ANALISIS KESADARAN MASYARAKAT TERHADAP BAHAYA INTERNET PHISHING MENGGUNAKAN K-MEANS CLUSTERING

Indrawan Ady Saputro[1*], Lilik Sugiarto[2], Febrianta Surya Nugraha[3]
[1,2,3]Program Studi Informatika, STMIK AMIKOM
indrawanadysaputro@gmail.com[1]

Submitted February 17, 2024; Revised July 10, 2024; Accepted July 17, 2024

Abstrak

Phishing adalah bentuk kejahatan siber yang menipu pengguna dengan pesan palsu yang mengandung tautan atau lampiran berbahaya, menyebabkan kerugian finansial, pencurian identitas, atau kerusakan sistem. Kesadaran keamanan yang tinggi diperlukan untuk menghindari phishing. Penelitian ini bertujuan menganalisis kesadaran masyarakat terhadap bahaya phishing di internet menggunakan metode k-means clustering. Data dikumpulkan melalui survei online terhadap 30 responden dari berbagai latar belakang dan diproses menggunakan algoritma k-means clustering untuk menghasilkan kelompok berdasarkan tingkat kesadaran. Hasil menunjukkan bahwa mayoritas responden memiliki kesadaran keamanan sedang terhadap phishing: 17% memiliki kesadaran tinggi, 57% sedang, dan 26% rendah. Faktor-faktor yang mempengaruhi kesadaran termasuk usia, pendidikan, pekerjaan, frekuensi penggunaan internet, dan sumber informasi. Evaluasi klaster menggunakan Silhouette Coefficient menunjukkan bahwa pengelompokan dengan 3 klaster (nilai 0,44) adalah yang paling efektif, dibandingkan dengan 2 klaster (nilai 0,37) dan 4 klaster (nilai 0,36). Penelitian ini memberikan wawasan tentang variasi kesadaran keamanan dan pentingnya peningkatan literasi digital untuk melindungi masyarakat dari phishing.

Kata Kunci : phishing, kesadaran keamanan, k-means clustering, internet, survei online

Abstract

Phishing is a form of cybercrime that tricks users with fake messages containing malicious links or attachments, causing financial loss, identity theft, or system damage. High security awareness is necessary to avoid phishing. This research aims to analyze public awareness of the dangers of phishing on the internet using the k-means clustering method. The data was collected through an online survey of 30 respondents from various backgrounds and processed using the k-means clustering algorithm to generate groups based on awareness levels. The revealed that the majority of respondents had moderate security awareness of phishing: 17% had high awareness, 57% moderate, and 26% low. Factors affecting the awareness include age, education, occupation, internet use frequency, and information sources. Cluster evaluation using Silhouette Coefficient indicated that grouping with 3 clusters (value 0.44) was the most effective, compared to 2 clusters (value 0.37) and 4 clusters (value 0.36). This research provides insight into the variation in security awareness and the importance of increasing digital literacy to protect the public from phishing.

Keywords : *phishing, security awareness, k-means clustering, internet, online surveys*

1. PENDAHULUAN

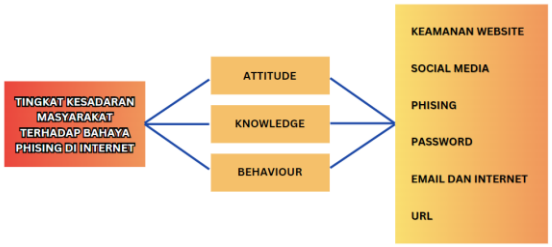
Internet dapat digunakan untuk berbagai keperluan, seperti mencari informasi, berkomunikasi, berbelanja, bermain, belajar, bekerja, dan lain-lain [1][2]. Menurut laporan dari Penyelenggara Jasa Internet Indonesia (APJII), jumlah pengguna internet di Indonesia mencapai 215,63 juta orang dalam rentang waktu 2022-2023 [3]. Namun, internet juga membawa berbagai risiko dan ancaman bagi penggunanya. Salah satu bahaya yang muncul perkembangan teknologi adalah web phising, yaitu teknik manipulasi psikologis yang digunakan oleh penyerang agar mendapatkan akses yang ilegal dari pengguna media sosial [4],[5]. Phishing adalah salah satu bentuk serangan siber untuk mencuri informasi pribadi korban

[6],[7]. Kejahatan phising sebenarnya sudah banyak bermunculan disekitar akan tetapi banyak dari masyarakat tidak mengetahui akan bahaya tersebut [8]. Cara kerja phising yaitu melakukan ancaman atau melakukan penjebakan seseorang dengan memancing korban untuk memasuki situs atau web tertentu yang dibagikan menggunakan tautan link [9].

