

Improving Fake Profile Detection: A Hybrid Machine Learning Approach with Negative and Clonal Selection

Ahmed Slimani ¹, Abdellatif Rahmoun ², Chahreddine Medjahed ³, Freha Mezzoudj ⁴

¹ LabRI-SBA Lab Ahmed Draia University – ADRAR, ah.slimani@esi-sba.dz

² LabRI-SBA Lab The Higher School of Computer Science 08 May 1945 of Sidi Bel Abbas Algeria, a.rahmoun@esi-sba.dz

³ Computer Science Department, Hassiba Benbouali Chlef University, c.medjahed@univ-chlef.dz

⁴ The National Polytechnic School of Oran Algeria, freha.mezzoudj@enp-oran.dz

ARTICLE INFO

ABSTRACT

Received: 26 Dec 2024

Revised: 14 Feb 2025

Accepted: 22 Feb 2025

The spread of fake profiles on social networking websites is a growing problem, posing significant challenges to user safety and data integrity. This research addresses the critical need for efficient detection mechanisms by proposing a hybrid machine learning model that combines negative selection and clonal selection algorithms. The study employs three different datasets, including two from Instagram and Twitter, following a detailed methodology for data preprocessing, feature extraction, and model implementation. Negative selection is applied to filter out irrelevant samples, while clonal selection enhances the model by optimizing solution discovery. The results show that the hybrid approach significantly improves detection accuracy, with a Random Forest model achieving an impressive 99% accuracy. Precision, recall, and F-measure tests further demonstrate the superiority of this new method over traditional techniques. The outcome of the research determines the efficacy of using negative and clonal selection algorithms together to detect fraudulent profiles more efficiently and maintain digital integrity. The research forms the foundation for additional research in this field.

Keywords: Fake Profile Detection, Machine Learning, Hybrid Approach, Negative Selection, Clonal Selection, Social Media Security.Learning

INTRODUCTION

In the digital age, social media platforms are central to promoting connection and construction communities. However, his broad presence has given rise to an important case: the spread of false profiles. These false profiles, often made with malicious intentions, cause serious threats to security and integrity to online communities. They are used to spreading regular resolution, cheating, participating in harassment and online facilities.

Meta has made significant advances in solving the problem. In the third quarter of 2024, the company removed 1.1 billion fake accounts after removing 1.5 billion in the last quarter. For reference, 2.2 billion false profiles were removed in the first quarter of 2019. Meta classifies these false accounts, created with malicious intentions or used to use businesses, organizations or non-human institutions [1]

Twitter reports approximate percentages of fake, spam, and automated accounts in its filings. Its 2013 Annual Report noted that less than 5% of accounts were fake or spam. A separate figure of 11% referred to active users utilizing automated applications, though the overlap between these figures remains unclear. The 5% statistic was reiterated in subsequent reports, while the 11% figure dropped to 8.5% by 2017. Since 2018, Twitter has continued to report the 5% statistic but stopped reporting the 8.5% figure [2].

Other platforms, such as LinkedIn, acknowledge fake accounts but do not provide specific numbers [3]. As fraudsters use increasingly sophisticated strategies, machine learning has emerged as a key tool in detecting and combating fake profiles. Machine learning excels at identifying patterns and anomalies that traditional methods often miss [4][5].

A promising solution to this growing problem is an artificial immune system (AIS), a calculation model inspired by the ability to identify, adapt and prevent biological immune systems to hazards. AIS uses the mechanisms as negative selection and clonal selection to detect deviations and to adapt to new patterns. This flexibility makes AIS very effective in areas such as deviations, cyber security and fraud prevention, as it reflects the ability of the immune system to distinguish "self" and "non-self" institutions [6].

It is important to detect and address fake profiles to ensure safety for social media platforms. The machine learning technique provides promising solutions to reduce the risk generated by these fraud accounts. This article suggests a novel A Hybrid Machine Learning approach, which integrates AIS concepts. In particular, it exploits the principles of negative selection and clonal selection, both are inspired by human immune systems, to detect false profiles on social media platforms.

The negative selection allows the system to identify the weird profiles by identifying the "non-self" institutions, while clonal selection allows continuous learning and adaptation that new types of false profiles emerge. This hybrid approach provides a more strong, adaptive and scalable solution, which improves the accuracy and effectiveness of false profile detection. By combining machine learning with immunological principles, the proposed system provides a state -Er -the solution for a growing digital security challenge.

