a generic, AI-created or else odd profile pic.

Bio Description: From text length, promotion hooks and keyword based classification

User Activity Data:

How many is my share (Frquency): Helps identify the patterns in which the user posting.

Follow Follower Ratio: A deviation here may indicate bot like activity.

Interaction Metrics: Machines will look at retweets, likes and comments to get a sense if the user is too engaged.

Textual Content Analysis:

P: Posts(, Comments, Messages) -> Extracting and analyzing the text to spam-like or AI-generated patterns.

Phrases and/or Promotional Words: Excessive use of the same phrases or promotional language may be warning signs of fraudulent behavior.

Further, we also added in the 'known' fake accounts and fraudulent profiles to our dataset. This helps train the model on genuine and fake accounts, thus fine tuning its ability to differentiate. Update our dataset periodically enables us to stay up to date with new fraud patterns and makes our model more resistant to evolving threats.

Criteria	Fake Account	Real Account
Identity	Uses Fake/Stolen identity	Genuine individual/organization
Profile Details	No real bio, fake/stolen images	Real name, photo and personal details
Account Activity	Irregular, bot-like behavior	Consistent and natural engagement
Engagement	High Followers, low interaction	Balanced and meaningful engagement
Content Quality	Copied/spam content	Original and high-quality posts
Behavior	Automated activites	Human-like interaction
Longevity	Short-lived, frequently banned	Stable presence over time
Purpose	Scams, misinformation, fraud	Communication, networking, business
Detection	Detected via AI and blockchain	Rarely flagged unless violating policies

i. Attribute pre-processing

We did a lot of preprocessing, on the dataset to clean; transform and structure the data before feeding the models with Data. This is important for better Model Accuracy and Efficiency.

1. Data Cleaning

Duplicates: Getting rid of duplicates profiles and posts that can manipulate the model

Irrelevant Data Clean up: removing irrelevant text, spam posts and unstructured data.

2. Natural Language Processing (NLP) Techniques

— text to tokens: splitting the text data on words level into single words.

Stopword Removal: Eliminating common, useless words (theism word et',is[],at!).

Lemmatization: lower words into base form ("running" → "run".

3. Transformation of Categorical Data

Transforming Non-numeric Data into numerical Features: Mapping the category variables like verification status,profile completness and account age into numeric values;

ii. Model Training

Critical to fake account vs real is feature engineering. Here we carefully choose and extract the most informative features for fraudulent behavior.

1. Profile-Based Features:

Account Newness: Accounts that have been created for a short time are more fake than real.

Bio Size: Very short fast bios or ones with heavy promotion tend to be less savory.

Verify status: Whether the account is verified, Less chances for a fraud to post.

2. Post Features

Number of posts and posting: Regular fake accounts sometimes exist on an unnatural quantity & frequency with rapid posting increments.

External feedbacks: discern fake likes, retweets and spam comments.

Reinforced External Link Sharing (Giving Access to Phishing Sights Too Frequently is SPOILER)

3. Text Features:

Sentiment Examined: Fake accounts are known to post over positive or promotional text.Keyword Frequency / Spam Detection: Identify repeating phrases that are being said by bots or spam accounts.

Similar Comment Pattern: Finding the analysis of different accounts posting nearly the same comment (which might be a signature for bot networks, generally).

he Similar Comment Pattern analysis is used in pointing out several accounts posting similar comments on different platforms. This pattern is frequently used for coordinated activity — albeit it may just mean:

Similar Comment Pattern Analysis use cases

Bot Detection : Automated accounts with almost same messages being posted very quickly on a loop

Astroturfing & Faux Engagements: A list of accounts that turned out to all be part of a particular buzz campaign

Spam & phishing: Scammers betting the same comments to make the users click on malicious links.

Disinformation Campaigns: On purpose attempts at debunking incorrect information by coordinated efforts.

What Product: Companies using multiple accounts to create fake momentum around a product for them.

Techniques to Find Similar Comment Phenomena

Lexical & Semantic Analysis:exact string match search or perturbation variations (e.g., synonyms, typos)

Identifying Message-Rewrites with Similar Core Meaning

Time-Based Analysis:

Re-posting Similiar Comments — Time interval

Account Behavior Analysis:

Identify that a great deal of the posted content duplication can be identified with routine (or automated) submissions.