Phishing adalah salah satu bentuk kejahatan siber yang mencoba menipu pengguna internet dengan mengirimkan pesan palsu yang mengandung tautan atau lampiran berbahaya. Pesan palsu ini biasanya meniru atau menyamar sebagai pesan resmi dari lembaga atau organisasi tertentu, seperti bank, perusahaan, pemerintah, atau media sosial. Tujuan dari phishing adalah untuk memperoleh informasi pribadi atau rahasia pengguna internet, seperti nama, alamat, nomor telepon, email, kata sandi. Informasi ini kemudian dapat digunakan oleh pelaku phishing untuk melakukan tindakan kriminal, seperti pencurian uang, penipuan, atau perusakan sistem [9]. Salah satu bentuk serangan yang dilakukan oleh penjahat dunia maya adalah dengan menempatkan tautan palsu di akun media sosial yang menggoda dengan ajakan atau iklan yang simpel. Dengan cara tersebut, penyerang dapat memperoleh informasi pengguna dan memanfaatkannya untuk mendapatkan keuntungan, seperti mengakses uang dari rekening pengguna atau melakukan transaksi online menggunakan rekening tersebut [10].

Phishing dapat menyebabkan kerugian finansial, pencurian identitas, atau kerusakan sistem bagi pengguna internet. Oleh karena itu, pengguna internet perlu memiliki kesadaran keamanan yang tinggi untuk menghindari phishing [11],[12]. Kesadaran keamanan adalah tingkat pemahaman dan kewaspadaan pengguna internet terhadap potensi bahaya dan ancaman yang ada di internet, serta cara-cara untuk mencegah dan mengatasinya. Faktor dalam kesadaran keamanan antara lain : pengetahuan, sikap, dan perilaku pengguna internet terhadap phishing.

Kesadaran masyarakat terhadap risiko phishing di internet dievaluasi berdasarkan pengetahuan (knowledge), sikap (attitude), dan perilaku (behavior). Area yang tercakup dalam konsep pengukuran ketiga dimensi tersebut meliputi pemahaman tentang keamanan situs web, phishing, media sosial, spam/virus, URL, serta pengelolaan username dan password. Konsep pengukuran ini tersaji pada gambar 1.



(Sumber: peneliti, 2023)

Gambar 1. Pengukuran Dimensi KAB

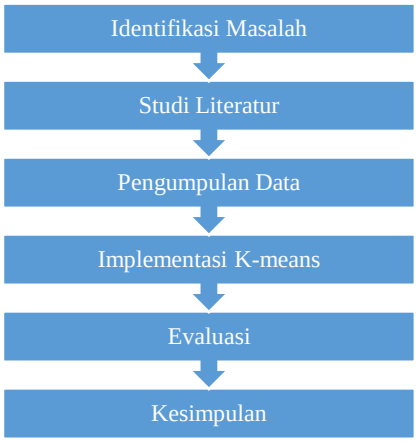
Metode data mining atau pengolahan data untuk mengelompokan data berdasarkan atribut atau karakteristik yang sama. Proses yang terjadi dalam algoritma k-means adalah menggunakan serangkaian kumpulan awal, setiap titik akan ditempatkan dalam salah satu dari kumpulan tersebut [13]. Kemudian, setiap pusat gugus akan digantikan dengan titik rata-rata di dalam klaster mereka masing-masing. Langkah-langkah ini sederhana dan diulang hingga mencapai konvergensi. Setiap titik akan ditempatkan ke dalam klaster yang memiliki jarak Euclidean terdekat dengan titik tersebut [14],[15], [16].