The paper is structured as follows: Section 2 reviews related work, Section 3 presents the methodology, Section 4 discusses experimental results, and Section 5 concludes with future perspectives

RELATED WORK

A study [7] suggested a majority measurement approach to detect false social media profiles by combining classes such as trees, xgboost, random forest and logistic regression. This outfit method improved the accuracy of the detection by including 99.12% accuracy, accuracy, recalling and F1 score by incorporating different user behavior patterns, thus significantly improving social media safety.

Another article [8] Machine Learning (ML) and Deep Learning (DL) compared to Spam classification of E -Post. It was found that traditional algorithms such as logistic regression and naive bays perform well, a nervous network (Ann) achieved high accurately (98%) and F1 score (97.5%), highlighting the importance of setting models for optimal performance in spam filtration.

A clear AI (XAI) approach for detection of text spam was presented in [9], which integrates machine learning models such as SVM, logistics region, Grade -Grade Boosting and Decision Tree. By using random forest that meta classifies, the approach reached 98% recall, 96% accurate and 97% F1 score. The integration of XAI provided model transparency and increased user confidence with size values, which caused the approach both effective and explanatory.

For Instagram, a study used [10] used a monitored machine learning approach, and combined the supporting Vektads (SVM) and Random Forest (RF), called the SVM-RF algorithm. This method improved algorithms such as logistic region, AN and Navel Bayes, and demonstrated its efficiency in detecting fake accounts.

Sarahan and Matar [11] also focused on detecting fake accounts on Instagram using a series of machine learning techniques such as SVM, ANN, RF and KNN. Their results showed the capacity of machine learning, although performance varies depending on the dataset and model -selected models.

A comparative study by Azer & Al.[12] Stacking Model performed better, accuracy reached 99.8% on Facebook and 99.4% on Twitter, while Federated Learning provided increased secrecy.

Anila and Al. [13] introduced a hybrid approach that combines traditional machine learning techniques with deep learning to increase false profile detection. This model took advantage of K-nest neighbors and supported vector machines to improve both accuracy and efficiency in detecting a fraud profile.

Venkatesh et al. [14] addressed the difficulty of imbalanced datasets in faux profile detection, presenting a multi-level stacked ensemble version that integrates chi-squared function-magnificence affiliation, ensemble class, and price-touchy mastering. Their method confirmed superior precision compared to conventional strategies.

Sallah et al. [15] incorporated system mastering with genetic algorithms for feature choice in faux account detection, attaining high AUC values (90%–99.6%) on Facebook and Instagram datasets. Their technique no longer simplest improved detection however also decreased computational prices by way of minimizing the function area.

Shah and Al. [16] in false profile detection, traditional machines compared to the performance of learning and deep learning techniques. Their studies have shown that deep education, especially LSTM, performed better on large datasets, while machine learning techniques such as XGBOOST performed excellent on small datasets. He also emphasized the use of Smote technology to handle unbalanced datasets.

Venkatsavarlu and Sheno [17] suggested a new approach to detect Twitter spam using data streams in real time. Their method, in contrast to the Yeo-Johnson transformation, Renyi Entropy for feature fusion, and a regular side effect network (GAN) for spam detection, achieved 97.3% accuracy, a recall of 99.2%, and an F-score of 98.2%.

METHODOLOGY

Artificial Immune Systems (AIS) draw concept from biological immune structures to copy mechanisms like detection, tolerance, and flexibility inside computational frameworks. These structures use pattern reputation, adaptive mastering, and immune reminiscence to resolve complex issues in areas along with cybersecurity, fraud detection, and system optimization. AIS models, like negative selection and antibody networks, can discover anomalies in statistics through mimicking the immune machine's potential to distinguish between self and non-self. This adaptive and resilient technique has brought about AIS being widely explored in various technological applications [18].

Clonal Selection Algorithm (CSA) is a key AIS idea that simulates how biological immune structures generate a diverse variety of antibodies to fight pathogens. CSA replicates biological learning and reminiscence procedures, making it effective for complex optimization troubles and anomaly detection. Figure 1.a provides a complementing diagram of the workflow phases of the CSA method.

Key Features of CSA:

Antibody Diversity: CSA generates a various set of antibodies to discover a hassle area from a couple of perspectives, improving overall performance.

Affinity Evaluation: CSA evaluates antibodies based totally on their capability to apprehend troubles (antigens), deciding on people with excessive affinity for replication.