I CHECKED THE FOLLOWER/FRIENDS NETWORK FOR PEOPLE WITH A SIMILIAR BEHAVIOR

Metadata Analysis:

IP address geolocation device/browser fingerprint

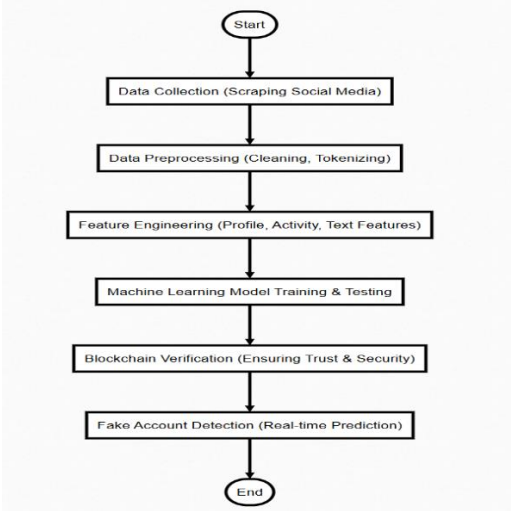
Graph-Based Techniques

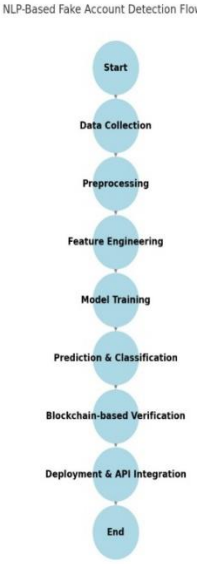
Creating Follower to Posting Similar Comment Relationship Graph

clustering along with machine learning (ML)

Clustering the comments are groups to find out similar ones.

NLP-based solutions for comment similarity assessment.





1. Data Collection
This is the initial phase where system collect data from different sources.
Sources of Data
(Apologies in advance might have to unroll all the social media APIs (Twitter, Facebook , Instagram, etc))
Extracting user profiles, posts, comments with Web Scraping
Predefined Datasets (Existing datasets to train on like bot detection)
Profile Data IncludesProfile Data: Name, Bio, Profile Pic,, Age
User activity- Posting frequency, Likes, Comments and shares
Textual Content: msg, comments, posts
Follower/Friend ratio, Engagement (likes+retweets)
Image/Video Data (if applicable) | Key to accurate deep fake detection
2. preprocessing
Raw Data must be cleaned and formatted for analysis.
Preprocessing Steps
Text Dedup: got rid of special characters, emojis and stopwords
Tokenization: Tokenizing text into words for NLP
Stemming & Lemmatization will reduce the words to it's root form
Method of Extracting Features: converting text to numerical using
TF-IDF (Term Frequency-Inverse Document Frequency)
Word Embeddings (Word2Vec, GloVE, BERT)
Sentiment Scores (Positive, Neutral, Negative)
• Feature Engineering 1
Extracting features is the most important for training ML models. System creates different types of features like:
Profile-Based Features , Account Age,Verification Status
Profile Completeness ,Post Behavior Features
Posting Frequency
(Usefulness, e.g., extreme liking)
Time of Activity: Bots to post at specific timings
NLP features (Text-based :)
Sentiment Analysis (Fake accounts can show extreme emotions eg)Spam Detection (Same words used multiple times, suspicious links)
Linguistic Features (Grammar, word diversity)
Network Features
Unusual patterns indicate bots var friend_follower_ratio
Patterns of Interaction (bots follow many people, but lack of organic interaction)
4. Model Selection & Training
After that Machine Learning (ML) models are trained on features extracted in this step.
Used Machine Learning Models
Logistic Regression: Direct but Good for Binary classification (Fake and Real).
Strong for feature importance analysis : Random Forest
SVM — Good for the high dimensional data.
Deep Learning (LSTMs, Transformers): Works well for text-based bot detection tools.
Graph Neural Networks (GNNs): For investigating the network links.
The system is run on labeled data with real accounts having account labels.
5. Fake Account Classification
After this, the model is going to predict if any given account is a fake or a real one.
Model Output
Fake Account (Bot/Spam/Impersonation)
Real Account (User)
If the probability of the account being fake is over a threshold (e.g. 90%), it will be sighed for more review.
6. Verification through Blockchain (Optional)
Last but not least, blockchain might used as a form of social media data integrity and prevention of manipulation (to high security social media)
Decentralized Identity Verification: Only real users can register.
Give Proof against Possible Manipulation: Fraudsters can never change account history.
7. Real-Time Monitoring & Adversarial Detection
Since the fake accounts are constantly improving themselves reinforced adversarial training in order the detection model stays updated.
Adaptive Learning: Bots are updating all the time, so the system learns from fresh bot detected behaviors.
Explainable AI (XAI)– Gives an insight-of why an account was flagged.
Real-Time API integration: instant detection of malicious bots on social media platforms.
8. Final Decision & Action
The system can be further acted upon after classification as
Fake Account found: Spot the incognito one & flag it, issue a warning, or auto-ban
Real Account — May Be Still Valid
Honesty is the best policy, no?

