

Analisa Pengaruh Pre-Processing Data Untuk Model Deteksi Akun Palsu Pada Media Sosial

Ni Nyoman Eny Perimawati¹⁾, Roy Rudolf Huizen²⁾, Dandy Pramana Hostiadi²⁾

Magister Sistem Informasi^{1,2,3)}

Institut Teknologi dan Bisnis STIKOM Bali

Denpasar, Indonesia

e-mail: enycaca@gmail.com¹⁾, dandy@stikom-bali.ac.id²⁾, roy@stikom-bali.ac.id³⁾

Abstrak

Media sosial telah menjadi bagian penting dalam interaksi Masyarakat yang dapat berdampak positif maupun negative. Salah satu media sosial yang banyak digunakan adalah media sosial X. Para pengguna media sosial dihadapkan pada masalah yang muncul diantaranya adalah adanya akun palsu yang dapat merusak reputasi pengguna dan menyebarkan informasi salah (hoax). Untuk mengatasi masalah tersebut penelitian ini mengumpulkan data melalui API Twitter sebanyak 1240 tweet dan membahas berbagai teknik pre-processing yang diperlukan, seperti pembersihan data, normalisasi, dan pelabelan. Hasil analisis menunjukkan bahwa model yang diterapkan setelah pre-processing mencapai akurasi 99,36%, dengan precision dan recall yang tinggi untuk kedua kelas akun. Sebaliknya, model tanpa pre-processing hanya mencapai akurasi 95,45%, menunjukkan penurunan dalam kemampuan deteksi. Hal ini memberikan gambaran pentingnya pre-processing data dalam pengembangan model machine learning untuk deteksi akun palsu. Penelitian ini diharapkan dapat memberikan kontribusi terhadap pendekatan yang lebih efektif dalam mengidentifikasi akun palsu di platform media sosial.

Kata Kunci: Pre-procesing, Akun Palsu, Media Sosial, Twitter, Logistic Regression

1. Pendahuluan

Media sosial telah menjadi bagian yang tidak terpisahkan dari kehidupan sehari-hari. Masyarakat menggunakan media sosial untuk berinteraksi, berbagi informasi, dan membangun komunitas. Hal ini memberikan dampak positif dan juga negatif seperti mulai banyaknya bermunculan akun palsu yang dapat merusak kepercayaan dan reputasi pengguna. Akun-akun ini sering digunakan untuk penyebaran informasi yang salah, penipuan, dan manipulasi opini publik pada masyarakat secara keseluruhan [1]. Oleh karena itu, penting untuk mengembangkan metode yang efektif untuk mendeteksi akun-akun tersebut dan menjaga integritas informasi yang beredar.

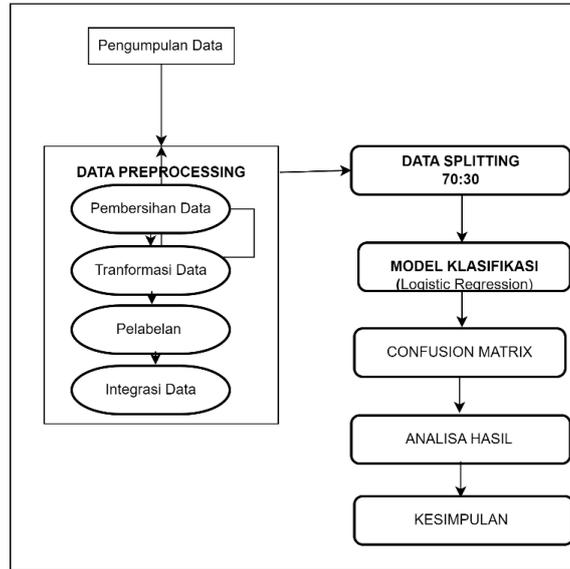
Deteksi akun palsu menjadi penting dalam era informasi saat ini di mana berita palsu dan hoaks dapat menyebar dengan cepat dan luas. Penelitian menunjukkan bahwa akun palsu dapat mempengaruhi perilaku pengguna, menciptakan ketidakpastian, dan menurunkan kualitas diskusi publik [2]. Logistic Regression merupakan salah satu algoritma yang banyak digunakan dalam masalah klasifikasi termasuk deteksi akun palsu. Algoritma ini berfungsi dengan memodelkan probabilitas suatu kelas berdasarkan fitur-fitur yang ada dalam data, sehingga sangat cocok untuk aplikasi di mana hasilnya bersifat biner [3].