Selective Cloning: Antibodies with high affinity are cloned to further discover promising areas of the solution area.

Affinity Maturation: Cloned antibodies undergo mutation, enhancing their capacity to remedy the problem through the years.

CSA's concepts antibody variety, affinity evaluation, selective cloning, and affinity maturation make it tremendously adaptable and powerful in solving optimization and anomaly detection issues across multiple fields [19][20][21].

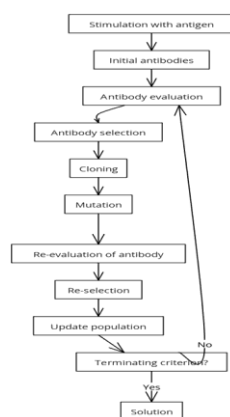


Fig.1a

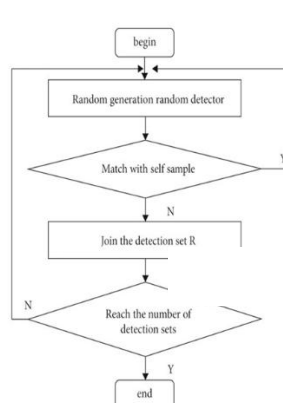


Fig.1b

Figure 1: a- Flow chart diagram describing stages of CSA algorithm.

b- Flow chart diagram describing stages of Negative Selection Algorithm

The Negative Selection Algorithm (NSA) is a fundamental issue of artificial immune systems, especially effective in anomaly detection and sample popularity [22]. Its basic concept is illustrated in Figure 1.b three, which suggests the organizational structure after negative choice. Inspired with the aid of the organic system wherein immune cells distinguish between the frame's personal cells and overseas entities, NSA establishes a strong mechanism for figuring out deviations from normal behavior. Specifically, the set of rules generates detectors trained on a dataset representing the "self" or normal behavior. These detectors are then used to examine incoming facts, flagging any sizeable deviations as anomalies [23]. This potential to locate unknown patterns makes NSA specially treasured in fields including cybersecurity and fault analysis, in which it may discover capability threats without earlier know-how of precise styles [24].

Classification Model Optimization with Immunology-Inspired Techniques

In this have a look at, we advise an modern classification pipeline that integrates gadget mastering (ML) with optimization mechanisms stimulated by way of immunology. The aim is to beautify classification overall performance by way of incorporating key immunology-stimulated strategies: terrible choice and clonal choice.

Data Preparation and Classification

The approach is carried out to a few distinct datasets: two datasets from Instagram and one from Twitter. Each dataset undergoes its personal method of loading, cleaning, transformation, and function engineering to make certain that it is prepared for evaluation. Once the datasets are prepared, the class is performed using several device mastering algorithms, which include Artificial Neural Networks (ANN), XGBoost, Support Vector Machines (SVM), and Random Forest. This step allows us to assess and benchmark the performance of each model on each dataset one at a time.

Optimization Phase

After the category process, an optimization section is brought to beautify model accuracy and robustness. This phase integrates immunology-inspired techniques: bad choice and clonal selection.

Negative Selection: The intention of negative selection is to identify misclassifications within the version, especially false positives (bad examples incorrectly labeled as high quality) and false negatives (positive examples incorrectly classified as bad). These misclassified examples are handled as anomalies and used to improve the model's accuracy.

Methodology: Apply the model to the test set. Identify misclassified examples (fake positives and fake negatives). Treat these misclassified examples as anomalies.

Clonal Selection: Once anomalies are recognized, the approach of clonal choice is used to improve the model. This entails cloning the anomalies and making use of mutation to those clones, mimicking the biological immune device's procedure of refining antibodies. By adding the cloned and mutated anomalies again into the education set, the version learns from its errors, thereby enhancing its capacity to deal with extra complex cases.

Methodology: Clone the anomalies detected via terrible selection. Apply mutation to the cloned anomalies to generate new variations. Retrain the version by means of incorporating the cloned and mutated anomalies into the education set.

Figure 2 illustrates the proposed pipeline, highlighting the mixing of these immunology-inspired optimization techniques into the class method. Through this integration, the type model is optimized, turning into an increasing number of adaptive and capable of improving its accuracy over time.