Penelitian tentang kesadaran keamanan terhadap phishing masih jarang dilakukan di Indonesia, terutama di kalangan masyarakat umum. Sementara itu, kelompok masyarakat umum rentan menjadi korban phishing karena kemungkinan kurangnya pengetahuan dan keterampilan yang memadai terkait

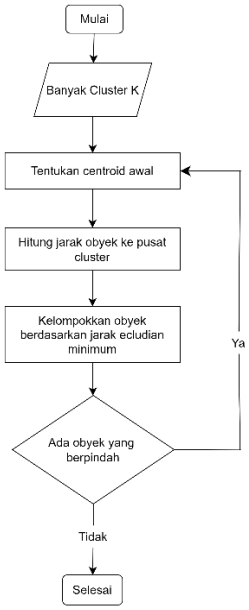
keamanan internet. Oleh karena itu, tujuan dari penelitian ini adalah untuk menganalisis tingkat kesadaran keamanan masyarakat terhadap ancaman phishing di internet dengan menerapkan metode pengelompokan k-means clustering. Penelitian ini berbeda dari penelitian terdahulu dengan fokus khusus pada masyarakat umum dan penggunaan metode analisis yang inovatif untuk mengidentifikasi tingkat kesadaran terhadap phishing.

2. METODE PENELITIAN

Metode analisis kesadaran masyarakat terhadap phishing melibatkan identifikasi masalah, studi literatur, pengumpulan data, penerapan metode K-Means, dan penarikan kesimpulan. Penelitian dimulai dengan menentukan konteks dan relevansi masalah, diikuti dengan tinjauan literatur untuk kerangka teoritis. Data dikumpulkan untuk mengeksplorasi kesadaran masyarakat terhadap phishing, kemudian diterapkan metode K-Means untuk mengelompokkan tingkat kesadaran. Evaluasi dilakukan menggunakan uji reliabilitas data dan Silhouette Coefficient. Hasil analisis dan evaluasi kemudian diakhiri dengan penarikan kesimpulan. Berikut alur penelitian pada penelitian ini tersaji pada gambar 2.



Gambar 2. Alur Penelitian



Gambar 3. Flowchart K-Means Clustering

Pada gambar 3 diatas terdapat flowchart K-Means clustering. Berikut adalah langkah-langkah penelitian untuk analisis tingkat kesadaran masyarakat menggunakan algoritma K-Means:

1. Menentukan tujuan dari Jumlah Cluster (K):
2. Pilih jumlah cluster yang diinginkan sesuai dengan karakteristik data atau tujuan analisis. K merupakan jumlah kelompok yang akan dibentuk.
3. Inisialisasi Centroid:
4. Buat K centroid acak menjadi titik awal pembentukan cluster.
5. Hitung Jarak ke Centroid
6. Tentukan Cluster : memilih centroid yang memiliki jarak terdekat. Data akan ditempatkan dalam cluster yang sesuai dengan centroid terdekat.
7. Perbarui Posisi Centroid : Tentukan lokasi centroid yang baru dengan mengkalkulasikan nilai rata-rata dari data yang tergabung dalam setiap centroid.
8. Periksa Konvergensi: Apabila posisi centroid baru masih berbeda dengan posisi centroid sebelumnya, langkah 3 akan diulang hingga konvergensi terjadi.

9. Tentukan Jumlah dan Nilai Kembali: penentuan dari jumlah cluster yang dibentuk dan nilai K. Hal ini dilakukan berulang kali untuk mencari kombinasi yang paling optimal.
10. Hitung Jarak dengan Persamaan Euclidean: Lakukan perhitungan jarak antara setiap input dan centroid menggunakan rumus Euclidean, di mana *D*, dengan i=1 hingga n, dijabarkan dalam persamaan (1).

$$(x, y) = |x - y| = \sqrt{\sum} (xi - yi)^2 / n \quad \text{.....(1)}$$

3. HASIL DAN PEMBAHASAN

Dalam penelitian ini, metode k-means diterapkan pada analisis kesadaran informasi tentang bahaya phishing di internet menggunakan sampel sebanyak 30 responden. Data usia responden yang mengisi kuesioner terdistribusi seperti yang tercantum dalam Tabel 1.

Tabel 1.Sebaran Usia Responden

Usia	Jumlah Responden
<18 Tahun	2
25-34 Tahun	15
35-44 Tahun	5
45-54 Tahun	3
>55 Tahun	5

(Sumber : data peneliti, 2023)

Data yang diterapkan berasal dari pengumpulan informasi melalui formulir google, dengan total 30 responden yang digunakan sebagai dataset yang akan dianalisis menggunakan metode k-means clustering. Berikut pada tabel 2 merupakan hasil kuesioner tingkat kesadaran masyarakat terhadap phising di Internet.