Pre-processing data menjadi langkah penting dalam pengembangan model deteksi akun palsu. Proses ini meliputi berbagai teknik, seperti penghapusan noise, normalisasi, dan pengurangan dimensi, yang bertujuan untuk membersihkan dan menyiapkan data agar lebih siap untuk analisis [4]. Dengan menerapkan teknik pre-processing yang tepat diharapkan model Logistic Regression dapat mencapai tingkat akurasi yang lebih tinggi dalam mengidentifikasi akun palsu.

Penelitian ini bertujuan untuk menganalisis pengaruh teknik pre-processing data terhadap performa model deteksi akun palsu di media sosial X menggunakan Logistic Regression. Dengan mengeksplorasi berbagai metode pre-processing, penelitian ini diharapkan dapat memberikan wawasan yang mendalam mengenai bagaimana setiap langkah dapat mempengaruhi hasil akhir, serta kontribusinya terhadap pengembangan sistem deteksi yang lebih efektif dan efisien. Hasil dari penelitian ini diharapkan dapat membantu memperbaiki pendekatan yang ada dalam deteksi akun palsu pada platform media sosial X.

2. Metode Penelitian

Tahapan metode penelitian yang dilakukan dalam penelitian ini dimulai dari pengumpulan data dengan melakukan crawling data pada media sosial X, *Preprocessing* data meliputi pembersihan data, transformasi data, pelabelan, integrasi data. Setelah itu, membangun model klasifikasi dengan algoritma klasifikasi dengan Logistic Regression, mengevaluasi model klasifikasi dengan *confusion matrix*, Analisa hasil dan Kesimpulan hasil. Tahapan penelitian ditunjukkan pada gambar 3.



Gambar 3. Tahapn Penelitian

2.1. Pengumpulan data

Pengumpulan data dilakukan dari media sosial X dengan menggunakan API Twitter untuk mendapatkan informasi tentang akun, tweet dan interaksi pengguna. Adapun dari percobaan penggunaan API twitter yang dilakukan dengan menggunakan google colab dengan Python disimpan dalam dataset.csv dimana sekali proses penarikan mendapatkan 1240 tweet. Gambar 1 merupakan sampel data hasil Crawling data pada media sosial X yang disimpan pada file CRAWLING_DATA_TWITER(X).csv.

No	conversation_id_str	created_at	favorite_count	full_text	id_str	image_url	in_reply_to_screen_name	lang	location	quote_count	reply_count	retweet_count	tweet_url	user_id	username
1	1.84E+18	Mon Sep 08 10:10:18	0	@Mdy_Asm	1.84E+18		Mdy_Asm	in		0	0	0	https://x.com/1.10E+18/metek		
2	1.84E+18	Mon Sep 08 10:10:18	0	ð—€	1.84E+18	https://pbs.twimg.com/profile_pictures/normal/1.83E+18/intelkamresgto				0	0	0	https://x.com/1.83E+18/intelkamresgto		
3	1.84E+18	Mon Sep 08 10:10:18	1	KerahKe	1.84E+18		KerahKerii	in	Chiyoda-k	0	0	0	https://x.com/1.70E+18/topkoleksyen		
4	1.84E+18	Mon Sep 08 10:10:18	0	@kucingdi	1.84E+18		kucingdari	in		0	1	0	https://x.com/3.2E+09/moon_sera11		
5	1.84E+18	Mon Sep 08 10:10:18	1	@tjtrosoe	1.84E+18		tjtrosoe	in	Menteng,	0	0	0	https://x.com/1.61E+18/anegara862		
6	1.84E+18	Mon Sep 08 10:10:18	0	Up terus	1.84E+18			in		0	0	0	https://x.com/1.10E+18/metek		
7	1.84E+18	Mon Sep 08 10:10:18	2	@HusinSh	1.84E+18		HusinShih	in	Jauh dima	0	0	0	https://x.com/1.5E+08/Shakutogery		
8	1.84E+18	Mon Sep 08 10:10:18	0	Pak kenapa	1.84E+18			in	Malang	0	0	0	https://x.com/2.8E+08/WalinHuda		
9	1.84E+18	Mon Sep 08 10:10:18	8725	hari ini ad	1.84E+18			in		16	264	217	https://x.com/1.49E+18/SW_RaishaJKT48		

Gambar 1. Sampel Dataset hasil Crawling

Pada tabel 1 adalah atribut data yang didapatkan hasil penarikan Data dengan Twitter API.