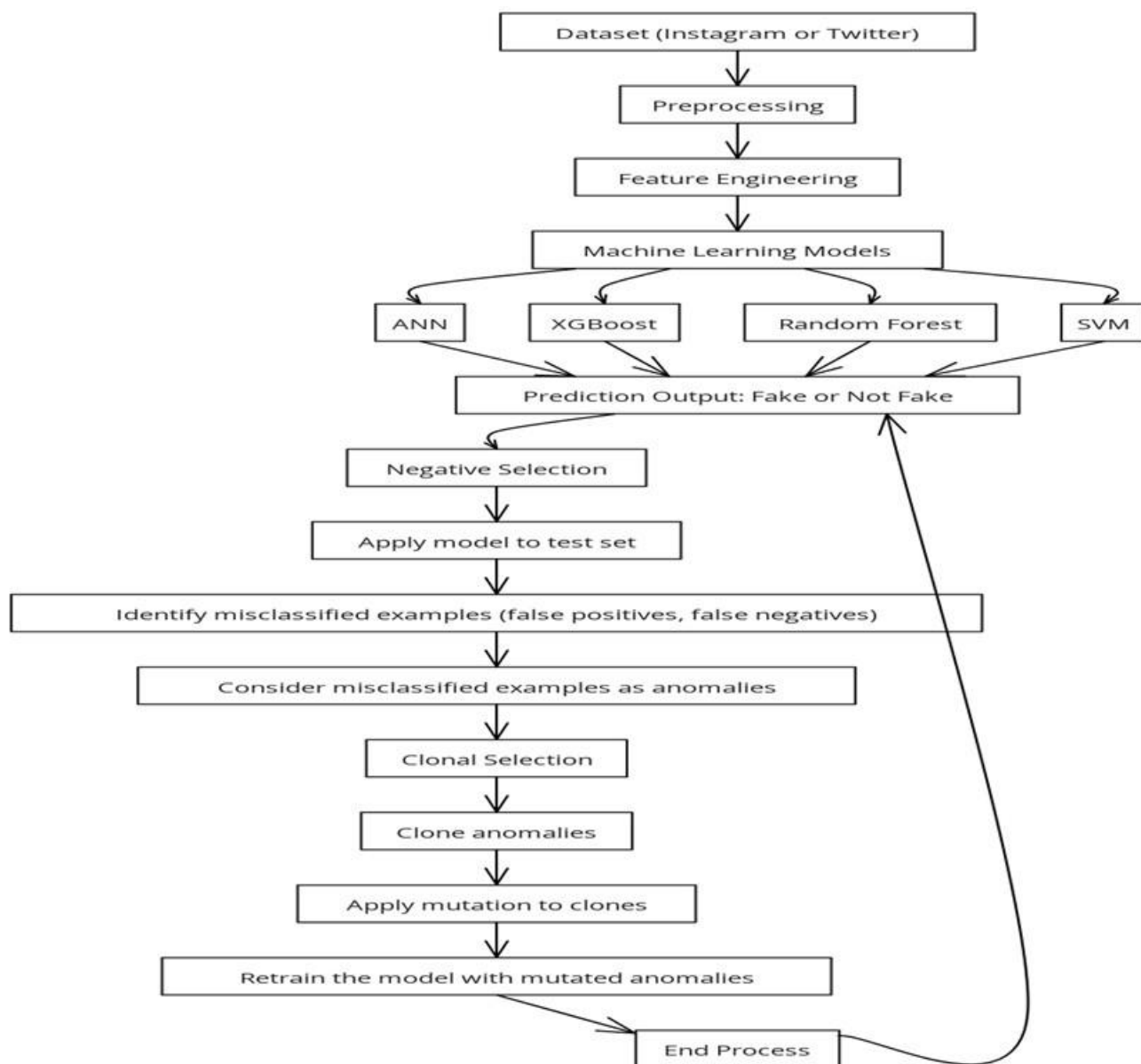


Figure 2: Immunology-Inspired Classification and Optimization Pipeline for Fake Profile Detection

Used Dataset

This observe utilizes three publicly available datasets. The first, created via Jafari [25], consists of seven-hundred fake Instagram bills acquired through purchase, with facts collected the use of a C# net scraper. Jafari additionally blanketed real money owed from his non-public Instagram. The 2nd dataset, from Bakhshandeh [26], identifies faux accounts thru guide inspection, though a few classification mistakes had been cited. Data became collected between March 15 and March 19, 2019, using a web crawler. The third dataset pertains to the Twitter element, divided into classes: faux and actual money owed. Each row represents a person profile, with various attributes. Some of those functions have been used to teach and compare our category models [27].

