



SNESTIK

Seminar Nasional Teknik Elektro, Sistem Informasi,
dan Teknik Informatika

<https://ejurnal.itats.ac.id/snestik> dan <https://snestik.itats.ac.id>



Informasi Pelaksanaan :

SNESTIK V - Surabaya, 26 April 2025

Fakultas Teknik Elektro dan Teknologi Informasi, Institut Teknologi Adhi Tama Surabaya

Informasi Artikel:

DOI : 10.31284/p.snestik.2025.7343

Prosiding ISSN 2775-5126

Fakultas Teknik Elektro dan Teknologi Informasi-Institut Teknologi Adhi Tama Surabaya
Gedung A-ITATS, Jl. Arief Rachman Hakim 100 Surabaya 60117 Telp. (031) 5945043
Email : snestik@itats.ac.id

DETEKSI AKUN PALSU MEDIA SOSIAL X (TWITTER) BERBASIS *MACHINE LEARNING*

Adisha Putri Az Zahra, Rimuljo Hendradi, Faried Effendy

Program Studi Sistem Informasi, Universitas Airlangga

e-mail: adisha.putri.az-2021@fst.unair.ac.id

ABSTRACT

In early 2024, X was ranked as the sixth most used social media in Indonesia. However, the growing has led to challenges, including an increase in fake accounts on X. Fake accounts are often used to spread false information (hoaxes), manipulate public opinion, incite cyberbullying and distribute spam that can lead to criminal activities such as fraud. Therefore, this study aims to identify patterns associated with fake accounts and develop a classification model for detecting fake accounts on X. The research stages include data collection, data preprocessing, feature selection, clustering, classification and model evaluation. Using a dataset of 8291 rows, the optimal K value for clustering was determined to be K=4. Cluster analysis revealed that Cluster 1 represents fake accounts, characterized by their short account age (often newly created), 99% using the default profile picture and a significantly higher number of followings compared to followers. The classification model was developed using the labeled dataset post-clustering, where Cluster 1 was labeled as fake accounts and Cluster 0, 2 and 3 were labeled as genuine accounts. Three algorithms—Random Forest, AdaBoost, and KNN—were tested across four data-splitting schemes. Random Forest model achieved the highest recall, F1-score, and accuracy at 99.35%, 98.71%, and 98.55%, respectively.

Keywords: Machine learning; fake account detection; x (twitter); kmeans clustering; classification

ABSTRAK

Pada awal tahun 2024, X menduduki peringkat keenam sebagai media sosial yang paling banyak digunakan di Indonesia. Jumlah pengguna yang terus meningkat ini dapat menimbulkan masalah seperti meningkatnya

jumlah akun pengguna palsu pada media sosial *X*. Akun palsu kerap kali digunakan sebagai media penyebaran informasi palsu (hoax), manipulasi opini publik, memicu tindak kriminal *cyber bullying* (perundungan siber) hingga penyebaran *spam* yang dapat berujung pada tindak kriminal penipuan. Dengan demikian, penelitian ini bertujuan untuk mengidentifikasi pola pada akun palsu serta pengembangan model klasifikasi deteksi akun palsu pada media sosial *X*. Tahapan pada penelitian ini meliputi proses pengumpulan data, data *preprocessing*, seleksi fitur, proses klusterisasi, tahap klasifikasi hingga evaluasi model. Dengan dataset sejumlah 8291 baris, nilai *K* optimal untuk proses *clustering* yakni *K*=4. Dari proses analisis kluster, *cluster* 1 menunjukkan kelompok akun dengan karakteristik akun palsu. Hal ini ditandai dengan usia akun yang tergolong masih muda hingga baru, sebanyak 99% akun masih menggunakan profil default, serta jumlah *following* yang jauh lebih banyak dari jumlah *followers*. Pengembangan model klasifikasi akan menggunakan dataset yang telah diberikan label setelah proses *clustering*. *Cluster* 1 akan dilabeli sebagai akun palsu serta *cluster* 0, 2 dan 3 diberi label sebagai akun asli. Dengan menggunakan tiga jenis algoritma yakni *Random Forest*, *AdaBoost* dan *KNN* serta empat skema pembagian dataset diperoleh *recall*, *F1-Score* dan akurasi tertinggi pada model *Random Forest* dengan persentase berturut-turut 99,35%, 98,71% dan 98,55%.