Table 1. Atribut pada CRAWLING_DATA_TWITER(X).csv

No	Nama atribut	Deskripsi
1	conversation_id_str	identifikasi unik untuk sesi percakapan
2	created_at	Tanggal dan waktu saat tweet dibuat
3	favorite_count	Jumlah favorit yang diterima tweet
4	full_text	Teks lengkap dari tweet
5	id_str	ID unik untuk tweet
6	image_url	URL gambar yang terkait dengan tweet
7	in_reply_to_screen_name	Nama pengguna yang di-reply
8	lang	Bahasa yang digunakan dalam tweet

9	location	Lokasi pengguna saat membuat tweet
10	quote_count	Jumlah kutipan yang diterima tweet
11	reply_count	Jumlah balasan yang diterima tweet
12	retweet_count	Jumlah retweet yang diterima tweet
13	tweet_url	URL untuk tweet
14	user_id_str	ID unik untuk pengguna
15	username	Nama pengguna yang membuat tweet

2.2. Pre-Processing

Setelah data terkumpul tahap selanjutnya adalah melakukan pre-processing data. Pada tahapan ini terdiri akan dilakukan proses pembersihan data dengan menghapus data duplikat noise, karakter khusus, dan informasi yang tidak diperlukan dari dataset. Hal ini dilakukan untuk memastikan bahwa data yang digunakan dalam analisis bersih dan relevan [3]. Kemudian akan dilakukan transformasi data yaitu dengan mengubah format data yang sesuai dengan Analisa yang dilakukan meliputi proses normalisasi teks dan perubahan tipe data dan pemformatan data [4].

Pelabelan dilakukan secara manual dengan terlebih dahulu membuat akun fiktif sebanyak 100 akun yang nantinya dilabeli sebagai akun palsu sedangkan akun dari twitter akan dilabeli sebagai akun asli hal ini dilakukan untuk menjaga privasi data dan legalitas. Pelabelan dengan sistem ini pernah oleh Estée Van Der Walt and Jan Eloff (2018) dalam penelitiannya yang berjudul “ Using Machine Learning to Detect Fake Identities: Bots vs Humans” dimana dalam pelabelan peneliti membuat 15.000 akun fiktif dengan atribut yang dirancang untuk mencerminkan penipuan identitas. Akun asli yang dikumpulkan dari Twitter dianggap sebagai akun "manusia", sementara akun fiktif yang dibuat dianggap sebagai akun "palsu".

Integrasi data merupakan proses menyatukan data dari berbagai sumber kedalam satu database baru. Proses penggabungan ini dapat terjadi pada atribut yang sama atau melibatkan penambahan atribut yang hasilnya akan meningkatkan informasi lebih komprehensif. Pada tahap ini data crawling Twitter (X) yang sebelumnya telah dilabeli sebagai akun asli dengan akun fiktif yang dibuat secara manual dan telah dilabeli sebagai akun palsu akan digabungkan menjadi satu dataset Tunggal.

2.3. Data Spitting

Setelah proses pre-processing selesai data dibagi menjadi dua bagian dengan proporsi 70:30. Sebanyak 70% dari data digunakan untuk melatih model klasifikasi, sementara 30% sisanya digunakan sebagai data uji untuk mengevaluasi performa model. Pembagian ini bertujuan untuk memastikan bahwa model dapat generalisasi dengan baik pada data yang belum pernah dilihat sebelumnya (Kaur & Kumar, 2022).

2.4. Model Klasifikasi

Model klasifikasi yang digunakan pada penelitian ini algoritma Logistic Regression diterapkan pada data yang telah diproses. Logistic Regression dipilih karena kesederhanaannya dan efektivitasnya dalam menangani masalah klasifikasi biner [9]. Model dilatih dengan data pelatihan dan dioptimalkan untuk memaksimalkan akurasi dalam mengidentifikasi akun palsu

2.5. Confusion matrix

Dalam penelitian ini evaluasi kinerja model dilakukan dengan menggunakan confusion matrix. Confusion matrix memberikan gambaran visual tentang performa model, termasuk jumlah true positives, true negatives, false positives, dan false negatives. Confusion matrix yang digunakan adalah akurasi, presisi, recall, dan F1-score.

2.6. Analisa hasil

Hasil dari evaluasi model dianalisis untuk menentukan seberapa baik model dapat mendeteksi akun palsu di media sosial. Analisis ini mencakup diskusi tentang kekuatan dan kelemahan model, serta potensi perbaikan yang dapat dilakukan di masa mendatang.