Tabel 2. Hasil Kuisioner

Responden	Knowledge	Attitude	Behaviour
1	3,4	2,2	3,2
2	4,4	3,4	1,8
3	4	3,4	3,8
4	3	3,4	2,4
5	2,4	3,8	2,6
6	5	5	5
7	4,8	5	5
...	....	...	....
30	4,2	3	2,2

(Sumber: data peneliti, 2023)

Hasil analisis deskriptif dari responden 30 masyarakat terdapat nilai rata-rata serta standar deviasi untuk semua variabel yang dapat dilihat pada tabel 3.

Tabel 3. Hasil Analisis Deskriptif

Descriptive Statistics

	N	Min	Max	Mean	Std. Deviation	Variance
1	30	1	4	3.36	0.76489	0.585
2	30	1	4	3.16	0.79148	0.626
3	30	1	4	3.30	0.83666	0.700
4	30	1	4	3.00	0.69481	0.483
5	30	1	4	3.10	0.88474	0.783
6	30	1	4	3.03	0.76489	0.585
7	30	1	4	3.20	0.84690	0.717
8	30	1	4	3.30	0.70221	0.493
9	30	2	4	2.96	0.76489	0.585
10	30	1	4	3.16	0.64772	0.420
11	30	1	4	3.30	0.83666	0.700
12	30	1	4	3.13	0.68145	0.464
13	30	1	4	3.13	0.93710	0.878
14	30	2	4	3.40	0.67466	0.455
15	30	1	4	3.10	0.92289	0.852
Valid N (listwise)	30					

(Sumber : data peneliti, 2023)

Dari hasil uji reliabilitas pada kuesioner, dapat disimpulkan bahwa nilai Cronbach's alpha untuk variabel ini berada di atas nilai dasar, dengan perbandingan 0,960 yang lebih tinggi daripada 0,403. Hal ini memberikan keyakinan bahwa semua

pernyataan yang terdapat dalam kuesioner dapat diandalkan. Informasi terperinci mengenai hasil uji reliabilitas dapat ditemukan dalam Tabel 4.

Tabel 4. Hasil Uji Reliabilitas

Reliability Statistics

Cronbach's Alpha	N of Items
0.960	15

(sumber data peneliti, 2023)

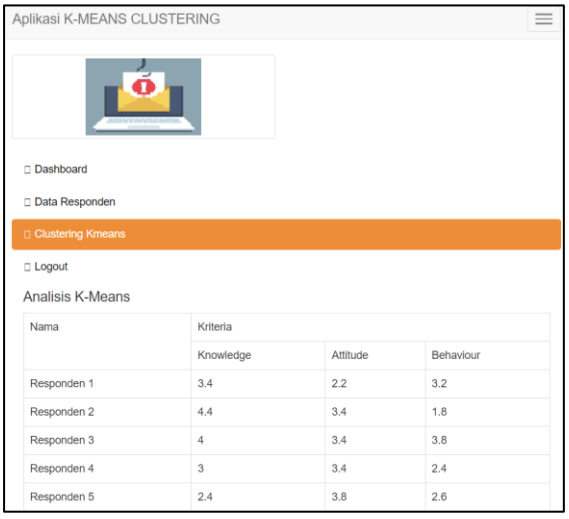
Perhitungan k-means dimulai dengan mencari pusat cluster. Kemudian melakukan perhitungan jarak antar centroid, melakukan iterasi agar mendapatkan pengelompokan sehingga mendapatkan akhir dari pusat cluster setelah itu mendapatkan hasil dari pengelompokan.