Kata kunci: Pembelajaran mesin; deteksi akun palsu; *x* (twitter); *k* means clustering; klasifikasi.

PENDAHULUAN

Dilansir dari situs Katadata, media sosial *X* (dahulu *Twitter*), menduduki peringkat keenam sebagai media sosial yang paling banyak digunakan di Indonesia pada awal tahun 2024 [1], dengan total pengguna sebanyak 24,69 juta atau 8,9% dari total populasi Indonesia. Jumlah ini meningkat dari data pengguna *X* pada awal tahun 2023 sebanyak 2,87% [2]. Seiring dengan meningkatnya pengguna media sosial *X*, hal tersebut juga dapat menimbulkan masalah seperti meningkatnya jumlah akun pengguna palsu pada media sosial *X* [3].

Akun palsu adalah akun dengan kredensial palsu yang tidak dimiliki atau dikelola oleh manusia asli [4], [5]. Dimana akun palsu ini melanggar aturan pada media sosial karena dapat menyebarkan berita palsu hingga melakukan penyebaran *spam*. Selain itu, akun palsu merupakan akun yang dibuat dengan tujuan meningkatkan popularitas pengguna lain dimana akun ini cenderung memiliki sedikit pengikut (jumlah *followers*) dan jumlah pengikut (*following*) yang jauh lebih banyak [6]. Akun palsu kerap kali digunakan sebagai media penyebaran informasi palsu (hoax), manipulasi opini publik, tindak kriminal *cyber bullying* (perundungan siber), penyebaran tautan *phishing* hingga penyebaran *spam* yang dapat berujung pada tindak kriminal penipuan [3], [7], [8].

Akun palsu biasanya dapat diidentifikasi oleh pengguna dengan melihat jumlah *following* yang perbandingannya jauh lebih besar dengan jumlah *followers* [9]. Selain itu, Erşahin, et al., mengungkapkan, akun palsu kerap kali menggunakan tema, foto profil secara default serta tidak adanya deskripsi atau bio akun [4]. Namun, akun palsu yang kini semakin marak dan berkembang dengan lebih canggih sering kali sulit diidentifikasi oleh pengguna biasa [10]. Oleh karena itu, perlu adanya tindakan preventif dengan pendeteksian akun palsu secara otomatis [11]. Dalam konteks ini, salah satu metode yang kerap digunakan para peneliti dalam mendeteksi akun palsu dari berbagai macam media sosial adalah *machine learning* [12].

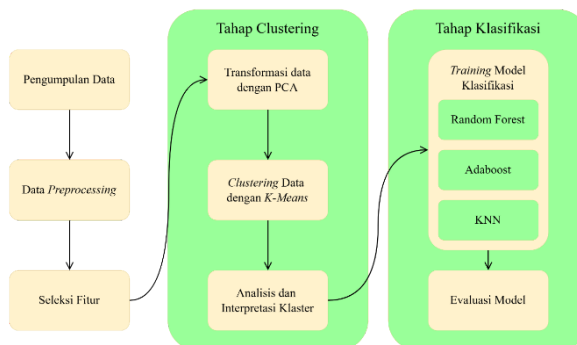
Adewole, et al., (2020) serta Alsubaei (2023), memanfaatkan *machine learning* dalam melakukan identifikasi pola akun palsu serta pengklasifikasian akun palsu pada platform media sosial. Penelitian oleh Adewole, et al., (2020) memanfaatkan PCA dan *K-Means clustering* untuk mengidentifikasi kluster akun-akun *spam* dari media sosial *X*. Hasil kluster dianalisis dan diberikan label non-*spam* atau *spam*. Data berlabel akan digunakan untuk membangun model klasifikasi deteksi akun-akun *spam* dengan membandingkan tiga algoritma klasifikasi *machine learning* yakni Multilayer Perceptron (MLP), SVM dan *Random Forest*. Hasil penelitian menunjukkan algoritma *Random Forest* mengalami peningkatan signifikan dari penelitian sebelumnya dengan nilai akurasi yang diperoleh adalah 96,30% [8]. Sebanyak 1337 akun palsu dan 1481 akun asli pada dataset penelitian Alsubaei (2023) digunakan dalam melatih model dengan tujuan pembuatan bot detector akun palsu dari platform *Twitter*. Dari proses klasifikasi