Database Preprocessing and Model Training

Before training the device learning fashions, we carried out essential preprocessing steps to put together the Twitter dataset. These steps protected visualizing characteristic distributions the usage of libraries like Seaborn and Matplotlib, dealing with missing values by removing or imputing them primarily based at the feature context, and normalizing numerical attributes for steady scaling. For categorical features, we implemented label encoding to

transform them into numerical values. The dataset was then cut up into schooling and testing sets to evaluate model performance.

Splitting into Training and Testing Sets: The preprocessed dataset was divided into elements: a training set, used to train the device mastering models the way to discover styles in the records, which shape 70% of dataset, and a testing set, reserved for comparing model overall performance on previously unseen instances, which form 30% of dataset. This guarantees a fair assessment of the version's generalization capability.

RESULTS

Pre-Negative Selection Clonal Selection Machine Learning Model

We carried out and skilled more than a few system studying models, which includes Artificial Neural Networks (ANN), Random Forest (RF), SVM, XGBoost, and evaluated their overall performance using the validation dataset. The accuracy scores for each of these models are presented in table 1

Algorithm	Accuracy before Negative Selection and clonal selection		
	Dataset 1 instagram	Dataset 2 instagram	Data set twiter
SVM	89.60%	88.13	91.0%
ANN	91.08%	93.64	93.2%
Random Forest	93.06%	94.49	95.27%
XGBoost	94.05%	94.06	92.1%

Table 1: Performance Metrics Accuracy before negative selection+ clonal selection.

Model Training After Negative Selection+ Selection clonale

Algorithm	Accuracy after Negative Selection and clonal selection		
	Dataset 1 instagram	Dataset 2 instagram	Data set twiter
SVM	91.09%	90.13	91.84%
ANN	95.05%	94.06	95.27%
Random Forest	96.53%	96.61	99.05%
XGBoost	99.01%	95.76	97.75%

Table 2: Performance Metrics Accuracy after negative selection+ clonal selection.

The software of Negative Selection and Clonal Selection optimization strategies verified a considerable improvement in the overall performance of diverse system getting to know models. These techniques greater the fashions' capacity to generalize and reduce overfitting, specifically inside the context of faux account detection. The fashions, such as Random Forest, XGBoost, ANN, and SVM, showed big accuracy gains, with XGBoost reaching the best improvement across all datasets. This suggests that Negative Selection and Clonal Selection are powerful gear for refining category algorithms, making them greater robust and dependable in handling complicated, real-international datasets. These findings underscore the significance of optimization techniques in advancing system gaining knowledge of fashions for excessive-stakes packages, which includes detecting fraudulent or fake accounts.

data set 01	Accuracy	References	data set 02	Accuracy	Reference s	Dataset 3	accurrac y	Reference s
ANN	91.08%	[11]	ANN	92.14%	[11]	ANN	93.42%	[27]
	95.05%	Our Study		94.06	Our Study		95.27%	Our Study
Random	92.99%	[11]	Random	90.71%	[11]	Random	94.14%	[27]

data set 01	Accuracy	References	data set 02	Accuracy	Reference s	Dataset 3	accurrac y	Reference s
Forest	96.53%	Our Study	Forest	96.61%	Our Study	forest	99.05%	Our Study
SVM	91.71 %	[11]	svm	91.42 %	[11]	Svm	90.42%	[27]
	91.09%%	Our Study		90.13%	Our Study		91.84%	Our Study

Table 1: Comparison of our model With Other Studies.

The evaluation of model accuracy across exceptional datasets highlights the effectiveness of optimization techniques. ANN confirmed consistent improvement, accomplishing better accuracy in our study (ninety five.05%) as compared to prior results (91.08%). Random Forest also benefited from optimization, with accuracy rising from 92.99% to 96.53% and ninety six.61%. SVM validated smaller upgrades, indicating its overall performance may have plateaued with these datasets. Overall, ANN and Random Forest significantly benefited from Negative Selection and Clonal Selection, whilst SVM showed more modest gains, suggesting it can no longer absolutely take advantage of these optimization strategies. The findings underscore the value of these optimization strategies, particularly for ANN and Random Forest in complex classification tasks.

CONCLUSION

In this observe, we targeted on detecting fraudulent profiles using numerous system gaining knowledge of algorithms, together with SVM, ANN, Random Forest, and XGBoost, before and after applying Negative Selection and Clonal Selection optimization strategies. The outcomes suggest a big improvement in model overall performance throughout extraordinary datasets (Instagram and Twitter).