Tabel 5. Hasil Pengelompokan dengan K-Means Clustering Klaster 2,3,4

Cluster Membership			
Case Number	Cluster 2	Cluster 3	Cluster 4
1	2	1	3
2	2	2	2
3	2	2	1
4	2	2	2
5	2	2	2
6	1	3	4
7	1	3	4
8	2	1	3
9	2	1	3
10	2	1	3
11	1	3	4
12	1	3	4
13	1	3	4
14	2	2	2
15	2	2	2
16	2	2	2
17	2	2	2
18	2	2	2
19	2	2	2
20	2	2	1
21	1	3	4
22	2	2	1
23	2	2	2
24	2	2	1
25	2	2	2
26	1	3	1
27	2	1	3
28	1	3	1
29	2	2	1
30	2	2	2

(Sumber: data peneliti, 2023)

Dari hasil pengelompokan menggunakan klaster 2,3 dan 4 dengan metode k-means clustering melalui aplikasi SPSS, didapatkan hasil bahwa pada klaster 2 yaitu 8 responden (27%) memiliki pemahaman tinggi dan 22 responden (73%) memiliki pemahaman rendah. Hasil klaster 3 sebanyak lima responden (17%) memiliki tingkat pemahaman yang tinggi, 17 responden (57%) tergolong dalam kategori pemahaman sedang, sementara delapan responden (26%) masuk dalam kategori pemahaman rendah.. Sedangkan pada klaster 4 didapatkan hasil pemahaman tinggi 7 responden (23%), sedang 12 responden (40%), pemahaman rendah 5 responden (17%) dan pemhaman sangat rendah 6 responden (20%). Analisis ini menunjukkan variasi dalam pemahaman responden terhadap materi yang disajikan. Perbedaan ini dapat memberikan wawasan yang berharga dalam mengembangkan strategi pembelajaran yang lebih efektif di masa depan. Selain menggunakan SPSS juga dilakukan pembuatan aplikasi web menggunakan bahasa pemrograman PHP. Tampilan pada aplikasi web dapat dilihat pada gambar 4.

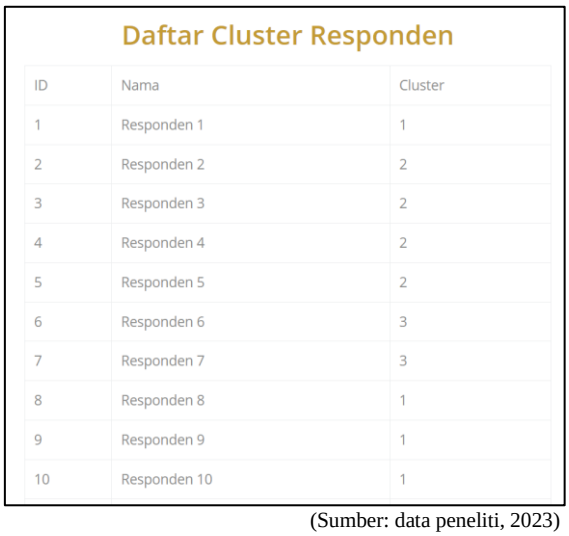


(sumber data peneliti, 2023)

Gambar 4. Tampilan Pada Aplikasi Web

Pada penelitian ini, pada iterasi ketiga dan keempat, tidak terjadi perubahan. Kelompok pertama memiliki 5 anggota,

kelompok kedua memiliki 17 anggota, dan kelompok ketiga memiliki 8 anggota. Hasil pengelompokan metode k-means clustering melalui aplikasi web yang telah dibuat menunjukkan kesamaan dengan hasil yang diperoleh dari aplikasi SPSS. Ditemukan bahwa 5 dari responden (sebesar 17%) memiliki tingkat pemahaman tinggi, 17 responden (57%) berada dalam kategori pemahaman sedang, dan 8 responden (26%) berada dalam kategori pemahaman rendah. Hal ini mengindikasikan bahwa mayoritas kesadaran masyarakat berada dalam kategori sedang. Rincian hasil pengelompokan dapat dilihat dalam Gambar 5.



Daftar Cluster Responden

ID	Nama	Cluster
1	Responden 1	1
2	Responden 2	2
3	Responden 3	2
4	Responden 4	2
5	Responden 5	2
6	Responden 6	3
7	Responden 7	3
8	Responden 8	1
9	Responden 9	1
10	Responden 10	1

(Sumber: data peneliti, 2023)

Gambar 5. Hasil Pengelompokan yang Diperoleh dari Sistem Website

Tahap selanjutnya setelah dilakukan implementasi kmeans pada tiga klaster, yaitu klaster 2, 3, dan 4, kemudian melakukan evaluasi kualitas masing-masing klaster diuji menggunakan *silhouette coefficient*. Berikut hasil evaluasi menggunakan *silhouette coefficient* tersaji pada tabel 6.