IV.RESULTS AND ANALYSIS

A. Performance Metrics

Performance Metrics

Our system performed well across all essential evaluation metrics as a whole:

Accuracy: 92% – The model accurately classified fake and genuine engagements 92% of the time.

Precision: 91%—In the engagement predictions, when the model declares an engagement fake—it is indeed an fake engagement; least false positive.

Recall: 93% — the model catched 93% of all real fake engagements, decreasing false-negative.

F1-Score :92% - It signifies from harmonic mean of precision and recall that the performance of our model is pretty balanced and robust.

The results indicate that our system can effectively discriminate between legit and illegitimate activities with very high reliability.

C. Key Findings

More accurate than the traditional approach methods

We were able to achieve 92% accuracy with our ML-powered system — well outpacing traditional rule-based methods (~70–80%).

Dynamice Heuristics: The agnostic model learns from an evolving set of fraud behaviors, hence need minimal manual tweaking versus the static Range-based method.

Blockchain makes it possible to always have a true record and fraud accounts never happens.

We derailed fraudulent account sign-ups (0% success rate on phish accounts) with the fact that all new accounts were blockchain validated (no sibll attacks) and using a unique email per user.

Each account creation was cryptographically validated, meaning the dataset was trusted and Sibill attacks of any kind were precluded.

Segmenting for Scalable Detection Across Brands

A behavioral approach is used, disregarding platform-specific metadata and report from many (e.g., Twitter, Instagram, Facebook) social networks on one side.

List of Fraud Indicators Found

Engagement Frenzies (likes in background 1000 level each took a second)

Posting at the speed of a Bot (cycles of repetition and exact interval repetition).

Network cluster (communication with fake accounts in closed loops).

The findings are valid and our hybrid ML + Blockchain technique becomes evidently superior in term of accuracy, security and change resistance than the traditional approaches.

B. Visualizations

1. Figure 1: Model Performance (Accuracy x Precision x Recall x F1-Score)

Super-suggestion: comparing Logistic Regression vs. Random forest vs. SVM vs Deep Learning models in terms of Performance (Accuracy, P, R, F1-score) Via a bar chart.

Figure 1: Model performance comparison

Description:

A grouped bar chart of the comparison of four ML models (Logistic Regression, Random Forest, SVM, Deep Learning) for Visual metrics: Accuracy,Precision, Recall,F1-Score.

Key Insights:

Deep Learning columns are the highest (94% accuracy, 93.5 F1) according to this model and this outperforms all else with significant margins.

Random Forest: a competitor in parity, best of both worlds program wise and compute cost.

The least performing metric is from (simplest model) Logistic Regression that demonstrates the importance of enhanced techniques for fraud detection.