2.7. Kesimpulan

Tahapan akhir dari penelitian adalah menarik kesimpulan berdasarkan hasil analisis. Kesimpulan mencakup ringkasan temuan utama, implikasi dari hasil penelitian, serta rekomendasi untuk penelitian lebih lanjut dan praktik di bidang deteksi akun palsu.

3. Hasil dan Pembahasan

3.1 Pre Processing Data

Pada tahap Preprocessing data akan di proses data sebanyak 1219 baris dengan 15 atribut pertama selanjutnya dilakukan eliminasi atribut conversation_id_str dan Creat_at dan penghapusan data duplikat dengan kriteria data yang akan dihapus adalah data yang memiliki atribut username dan user_id_str yang

sama, dan atribut timestamp (`created_at`) yang diidentik. Penghapusan data akan dilakukan menggunakan coding python dengan metode drop duplicate dengan kriteria khusus.

Tabel 2. Atribut setelah eliminasi

Nama atribut	Deskripsi
<code>favorite_count</code>	Jumlah favorit yang diterima tweet
<code>full_text</code>	Teks lengkap dari tweet
<code>id_str</code>	ID unik untuk tweet
<code>image_url</code>	URL gambar yang terkait dengan tweet
<code>in_reply_to_screen_name</code>	Nama pengguna yang di-reply
<code>lang</code>	Bahasa yang digunakan dalam tweet
<code>location</code>	Lokasi pengguna saat membuat tweet
<code>quote_count</code>	Jumlah kutipan yang diterima tweet
<code>reply_count</code>	Jumlah balasan yang diterima tweet
<code>retweet_count</code>	Jumlah retweet yang diterima tweet
<code>tweet_url</code>	URL untuk tweet
<code>user_id_str</code>	ID unik untuk pengguna
<code>username</code>	Nama pengguna yang membuat tweet

Jumlah data setelah dilakukan proses penghapusan duplikasi dari awalnya 1219 baris menjadi 423 baris. Data selanjutnya dilakukan transformasi tipe data dengan mengubah kolom `created_at` dari string ke tipe data tanggal. Selain itu untuk kolom `favorite_count`, `reply_count`, dan `retweet_count` akan dirubah dari string ke tipe data numerik. Normalisasi teks dilakukan dengan menggunakan `CountVectorizer` untuk mengubah teks tweet menjadi representasi numerik. Hasil dari vektorisasi teks ini digabungkan dengan fitur numerik seperti `favorite_count`, `reply_count`, dan `retweet_count`. Data yang telah diproses ini kemudian digunakan untuk pelatihan model.

Proses pelabelan dilakukan dengan cara manual dimana data hasil pembersihan sebanyak 423 baris akan dilabel 1 dimana ini mewakili akun asli. selanjutnya dibuat 100 akun fiktif (palsu) dengan kriteria informasi profil yang tidak konsisten, seperti nama pengguna dan nama lengkap yang tidak cocok atau penggunaan gambar profil yang tidak sesuai. Aktivitas akun juga berperan; akun yang memposting dengan sangat sering atau tidak pernah memposting sama sekali. Selain itu, keterlibatan sosial yang rendah, seperti rasio pengikut dan teman yang tidak wajar serta minimnya interaksi dengan pengguna lain, dapat menunjukkan bahwa akun tersebut tidak asli. Selanjutnya data tersebut disimpan dalam dataset `Fake_akun`. CSV. Akun palsu secara manual yang dirancang untuk menyerupai akun manusia dan Akun ini dilabeli dengan 0.

Setelah proses pelabelan kedua dataset selesai yaitu dataset akun yang telah dibersihnya dengan nama `cleaned_dataset_tweets.csv` dan data akun palsu `fake_dataset_tweets.csv` selesai maka akan dilakukan integrasi atau penggabungan keduanya. Dimana dari hasil penggabungan ini tercipta dataset baru bernama `merged_dataset.csv` dengan jumlah data sebanyak 523 baris.

3.2 Klasifikasi dan Evaluasi Model

3.2.1 Dengan Pre-Processing

Sebelum membangun model klasifikasi, data dibagi (`Splitting Data`) menjadi 70% data training dan 30% data testing dengan jumlah data training 366 baris dan data testing 157 baris data. Jumlah data masing-masing dapat dilihat pada tabel 3.