Tabel 6. Hasil evaluasi dengan *silhouette coefficient*

Klaster	Nilai *Silhouette Coefficient*
2	0,37
3	0,44
4	0,36

(Sumber: data peneliti, 2023)

Berdasarkan nilai *Silhouette Coefficient* untuk masing-masing jumlah klaster, pengelompokan dengan 3 klaster menunjukkan hasil yang paling baik, dengan nilai tertinggi sebesar 0,44, yang mengindikasikan struktur klaster yang lebih jelas dan terdefinisi. Sebaliknya, klaster 2 memiliki nilai 0,37, menunjukkan kualitas clustering yang cukup baik namun masih bisa diperbaiki, sedangkan klaster 4, dengan nilai 0,36, menunjukkan kualitas yang sedikit lebih rendah dibandingkan dengan klaster 2 dan 3. Dengan demikian, pengelompokan dengan 3 klaster adalah yang paling efektif, memberikan hasil clustering yang optimal dibandingkan dengan 2 atau 4 klaster.

Berbagai faktor dapat memengaruhi hasil ini, di antaranya adalah ketidaksadaran akan ciri-ciri phishing, kurangnya kehati-hatian dalam menjaga kerahasiaan password, dan kecenderungan membuka tautan email tanpa pertimbangan yang matang. Pemahaman akan pentingnya keamanan data pribadi memegang peranan penting dalam menjaga keamanan digital masyarakat. Oleh karena itu, peningkatan literasi digital dan kesadaran tentang keamanan data menjadi sangat penting untuk membantu masyarakat menjalankan aktivitas online dengan lebih bijaksana. Dengan demikian, upaya untuk meningkatkan pemahaman tentang keamanan digital di kalangan masyarakat perlu ditingkatkan secara keseluruhan.

4. SIMPULAN

Penelitian ini menunjukkan bahwa sebagian besar responden memiliki kesadaran keamanan terhadap phishing, dengan 17% memiliki kesadaran tinggi, 57% sedang, dan 26% rendah. Faktor-faktor yang mempengaruhi kesadaran ini meliputi usia, pendidikan, pekerjaan, frekuensi penggunaan internet, dan sumber informasi keamanan. Evaluasi klaster mengungkapkan bahwa pengelompokan dengan 3 klaster adalah yang paling efektif,

dengan nilai Silhouette Coefficient tertinggi 0,44, menandakan struktur klaster yang paling jelas. Sebaliknya, klaster dengan 2 dan 4 klaster memiliki nilai lebih rendah, yaitu 0,37 dan 0,36. Klaster pertama, yang menunjukkan kesadaran tinggi, terdiri dari individu berusia 18-25 tahun, memiliki pendidikan tinggi, bekerja di sektor teknologi, menggunakan internet lebih dari 4 jam sehari, dan memperoleh informasi dari sumber terpercaya. Klaster kedua, dengan kesadaran sedang, mencakup kelompok yang lebih beragam dalam hal usia, pendidikan, pekerjaan, frekuensi penggunaan internet, dan sumber informasi. Klaster ketiga, yang memiliki kesadaran rendah, sering terdiri dari responden berusia 26-35 tahun, berpendidikan menengah, bekerja di sektor non-teknologi, menggunakan internet kurang dari 4 jam sehari, dan mendapatkan informasi dari media sosial. Hal ini menunjukkan bahwa berbagai faktor tersebut secara bersamaan mempengaruhi tingkat kesadaran terhadap keamanan phishing.

DAFTAR PUSTAKA

[1] L. Yana Siregar and M. Irwan Padli Nasution, "Perkembangan Teknologi Informasi Terhadap Peningkatan Bisnis Online," *HIRARKI J. Ilm. Manaj. dan Bisnis*, vol. 2, no. 1, pp. 71–75, 2020.

[2] P. P. A. P. Batmetan, John Reimon, Morisa F. Lumingkewas, Claudia Tumuyu, "Analisis Perilaku Keamanan Informasi Pengguna Sosmed Dikalangan Generasi Milenial," *Prodi Pendidik. Teknol. Inf. dan Komunikasi, Univ. Negeri Manad. Tondano. 95318*, 2019.