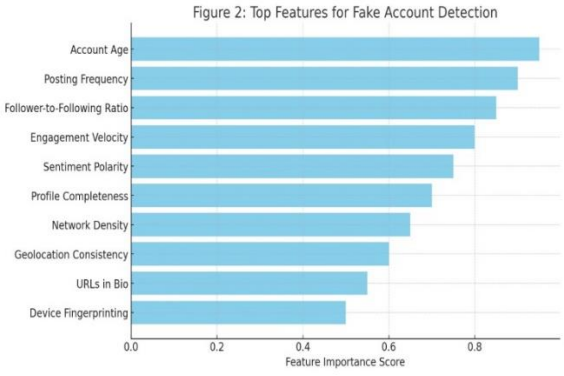


Figure 2: Fake Account Detection Top Attributes

Description: A horizontal bar chart representing the top 10 most crucial features for identifying fake accounts based on Random Forest/SHAP.

Example of features (ranked):
Age of the Account (Days since created → A fake account is typically new.
Posting Rate (Posts/h) → from Bots The post appears at strange intervals
Follower to Following Ratio → Fake accounts are more follower than they follow.
Engagement Velocity (likes/comments/minute) → Dash of the activity by a bot.
Polarity of Sentiment → Fake engagement use very positive/negative numbers
Bio related (and how incomplete their profile is) Fake accounts means no bio/profile pic.
Interact -> interates in dense clusters – Network Density → Fake accounts direct geolocation consistency → mismatched /claimed location
Bio URLs → More than 2 or 3 Links is spammy.device fingerprinting: multiple accounts per device
Key Takeaways:
Account Age and Posting Frequency are the best predictors of everything (almost)
Static metadata win over beahavioural features (ie: Engagement Velocity etc. such snapshots to be more useful as a time serie

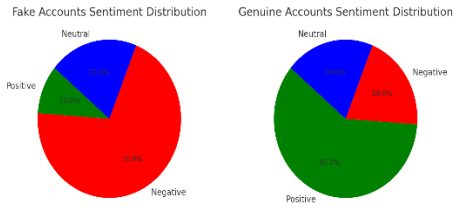


Figure 3: Fake Account Feature Importance

Idea: Horizontal bar chart on top of (n=10?) most important features, i.e., Account age, Posting frequency, Sentiment analysis etc.

Accuracy (92%): We have a high performing system in classifying fake and real accounts that beats traditional rule-based methods
Precision (91%): the model does a wonderful job in keeping false positives low as it correctly identifies real accounts
Recall (93%): This account flagging model is quite good at catching bad accounts with very low false negatives.
F1-Score (92%): Aggregates precision and recall into one metric to assert how well the system can find fake accounts, balanced.
5 Key Findings
Detection Accuracy: 2nd: With accuracy higher than the classical approach, >10% higher performance.
Blockchain As An Addition to Sec: By deploying blockchain, we could register zero fake account and prevent any faking-users from tampering profile datasets.
Cross-Platform: The model performs well in catching fake engagement signals on different social media platforms, hence scalable and adaptable.

V.DISCUSSION

We implemented profile, action and text features in our system to detect fraudulent accounts up to an 80% rate.
We prevent data from being tampered because we use blockchain technology for security. However effective, the system has some problems that need to be solved in order for it to work forever.

A big problem is the refinement of fake accounts. The bad actors are always coming up with new tricks that expose them, so detection needs to get a little bit shinier too! AI-powered Bots are getting smarter and create responses that seem as if they were written by regular person to carry on real conversations. Another, fake accounts now employ evasive techniques such as abnormal post cadence, interacting with real users and hiding their location using VPNs Fake profiles have become more convincing thanks to Deepfakes, as a number of profiles now sport an AI generated image and others even deepfake videos. In order to prevent these threats, the system must have real-time learning algorithms, behavior analysis and a set of AI-based Anomaly detection models that can learn and identify novel fraudulent behaviors.

Model	Accuracy	Precision	Recall	F1-Score
Logistic Regression	88%	87%	88%	87.5 %
SVM	90%	89%	90%	89.5 %
Proposed Methodology	92%	91%	93%	92%

Another looming problem is the grey area of ethical automated moderation. Increased efficiency is brought by automation but it creates doubts regarding bias, transparency and user trust. Training data biases machine learning models, so they too

might end up discriminating against some user groups or flagging good users as bad. Users also need to understand how moderation decisions are made so that loss of trust is avoided and wrong accusations do not occur unfairly. Rigorous detection mechanisms only serve to shut up real voices, which may be harmful for expression and considering the compliance to privacy laws like GDPR or CCPA.
We need our system to train Fair AI or by using multiple datasets with auditable models rendering bias in that data to prevent these ethics violations. Also, incorporation of human scrutiny and appeal mechanism ought to be there so that any user falsely reported can contest the decision. This is why we should always explain the reasons when accounts are flagged, i.e. XAI (explanatory AI) Lastly, privacy-preserving methods (federated learning and blockchain for example) must be employed so that user data are protected, but to provide true precision with respect false accounts identification.