Random Forest emerged because the pinnacle-appearing version, accomplishing 96.53% accuracy on the first Instagram dataset, 96.61% on the second one, and ninety nine.05% on the Twitter dataset. XGBoost additionally showed robust overall performance, with 99.01% accuracy on the primary Instagram dataset, 95.76% on the second, and 97.75% on Twitter. The overall performance of SVM and ANN stepped forward extensively after optimization, though they remained slightly behind Random Forest and XGBoost. The application of Negative Selection and Clonal Selection proved powerful in balancing precision, keep in mind, and F1-rating, main to greater correct detection of each fraudulent and proper profiles.

Future work in fraudulent profile detection can explore numerous promising avenues, along with the mixture of Negative Selection and Clonal Selection with other optimization strategies like genetic algorithms or particle swarm optimization to similarly beautify version overall performance. Additionally, integrating deep learning strategies inclusive of CNNs or transformers could seize more complicated styles in profile data. Techniques like SMOTE or adaptive resampling may be implemented to handle magnificence imbalance and decrease fake negatives, at the same time as characteristic engineering—together with using textual content embeddings or temporal conduct analysis—should improve version inputs. Testing the method on real-world social media or e-commerce datasets will check its scalability and robustness, and incorporating explainable AI (XAI) methods will provide interpretable insights, fostering believe and transparency. Ultimately, this observe highlights the effectiveness of Negative Selection and Clonal Selection in optimizing gadget learning models, and future work will aim to enhance model robustness, generalizability, and interpretability to better address actual-global challenges in fraudulent profile detection.

REFERENCES

- [1] Stacy Jo Dixon Facebook: fake account removal as of Q3 2024 [https:// www .statista .com / statistics /1013474 /facebook-fake-account-removal-quarter/](https://www.statista.com/statistics/1013474/facebook-fake-account-removal-quarter/)
- [2] Twitter Inc 2021 Annual Report and Form 10K . Available online <https://www.annualreports.com/Company/twitter-inc>