diperoleh akurasi terbesar yakni 99,42% pada skema pelatihan 80:20 menggunakan *classifier Random Forest* [13].

Berdasarkan pemaparan yang telah disebutkan sebelumnya, penelitian ini berfokus pada identifikasi pola akun palsu serta pengembangan model deteksi akun palsu pada media sosial X khususnya pada konten berbahasa Indonesia dengan memanfaatkan *machine learning*. Terdapat dua jenis *machine learning* yang akan digunakan pada penelitian ini yakni unsupervised learning dengan algoritma *K-Means* serta *supervised learning* dengan tiga algoritma diantaranya *Random Forest*, *AdaBoost* dan *KNN*. Penggunaan unsupervised learning dimanfaatkan dalam mengidentifikasi pola pada dataset yang tidak memiliki atribut “Target” dengan menghitung kemiripan antar data sehingga pola dengan tingkat kemiripan maksimum akan dikelompokkan menjadi satu cluster. Dari proses clustering, akan diperoleh output berupa dataset dengan atribut “Target” yang akan dimanfaatkan dalam pembuatan model deteksi akun palsu menggunakan algoritma supervised learning. Dengan demikian, penelitian ini diharapkan dapat mengeksplorasi penggunaan machine learning dalam pendeteksian akun palsu pada platform media sosial X berbasis bahasa Indonesia dengan mempelajari pola-pola unik dari akun palsu pada platform media sosial X berbahasa Indonesia serta dapat berkontribusi dalam upaya mengurangi penyebaran akun palsu di media sosial X, khususnya di Indonesia.

METODE

Penelitian ini memiliki beberapa tahapan dengan fokus utama penelitian yakni melakukan *clustering* untuk identifikasi pola pada akun palsu media sosial X serta pembuatan model klasifikasi untuk mendeteksi akun palsu media sosial X yang ditunjukkan pada Gambar 1.



Gambar 1. Tahapan penelitian.

Pengumpulan Data

Jenis dataset yang digunakan pada penelitian ini merupakan dataset primer yang diperoleh langsung melalui *crawling* pada *website X* menggunakan *library Python, Twikit*. Pengumpulan data dilakukan selama 10 hari pada rentang waktu dari tanggal 22 Juli 2024 hingga tanggal 10 September 2024. Dari pengumpulan data tersebut diperoleh dataset sebanyak 64350 baris. Dataset yang dikumpulkan memiliki 44 atribut yang diambil dari *tweets* atau cuitan pada media sosial X. Input yang digunakan pada proses pengumpulan data merupakan keyword atau kata kunci yang tersedia pada fitur *Trending* dari media sosial X.

Data Preprocessing

Pada tahap preprocessing, dataset akan dianalisis untuk menghilangkan inkonsistensi, ketidaklengkapan seperti penghapusan baris yang duplikat, pemilihan Bahasa Indonesia pada atribut “Lang”, pengkodean menggunakan digit biner antara 1 atau 0 untuk atribut dengan nilai ‘TRUE’ atau ‘FALSE’ serta pengecekan *missing value*. Selain itu, pemrosesan dan penambahan atribut pada dataset juga akan dilakukan pada tahap *preprocessing* untuk meningkatkan kualitas

data yang bertujuan menghasilkan efisiensi dalam proses *clustering* serta pembangunan model klasifikasi dan proses training.