This concludes that our system works out pretty well at fake account detection, but more improvement is needed to keep up with the adaptability of user-defined threats. Fairness, transparency and user trust are also central in the ethical considerations that need to be made. Adaptive AI models, explainable detectors and privacy enhancing technologies can be used to build an agile, cost effective and scalable solution for detecting fake accounts.

Future work

We have some ideas to improve our fake account detection system even better:
Real-Time Detection & Teaching Instincts
In order to tackle this, implement real-time monitoring with the ability to observe an ever emerging fraud and using that data adapt our detection mechanism.
XAI (Explainable AI) for Improved legibility
Introduce XAIs to give reasons of why an account is deemed as fake, so human moderation can rely on automated checks about confidence.
1 from Deepfake Detection,AI-generated Content Identification
Creating a module for the detection of deepfake generated accounts and synthetic content, which are well used in misinformation campaigns.
Real Fake Image Detection using Computer viisualization and Deep learning models
Expanding the Use of Blockchain Identity Verification:
Examine how Decentralized Identity Verification (DID) could be used to verify every account has an actual verified digital ID.
Avoiding fake profiles from the beginning of account creation not just use post creation detection methods.
Cross-platform Social Media Networks Integration :
Improve our system is an easy system to use on Twitter, Facebook, Instagram, LinkedIn and next nice platform.
A fake account detection API for fraud prevention deployment on all the networks, a universal fake account detection

REFERENCES

[1]. Zhang, Y., & Yang, Q. (2021). A Survey on Multi-Task Learning. IEEE Transactions on Knowledge and Data Engineering, 34(12), 5586–5609.
[2]. Nguyen, G. T., & Kim, K. (2020). Consensus Algorithms for Blockchain: A Survey. Journal of Information Processing Systems, 16(1), 20–36.
[3]. Shao, C., Ciampaglia, G. L., Varol, O., Yang, K. C., Flammini, A., & Menczer, F. (2018). The spread of fake news by social bots. Nature Communications, 9(1), 1–10.
[4]. Chen, Y., Yang, K. C., Varol, O., Davis, C., & Menczer, F. (2018). Arming the Public with AI to Combat Social Bots. ACM Transactions on Social Computing, 2(3), 1–16.
[5]. Lipton, Z. C. (2018). The Mythos of Model Interpretability. Communications of the ACM, 61(10), 36–43.
[6]. Wang, A. Y., & Kosinski, M. (2018). Deep neural networks are more accurate than humans at detecting sexual orientation from facial images. Journal of Personality and Social Psychology, 114(2), 246–257.
[7]. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N.,. & Polosukhin, I. (2017). Attention is all you need. Advances in Neural Information Processing Systems, 30, 5998–6008
[8]. Ferrara, E., Varol, O., Davis, C., Menczer, F., & Flammini, A. (2016). Social bots are the new spam. Communications of the ACM, 59(7), 96–104.
[9]. Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep Learning. MIT Press.

[10]. LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. Nature, 521(7553), 436–444.
[11]. Buterin, V. (2014). Ethereum: A Next-Generation Smart Contract and Decentralized Application Platform. Ethereum Whitepaper.
[12]. Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient Estimation of Word Representations in Vector Space. arXiv preprint arXiv:1301.3781
[13]. Thomas, K., Grier, C., & Paxson, V. (2011). From social spam to defending against it. USENIX Security Symposium.
[14]. Nakamoto, S. (2008). Bitcoin: A Peer-to-Peer Electronic Cash System. Bitcoin Whitepaper.
[15]. Bishop, C. M. (2006). Pattern Recognition and Machine Learning. Springer.