[3] APJII, "Hasil Survei Penetrasi dan Perilaku Pengguna Internet," *Asosiasi Penyelenggara Jasa Internet Indonesia*, 2022. https://apjii.or.id/survei (accessed Feb. 07, 2023).

[4] A. Muhariya, I. Riadi, and Y. Prayudi, "Cyberbullying Analysis on Instagram Using K-Means Clustering," *JUITA J. Inform.*, vol. 10, no. 2, p. 261, 2022, doi: 10.30595/juita.v10i2.14490.

[5] I. Riadi, Herman, and I. A. Rafiq, "Mobile Forensic Investigation of Fake News Cases on Instagram Applications with Digital Forensics Research Workshop Framework," *Int. J. Artif. Intell. Res.*, vol. 6, no. 2, 2022, doi: 10.29099/ijair.v6i2.311.

[6] D. Wahyudi, M. Niswar, A. Ais, and P. Alimuddin, "Website Phising Detection Application Using Support Vector Machine (Svm)," *J. Inf. Technol. Its Util.*, vol. 5, no. 2, p. 2022, 2022.

[7] B. P. Zen, R. A. G. Gultom, and A. H. S. Reksoprodjo, "Analisis Security Assessment Menggunakan Metode Penetration Testing dalam Menjaga Kapabilitas Keamanan Teknologi Informasi Pertahanan Negara," *J. Teknol. Penginderaan*, vol. 2, no. 1, pp. 105–122, 2020.

[8] A. S. Gulo, S. Lasmadi, and K. Nawawi, "Cyber Crime dalam Bentuk Phising Berdasarkan Undang-Undang Informasi dan Transaksi Elektronik," *PAMPAS J. Crim. Law*, vol. 1, no. 2, pp. 68–81, 2021, doi: 10.22437/pampas.v1i2.9574.

[9] M. Y. DM, Addermi, and J. Lim, "Kejahatan Phising dalam Dunia Cyber Crime dan Sistem Hukum di Indonesia," *J. Pendidik. dan Konseling*, vol. 4, p. 8022, 2022.

[10] A. Ramadhan, M. A. Alhafidh, and M. D. Firmansyah, "Penyebaran Link Phising Kuota Kemendikbud Terhadap Kesadaran Informasi Pribadi Di Kalangan Mahasiswa UNINUS," *Kampret J.*, vol. 1, no. 1, pp. 11–15, 2022, doi: 10.35335/kampret.v1i1.9.

[11] V. F. Putra Y, "Modus Operandi

Tindak Pidana Phising Menurut UU ITE," *Jurist-Diction*, vol. 4, no. 6, p. 2525, 2021, doi: 10.20473/jd.v4i6.31857.

[12] A. Safi and S. Singh, "A systematic literature review on phishing website detection techniques," *J. King Saud Univ. - Comput. Inf. Sci.*, vol. 35, no. 2, pp. 590–611, 2023, doi: https://doi.org/10.1016/j.jksuci.2023.01.004.

[13] Z. R. Alfy, A. D. Baihaqie, and Z. F. A'ini, "Analisis Cluster K-Means pada Indikator Indeks Pembangunan Teknologi, Informasi, dan Komunikasi," *STRING (Satuan Tulisan Ris. dan Inov. Teknol.*, vol. 8, no. 2, pp. 183–188, 2023.

[14] M. Habibi and P. W. Cahyo, "Clustering User Characteristics Based on the influence of Hashtags on the Instagram Platform," *IJCCS (Indonesian J. Comput. Cybern. Syst.*, vol. 13, no. 4, p. 399, 2019, doi: 10.22146/ijccs.50574.

[15] S. Harifi, M. Khalilian, J. Mohammadzadeh, and S. Ebrahimnejad, "Using metaheuristic algorithms to improve k-means clustering: A comparative study," *Rev. d'Intelligence Artif.*, vol. 34, no. 3, pp. 297–305, 2020, doi: 10.18280/ria.340307.

[16] B. Harahap, "Penerapan Algoritma K-Means Untuk Menentukan Bahan Bangunan Laris (Studi Kasus Pada UD. Toko Bangunan YD Indarung)," *Reg. Dev. Ind. Heal. Sci. Technol. Art Life*, pp. 394–403, 2019, [Online]. Available: https://ptki.ac.id/jurnal/index.php/readystar/article/view/82