Seleksi Fitur

Tahapan ini akan memilih atribut atau fitur yang digunakan baik dalam pengklasteran maupun proses klasifikasi dengan kriteria atribut paling relevan atau memiliki pengaruh signifikan dalam pendeteksian akun palsu pada media sosial X. Variance Threshold akan digunakan sebagai metode dalam seleksi fitur dengan menghitung nilai varians untuk setiap fitur menggunakan rumus pada persamaan 1.

$$\sigma^2 = \frac{\sum (x - \mu)^2}{N} \quad (1)$$

Persamaan 1 menunjukkan perhitungan varians dengan x mewakili nilai data (setiap data individu dalam dataset), μ adalah rata-rata dan N adalah jumlah baris data. Jika nilai varians berada di bawah ambang batas (threshold) maka fitur akan dihapus. Nilai *Variance Threshold* yang digunakan pada penelitian adalah nol, dimana fitur yang tidak memiliki variasi dalam data akan dihapus [14].

Tahap *Clustering* dan Klasifikasi

Tahapan ini memuat proses transformasi data dengan PCA dengan tujuan melihat persebaran kluster yang lebih rinci dengan adanya kontribusi dari masing-masing fitur pada dataset yang diwakilkan oleh *Principal Component 1* dan *Principal Component 2*, *clustering* dengan algoritma *K-Means* hingga proses analisis dan interpretasi kluster menggunakan acuan karakteristik akun X dari penelitian Dracewicz & Sepczuk [9] serta hasil rata-rata nilai kluster atau *centroid*.

Proses *training model klasifikasi* menggunakan dataset dengan atribut baru yakni “Target” yang merupakan *output* dari proses analisis kluster. Label dalam dataset terdiri dari dua kelas yakni ‘Akun Palsu’ dan ‘Akun Asli’. Model *machine learning* akan dilatih menggunakan empat skema pembagian data yakni 80:20, 70:30, 90:10, dan 67:33, dengan tujuan membedakan dua kelas pada atribut “Target” berdasarkan pola yang ada pada atribut-atribut input dari dataset. Dalam penelitian ini, akan digunakan tiga algoritma untuk membangun model deteksi akun palsu yakni *Random Forest*, *AdaBoost* dan KNN.

Random Forest adalah sebuah *class* dari algoritma *Decision Tree* yang berbasis pendekatan *ensemble* [8]. Dengan memuat sejumlah *decision tree* pada berbagai subset dari kumpulan data yang ada, algoritma *Random Forest* akan mengambil rata-rata berdasarkan mayoritas prediksi untuk meningkatkan akurasi prediksi dari kumpulan data tersebut [15]. Algoritma *AdaBoost* juga melibatkan training atau pelatihan dasar menggunakan algoritma dasar seperti *Decision Tree* [16]. Bobot sampel akan disesuaikan dengan prediksi *classifier* atau pengklasifikasi kemudian sampel tersebut akan digunakan untuk melatih *classifier* berikutnya. Oleh karena itu, bobot yang lebih besar akan diberikan pada sampel yang salah diklasifikasikan dan sampel yang diklasifikasikan dengan benar akan diberi bobot lebih rendah [16]. Sedangkan *K-Nearest Neighbors* (KNN) merupakan jenis algoritma *machine learning* yang memprediksi label atau target pada data *testing* dengan menghitung jarak antara data *testing* dengan semua titik *training*, kemudian menunjukkan jumlah titik k (*training*) yang dekat dengan data *testing* [17].