Tabel 3. Jumlah data training dan data testing

Data	Label 0(palsu)	Label 1(asli)
<i>Data training</i>	294	72
<i>Data testing</i>	129	28

Akurasi model mencapai 99.36%, dengan precision dan recall yang tinggi untuk kedua kelas. Hasil ini menunjukkan bahwa preprocessing yang dilakukan efektif dalam meningkatkan performa model. Tabel 4 menunjukkan performa model klasifikasi Logistic Regression dengan menggunakan dataset yang dilakukan proses pre-processing.

Tabel 4. Performa model klasifikasi Logistic Regression

	Precision	Recall	f1-score	support
Label (0)	0,97	1,00	0,98	28
Label (1)	1,00	0,99	1,00	129
accuracy			0,99	157
macro avg	0,98	1,00	0,99	157
weighted avg	0,99	0,99	0,99	157

3.2.2 Tanpa Pre-Processing

Pada pendekatan tanpa preprocessing dataset awal CRAWLING_DATA_TWITTER(X).csv tidak terfilter dari duplikat dan entri yang tidak relevan meskipun dilakukan penghapusan kolom conversation_id_str dan created_at. Pada proses ini hanya fitur numerik yang digunakan dalam model tanpa memasukkan informasi tekstual dari tweet. Ini mengurangi jumlah informasi yang tersedia bagi model untuk membuat prediksi sehingga model hanya memiliki 923 sampel untuk pelatihan dan 396 sampel untuk pengujian.

Akurasi model tanpa preprocessing adalah 95.45% meskipun masih tinggi hasil ini menunjukkan penurunan signifikan dibandingkan dengan model yang menggunakan preprocessing. Confusion matrix menunjukkan bahwa model salah mengklasifikasikan 17 akun asli sebagai palsu. Precision untuk kelas palsu (label 0) hanya 0.90, dengan recall 0.35, menunjukkan bahwa model kesulitan dalam mendeteksi akun palsu. Pada tabel 5 performa model klasifikasi Logistic Regression tanpa pre-processing.

Tabel 5. Performa model klasifikasi

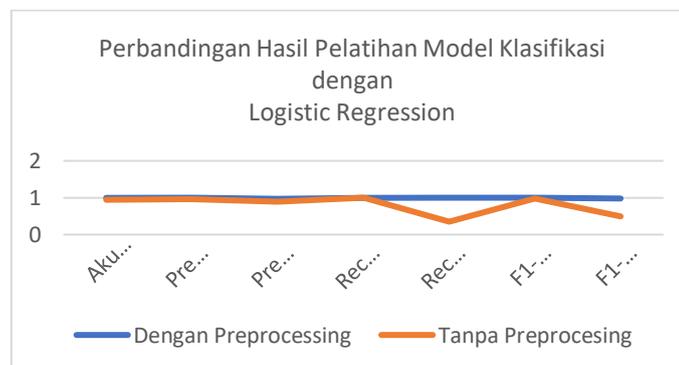
	Precision	Recall	f1-score	support
Label (0)	0,90	0,35	0,50	29
Label (1)	0,96	1,00	0,98	370
accuracy			0,95	395
macro avg	0,93	0,67	0,74	395
weighted avg	0,95	0,95	0,94	395

3.3 Evaluasi Model

Pada tabel 6 Menunjukan performa model Logistic Regression antara kondisi dengan proses pre-processing dan tanpa preprocessing. Dalam kondisi dengan preprocessing confusion matrix menunjukkan hasil yang sangat baik dengan 28 prediksi benar untuk kelas 0 dan 128 untuk kelas 1, serta hanya satu kesalahan dalam memprediksi kelas 1 sedangkan model tanpa preprocessing menunjukkan penurunan performa yang signifikan. Confusion matrix menunjukkan bahwa model ini salah mengklasifikasikan 17 sampel kelas 0 sebagai kelas 1, dengan 9 prediksi benar untuk kelas 0 dan 369 untuk kelas 1.