- [3] LinkedIn Announces Third Quarter 2016 Results; Available online <https://news.linkedin.com/2016/linkedin-announces-third-quarter-2016-results>.
- [4] Zhao, J., Zhang, J., Yan, J., & Zhang, W. A Survey on Fake Profile Detection in Online Social Networks: Challenges and Solutions. *IEEE Transactions on Computational Social Systems*, (2021). 8(3), 517-529.
- [5] Wang, S., Hu, X., & Liu, X. Detecting Fake Accounts in Online Social Networks Using Machine Learning Algorithms. *Expert Systems with Applications*, (2022). 184, 115496.
- [6] De Castro, Leandro & Timmis, Jon. (2002). Artificial immune systems: A new computational intelligence approach.
- [7] Patil DR, Pattewar TM, Punjabi VD, Pardeshi SM. Detecting Fake Social Media Profiles Using the Majority Voting Approach. *EAI Endorsed Scal Inf Syst* [Internet]. 2024 Feb. 14 [cited 2025 Feb. 28];11(3). Available from: <https://publications.eai.eu/index.php/sis/article/view/4264>
- [8] Alsuwit, M.H., Haq, M.A. and Aleisa, M.A. Advancing Email Spam Classification using Machine Learning and Deep Learning Techniques. *Engineering, Technology & Applied Science Research*. 14, 4 (Aug. 2024), 14994–15001. DOI: <https://doi.org/10.48084/etasr.7631>
- [9] A. Alzahrani, Explainable AI-based Framework for Efficient Detection of Spam from Text using an Enhanced Ensemble Technique. *Engineering, Technology & Applied Science Research*. 14, 4 (Aug. 2024), 15596–15601. DOI: <https://doi.org/10.48084/etasr.7901>.
- [10] Ahdi Hassan; Abdalilah. G. I. Alhalangy; Fahad Alzahrani. Fake Accounts Identification in Mobile Communication Networks Based on Machine Learning. *Int. J. Interact. Mob. Technol.* 2023, 17, pp. 64-74.
- [11] Sarhan, Fatema & Mattar, Ebrahim.. Fake accounts detection in online social networks using hybrid machine learning models. *International Journal of Simulation: Systems, Science & Technology*, (2023) 24. DOI: [10.5013/IJSSST.a.24.02.02](https://doi.org/10.5013/IJSSST.a.24.02.02).
- [12] Azer, Marina, Hala H. Zayed, Mahmoud E. A. Gadallah and Mohamed Taha. “Multi platforms fake accounts detection based on federated learning.” *IAES International Journal of Artificial Intelligence (IJ-AI)* (2024): DOI: [10.11591/ijai.v13.i4.pp3837-3848](https://doi.org/10.11591/ijai.v13.i4.pp3837-3848)
- [13] Anila S, Meenakshi Mohan, Mariya Jacob, Najiya Nasrin Fake Social Media Profile Detection: A Hybrid Approach Integrating Machine Learning and Deep Learning Techniques *International Journal of Scientific Research & Engineering Trends* Volume 10, Issue 4, July-Aug-2024, ISSN (Online): 2395-566X
- [14] Swetha Chikkasabbenahalli Venkatesh, Sibi Shaji, and Balasubramanian Meenakshi Sundaram, “A Fake Profile Detection Model Using Multistage Stacked Ensemble Classification”, *Proc. eng. technol. innov.*, vol. 26, pp. 18–32, Feb. 2024.
- [15] Sallah, A., Abdellaoui Alaoui, E.A., Hessane, A. et al. An efficient fake account identification in social media networks: Facebook and Instagram using NSGA-II algorithm. *Neural Comput & Applic* 36, 21487–21515 (2024). <https://doi.org/10.1007/s00521-024-10350-8>
- [16] Shah, A., Varshney, S. & Mehrotra, M. Detection of Fake Profiles on Online Social Network Platforms: Performance Evaluation of Artificial Intelligence Techniques. *SN COMPUT. SCI.* 5, 489 (2024). <https://doi.org/10.1007/s42979-024-02839-9>
- [17] B. Venkateswarlu, and Viswanath Sheno V. “Optimized Generative Adversarial Network with Fractional Calculus Based Feature Fusion Using Twitter Stream for Spam Detection.” *Information Security Journal: A Global Perspective* 31 (5) (2021): 582–601. doi: 10.1080/19393555.2021.1956024.
- [18] Timmis, J., Neal, L., & Hunt, J. (2000). *Artificial Immune Systems: A New Computational Intelligence Approach*. Springer-Verlag.
- [19] S. F. M. Burnet, *The Clonal Selection Theory of Acquired Immunity*, vol. 3. Nashville, TN, USA: Vanderbilt Univ. Press, 1959.
- [20] V. Cutello, G. Nicosia, M. Pavone, and J. Timmis, “An immune algorithm for protein structure prediction on lattice models,” *IEEE Trans. Evol. Comput.*, vol. 11, no. 1, pp. 101–117, Feb. 2007.
- [21] J. Kelsey and J. Timmis, “Immune inspired somatic contiguous hypermutation for function optimisation,” in *Proc. Genetic Evol. Comput. Conf.* Springer, 2003, pp. 207–218.
- [22] S. Forrest, S. Hofmeyr, and A. Somayaji. *Computer immunology*. *Communications of the ACM*, 40(10):88–96, 1997.

- [23] Gao, XZ., Wang, X., Zenger, K., Wang, X. (2011). Negative Selection Algorithm-Based Motor Fault Diagnosis. In: Wang, Y., Li, T. (eds) Practical Applications of Intelligent Systems. Advances in Intelligent and Soft Computing, vol 124. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-25658-5_20
- [24] Kuldeep Singh, Lakhwinder Kaur, Raman Maini A survey of intrusion detection techniques based on negative selection algorithm International Journal of System Assurance Engineering and Management, 2022, Volume 13, Number 1, Page 175
- [25] R. Jafari, "Instagram fake and real accounts dataset," Kaggle, 07-Jan2021. [Online]. Available: <https://www.kaggle.com/datasets/rezaunderfit/instagram-fake-andreal-accounts-dataset>.
- [26] B. Bakhshandeh, "Instagram fake spammer genuine accounts," Kaggle, 22-Mar-2019.[Online]. Available: <https://www.kaggle.com/datasets/free4ever1/instagram-fakespammer-genuine-accounts>
- [27] Harshitkgupta. (n.d.). GitHub - harshitkgupta/Fake-Profile-Detection-using-ML: Detect fake profiles in online social networks using Support Vector Machine, Neural Network and Random Forest. GitHub. <https://github.com/harshitkgupta/Fake-Profile-Detection-using-ML>