Evaluasi Model

Proses evaluasi tahapan klasifikasi pada penelitian ini dilakukan menggunakan *Confusion Matrix* dengan skema perbandingan dataset. Dengan membandingkan performa kinerja tiga model yakni *Random Forest*, *AdaBoost* dan KNN melalui metrik akurasi, presisi, *recall* hingga F1-score akan didapatkan model klasifikasi terbaik yang selanjutnya akan digunakan dalam pendeteksian akun palsu.

$$akurasi = \frac{TP+TN}{TP+FP+FN+TN} \quad (2)$$

$$presisi = \frac{TP}{TP+FP} \quad (3)$$

$$recall = \frac{TP}{TP+FN} \quad (4)$$

$$F1 \text{ Score} = 2 \cdot \frac{recall \times presisi}{recall + presisi} \quad (5)$$

Persamaan 2 menunjukkan perhitungan nilai akurasi yang mengukur seberapa akurat model klasifikasi membuat prediksi yang benar dari jumlah prediksi yang telah dibuat. Sedangkan, perhitungan nilai presisi ditunjukkan pada persamaan 3, dimana presisi akan mengukur seberapa baik model memprediksi benar untuk kelas positif dari semua prediksi positif yang dibuat oleh model. Persamaan 4 menunjukkan perhitungan nilai *recall* yang akan mengukur seberapa baik model dalam mengidentifikasi kelas positif dengan benar serta Persamaan 5 yang menunjukkan perhitungan nilai F1 Score yang menggambarkan keseimbangan antara presisi dan recall.

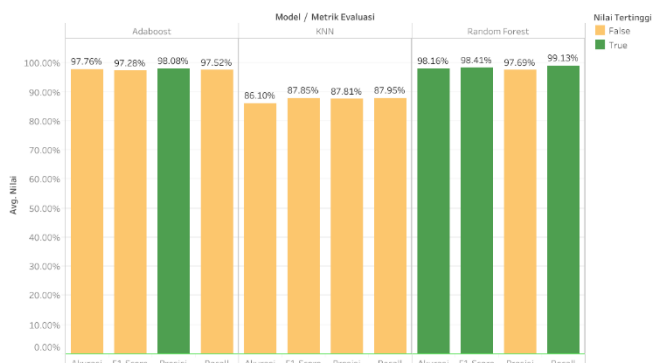
HASIL DAN PEMBAHASAN

Sebanyak 8029 baris dan 26 atribut berhasil didapatkan setelah proses data *preprocessing* hingga seleksi fitur. Dataset ini kemudian akan menjadi dataset final yang akan digunakan pada tahap *clustering* hingga evaluasi model. Dengan menggunakan metode *Elbow* didapatkan nilai K optimum pada dataset tersebut adalah K=4. Dari hasil *clustering*, diperoleh 3289 data pada cluster 0, 4756 data pada cluster 1, 127 data pada cluster 2 serta 119 data pada cluster 3. Hasil *clustering* akan ditambahkan pada dataset dengan nama atribut “Cluster” yang selanjutnya akan digunakan sebagai inputan pada tahapan analisis dan interpretasi klaster. Berdasarkan acuan karakteristik akun *X* dari penelitian Dracewicz & Sepczuk [9], untuk menganalisis pola pada *cluster* akan dilakukan pelabelan jenis akun pada dataset dengan indikator yang ditunjukkan pada Tabel 1.

Tabel 1. Indikator pelabelan jenis akun

Karakteristik	Jenis Akun			
	Bot	Cyborg	Real	Verified
Jumlah Followers	Very Low, Low	Very Low, Low	Low, Medium, High	High, Very High
Jumlah Following	Very High, High	Very High, High	Low, Medium, High	Medium
Usia Akun	New Account, Young Account	New Account, Young Account	Middle-aged Account, Old Account, Very Old Account	Middle-aged Account, Old Account, Very Old Account
Jumlah Tweets	Very High, High, Very Low	Very High, High, Very Low	Low, Medium, High	High, Very High
User Default Profile Image	0	0 atau 1	1	1
User Profile Banner Url	0	0 atau 1	1	1
User Description	0	0	0 atau 1	1
User Url	0	0	0 atau 1	0 atau 1
User Blue Verified	0	0	0	1
Jumlah	1954	2387	3033	377