Tabel 6. Perbandingan Confusion Matrix

	TP	TN	FP	FN
Dengan Preprocessing	28	128	1	0
Tanpa Preprocessing	9	369	1	17



Grafik 1. Perbandingan hasil pelatihan model

Grafik 1 adalah perbandingan Tingkat akurasi model Logistic Regression dengan proses preprocessing dan tanpa preprocessing. Untuk model dengan preprocessing precision 0,97 dan recall 1,00 untuk kelas 0, serta precision 1,00 dan recall 0,99 untuk kelas 1. Akurasi keseluruhan model ini mencapai 99,36%, mencerminkan kemampuan yang sangat baik dalam mendeteksi kedua kelas. Sedangkan model tanpa preprocessing precision untuk kelas 0 hanya mencapai 0,90 dengan recall yang rendah yaitu 0,35, menunjukkan bahwa model kesulitan dalam mendeteksi kelas 0. Meskipun kelas 1 memiliki precision 0,96 dan recall 1,00, akurasi keseluruhan model ini hanya 95,45%. Hal ini menandakan bahwa meskipun model masih dapat bekerja dengan baik ketidakmampuan untuk mendeteksi kelas 0 menggambarkan adanya bias yang perlu diperbaiki.

4. Kesimpulan

Dalam penelitian ini proses deteksi akun palsu di media sosial X dimulai dengan pengumpulan data dari API Twitter yang menghasilkan dataset awal yang terdiri dari 1219 baris. Selanjutnya model dilatih dengan menggunakan proses pre-processing data dan tidak dilakukan proses pre-processing data. Pada model yang menggunakan pre-processing menunjukkan hasil yang bagus yaitu dengan tingkat akurasi mencapai 99,36%. Precision untuk kelas palsu (label 0) hanya 0,97, dengan recall 1,00 dan F1-Score 0,98 sedangkan Precision untuk kelas Asli (label 1) 1,00 dengan recall 0,99 dan F1-Score 1,00. Sedangkan model yang dilatih tanpa preprocessing hanya mencapai akurasi 95,45%, dengan performa yang jauh lebih rendah dalam mendeteksi akun palsu. Hal ini menunjukkan bahwa proses pre-processing data sangat penting dalam pengembangan model machine learning. Untuk penelitian selanjutnya disarankan untuk mengeksplorasi teknik preprocessing tambahan dan algoritma klasifikasi lainnya serta memperluas dataset agar model dapat lebih generalis dan efektif dalam mendeteksi akun palsu di berbagai konteks media sosial.

Daftar Pustaka

- [1] Himawan, A., Santoso, E., & Nugroho, A. (2020). "Detection of Fake Accounts on Social Media Using Machine Learning Algorithms." *Journal of Computer Science and Technology*, 35(2), 123-135.
 - [2] Ferrara, E., Varol, O., Davis, C., Menczer, F., & Flammini, A. (2016). "The rise of fake news on social media." *Communications of the ACM*, 59(1), 84-92.
 - [3] Zhang, Y., & Wang, Y. (2019). "Machine Learning Techniques for Fake News Detection: A Survey." *IEEE Access*, 7, 103797-.
 - [4] 103812Kumar, A., & Singh, S. (2020). "Data Cleaning and Preprocessing for Machine Learning: A Review." *International Journal of Computer Applications*, 975, 8887.
 - [5] Al-Azzeh, M., & Al-Rousan, M. (2021). "Artificial Intelligence in Social Media: A Review of Fake News Detection Techniques." *Journal of King Saud University - Computer and Information Sciences*.
 - [6] Gupta, A., & Kumar, A. (2021). "A Comprehensive Review on Fake News Detection Techniques." *International Journal of Information Technology*, 13(3), 1231-1240.
 - [7] Sarker, I. H., & Ghosh, K. (2020). "Detection of Fake News on Social Media: A Machine Learning Approach." *International Journal of Computer Applications*, 975, 8888.
 - [8] Shahi, G. K., & Singh, P. (2021). "A Study on Fake News Detection and its Impact on Society." *International Journal of Information Technology and Computer Science (IJITCS)*, 13(5), 1-10.
 - [9] Ghanem, A., & El-Shafee, M. (2020). "Analyzing the Impact of Preprocessing Techniques on Twitter Sentiment Analysis." *Journal of Computer Science and Technology*, 35(4), 789-797.
 - [10] Lazer, D. M. J., Baum, M. A., Benkler, Y., et al. (2018). "The Science of Fake News: Addressing Fake News in a Post-Truth World." *Science*, 359(6380), 1094-1096.
 - [11] Estée Van Der Walt (2018). *Using Machine Learning to Detect Fake Identities: Bots vs Humans*
 - [12] Kumar, A., & Singh, S. (2021). "Data Cleaning and Preprocessing for Machine Learning: A Review." *International Journal of Computer Applications*, 975, 8887.
 - [13] Kaur, H., & Kumar, V. (2022). "Deep Learning Techniques for Fake News Detection: A Review."
-