Hasil distribusi tipe akun berdasarkan *cluster* menunjukkan *cluster* 1 memiliki jumlah tipe akun ‘Bot’ paling banyak di antara semua *cluster*. Hal ini kemudian didukung dengan analisis hasil rata-rata nilai klaster yakni pada atribut “Account Age (Days)” yang mendapatkan nilai rata-rata usia akun 1562,03 hari atau sekitar 4 tahun. Hal ini menunjukkan, dominasi akun yang relatif baru dibandingkan dengan *cluster* lainnya. Selain itu, pada *cluster* ini sebanyak 99% akun masih menggunakan profil default serta nilai rasio *followers* yang jauh lebih rendah dari semua *cluster*. Dengan demikian, *cluster* 1 akan diberikan target sebagai ‘Akun Palsu’ dengan total data yakni 4756 baris. Sedangkan pada *cluster* 0, 2 dan 3 menunjukkan karakteristik mirip dengan akun asli serta cenderung tidak memiliki perubahan nilai rata-rata klaster yang signifikan pada atribut usia akun, jumlah *followers*, rasio *followers* dan reputasi akun sehingga *cluster* 0, *cluster* 2 dan *cluster* 3 akan diberi target sebagai ‘Akun Asli’ dengan total data sebanyak 3535 baris.



Gambar 2. Visualisasi nilai rata-rata matriks evaluasi setiap model.

Dengan dataset yang sudah diberi target ‘Akun Palsu’ dan ‘Akun Asli’, kemudian dilakukan pembuatan model klasifikasi menggunakan tiga algoritma yakni *Random Forest*, *AdaBoost* dan KNN. Gambar 2 menunjukkan *Random Forest* menjadi model paling optimal dibandingkan dengan semua model. Rata-rata tertinggi dicapai oleh model *Random Forest* pada nilai recall sebesar 99,13%, akurasi dengan persentase 98,16% dan nilai F1-Score sebesar 98,41%. Sedangkan rata-rata presisi tertinggi diraih oleh model *Adaboost* dengan persentase sebesar 98,08%. Penelitian ini memiliki hasil akurasi yang lebih tinggi 2,35% dari penelitian oleh Adewole, et al. (2020) serta nilai recall yang lebih tinggi 0,2% namun lebih rendah 0,87% pada hasil akurasi dari penelitian oleh Alsubaei (2023).

KESIMPULAN

Hasil identifikasi pola pada akun palsu melalui metode *clustering* yakni karakteristik akun palsu mengacu pada usia akun yang masih baru hingga muda, penggunaan profil default atau belum dilakukan personalisasi hingga jumlah *following* yang lebih tinggi dari jumlah *followers*. Hal ini sesuai dengan kriteria yang terdapat pada *cluster* 1. Model klasifikasi yang dibuat dan dilatih dengan *machine learning* berhasil mencapai nilai tertinggi pada model *Random Forest* dengan rincian nilai akurasi sebesar 98,71%, recall dengan 99,35% serta F1-Score sebesar 98,71%. Adapun nilai presisi tertinggi dicapai oleh model *Adaboost* dengan nilai presisi sebesar 98,88%.

DAFTAR PUSTAKA

- [1] C. M. Annur, "Ini Media Sosial Paling Banyak Digunakan di Indonesia Awal 2024," databoks. [Online]. Available: <https://databoks.katadata.co.id/datapublish/2024/03/01/ini-media-sosial-paling-banyak-digunakan-di-indonesia-awal-2024>
- [2] S. Kemp, "Digital 2024: Indonesia," datareportal. [Online]. Available: <https://datareportal.com/reports/digital-2024-indonesia>
- [3] E. Van Der Walt and J. Eloff, "Using Machine Learning to Detect Fake Identities: Bots vs Humans," *IEEE Access*, vol. 6, pp. 6540–6549, 2018, doi: 10.1109/ACCESS.2018.2796018.
- [4] B. Erüahin, Ö. Aktaü, D. KÖlÖnç, and C. Akyol, "Twitter Fake Account Detection".
- [5] J. Ezarfelix, N. Jeffrey, and N. Sari, "Systematic Literature Review: Instagram Fake Account Detection Based on Machine Learning:," *Eng. Math. Comput. Sci. EMACS J.*, vol. 4, no. 1, pp. 25–31, Feb. 2022, doi: 10.21512/emacsjournal.v4i1.8076.
- [6] Y. Elyusufi, Z. Elyusufi, and M. A. Kbir, "Social networks fake profiles detection based on account setting and activity," in *Proceedings of the 4th International Conference on Smart City Applications*, Casablanca Morocco: ACM, Oct. 2019, pp. 1–5. doi: 10.1145/3368756.3369015.
- [7] M. Vyawahare and S. Govilkar, "Fake profile recognition using profanity and gender identification on online social networks," *Soc. Netw. Anal. Min.*, vol. 12, no. 1, p. 170, Dec. 2022, doi: 10.1007/s13278-022-00997-3.
- [8] K. S. Adewole, T. Han, W. Wu, H. Song, and A. K. Sangaiah, "Twitter spam account detection based on clustering and classification methods," *J. Supercomput.*, vol. 76, no. 7, pp. 4802–4837, Jul. 2020, doi: 10.1007/s11227-018-2641-x.
- [9] W. Dracewicz and M. Sepczuk, "Detecting Fake Accounts on Social Media Portals—The X Portal Case Study," *Electronics*, vol. 13, no. 13, p. 2542, Jun. 2024, doi: 10.3390/electronics13132542.
- [10] N. G. Kerrysa and I. Q. Utami, "Fake account detection in social media using machine learning methods: literature review," *Bull. Electr. Eng. Inform.*, vol. 12, no. 6, pp. 3790–3797, Dec. 2023, doi: 10.11591/eei.v12i6.5334.
- [11] P. Wanda, "RunMax: fake profile classification using novel nonlinear activation in CNN," *Soc. Netw. Anal. Min.*, vol. 12, no. 1, p. 158, Dec. 2022, doi: 10.1007/s13278-022-00983-9.
- [12] M. Smruthi and N. Harini, "A Hybrid Scheme for Detecting Fake Accounts in Facebook," *Int. J. Recent Technol. Eng. IJRTE*, vol. 7, no. 5S3, Feb. 2019.
- [13] F. S. Alsubaei, "Detection of Inappropriate Tweets Linked to Fake Accounts on Twitter," *Appl. Sci.*, vol. 13, no. 5, p. 3013, Feb. 2023, doi: 10.3390/app13053013.
- [14] M. Al Fatih Abil Fida, T. Ahmad, and M. Ntahobari, "Variance Threshold as Early Screening to Boruta Feature Selection for Intrusion Detection System," in *2021 13th International Conference on Information & Communication Technology and System (ICTS)*, Surabaya, Indonesia: IEEE, Oct. 2021, pp. 46–50. doi: 10.1109/ICTS52701.2021.9608852.
- [15] A. Parmar, R. Katariya, and V. Patel, "A Review on Random Forest: An Ensemble Classifier," in *International Conference on Intelligent Data Communication Technologies and Internet of Things (ICICI) 2018*, vol. 26, J. Hemanth, X. Fernando, P. Lafata, and Z. Baig, Eds., in *Lecture Notes on Data Engineering and Communications Technologies*, vol. 26, Cham: Springer International Publishing, 2019, pp. 758–763. doi: 10.1007/978-3-030-03146-6_86.

- [16] I. D. Mienye and Y. Sun, “A Survey of Ensemble Learning: Concepts, Algorithms, Applications, and Prospects,” *IEEE Access*, vol. 10, pp. 99129–99149, 2022, doi: 10.1109/ACCESS.2022.3207287.
- [17] J. Kufel *et al.*, “What Is Machine Learning, Artificial Neural Networks and Deep Learning?—Examples of Practical Applications in Medicine,” *Diagnostics*, vol. 13, no. 15, p. 2582, Aug. 2023, doi: 10.3390/diagnostics13152582.