# IMPLEMENTASI ALGORITMA *LOGISTIC REGRESSION* UNTUK DETEKSI *LINK PHISHING*

**SKRIPSI**

Diajukan untuk memenuhi syarat dalam memperoleh gelar Sarjana di Program Studi Pendidikan Sistem dan Teknologi Informasi Universitas Pendidikan Indonesia Kampus Purwakarta

**Oleh:**

**Hilya Anbiyani Fitri Muhyidin**

**NIM 2005517**

**PROGRAM STUDI S1**

**PENDIDIKAN SISTEM DAN TEKNOLOGI INFORMASI**

**UNIVERSITAS PENDIDIKAN INDONESIA**

**KAMPUS PURWAKARTA**

**2024**

# IMPLEMENTASI ALGORITMA *LOGISTIC REGRESSION* UNTUK DETEKSI *LINK PHISHING*

Oleh

**Hilya Anbiyani Fitri Muhyidin**

Sebuah skripsi yang diajukan untuk memenuhi salah satu syarat memperoleh gelar Sarjana Pendidikan Sistem dan Teknologi Informasi Kampus Daerah Purwakarta



Universitas Pendidikan Indonesia

Juni 2024

**LEMBAR PENGESAHAN**

**HILYA ANBIYANI FITRI MUHYIDIN**

**IMPLEMENTASI ALGORITMA *LOGISTIC REGRESSION***
**UNTUK DETEKSI *LINK PHISHING***

Disetujui dan disahkan oleh pembimbing:

**Pembimbing I:**

**Dr. H. Suprih Widodo, S.Si., M.T.**
**NIP. 19801217005021007**

**Pembimbing II:**

**Liptia Venica, S.T., M.T.**
**NIP. 920210919941203201**

Mengetahui,
**Ketua Program Studi Pendidikan Sistem dan Teknologi Informasi**
**Kampus Daerah Purwakarta**

**Ir. Nuur Wachid Abdulmajid, M.Pd., IPM., ASEAN Eng.**
**NIP. 920171219910625101**

## LEMBAR PERNYATAAN

## HALAMAN PERNYATAAN TENTANG KEASLIAN SKRIPSI DAN PERNYATAAN BEBAS PLAGIARISME

Saya yang bertanda tangan dibawah ini:

Nama : Hilya Anbiyani Fitri Muhyidin

NIM : 2005517

Program Studi : Pendidikan Sistem dan Teknologi Informasi

Dengan ini saya menyatakan bahwa skripsi dengan judul "**Implementasi Algoritma *Logistic Regression* untuk Deteksi *Link Phishing***" ini beserta seluruh isinya adalah benar-benar hasil karya saya sendiri. Saya tidak melakukan penjiplakan atau mengutip dengan cara-cara yang tidak sesuai dengan etika keilmuan yang berlaku dalam masyarakat keilmuan. Atas pernyataan ini saya siap menanggung resiko dan sanksi apabila dikemudian hari ditemukan adanya pelanggaran etika keilmuan atau ada klaim dari pihak lain terhadap keaslian karya saya ini.

Purwakarta, Juni 2024

Pembuat Pernyataan


Hilya Anbiyani Fitri Muhyidin

NIM. 2005517

# KATA PENGANTAR

Alhamdulillah, puji dan syukur penulis panjatkan kehadirat Allah SWT yang telah melimpahkan segala nikmat, sehat, rahmat, dan karunia serta mukjizat-Nya, sehingga penulis dapat menyelesaikan Skripsi dengan judul "Implementasi Algoritma *Logistic Regression* untuk Deteksi *Link Phishing*".

Skripsi ini berisi tentang penelitian dalam membangun algoritma model *Logistic Regression* dalam mendeteksi potensi *link* yang terindikasi *phishing* secara akurat. Peneliti berharap hasil penelitian ini mampu meningkatkan keamanan siber dan mengurangi risiko *phishing* bagi pengguna *internet*.

Penulis telah berusaha dengan sebaik mungkin dengan kemampuan yang ada dalam menyelesaikan skripsi ini untuk mendapatkan hasil yang sebaik-baiknya. Namun penulis menyadari bahwa penulisan Skripsi ini masih jauh dari kesempurnaan. Oleh karena itu, dengan segala kerendahan hati, penulis sangat menghargai segala kritik serta saran yang bersifat membangun ke arah perbaikan dan penyempurnaan skripsi ini. Semoga Skripsi ini dapat memberikan manfaat bagi banyak pihak, khususnya bagi penulis dan umumnya bagi pembaca. Akhir kata, penulis berharap semoga Allah SWT senantiasa membalas kebaikan semua pihak yang telah membantu. Aamiin Yaa Rabbal Aalamiin.

Purwakarta, Juni 2024

Penulis

## UCAPAN TERIMA KASIH

Dalam menyusun skripsi ini, penulis tidak luput dari berbagai kesulitan dan hambatan, namun atas bantuan dan dorongan dari erbagai pihak akhirnya penulis dapat menyelesaikan skripsi ini. Oleh karena itu, pada kesempatan ini, penulis ingin menyampaikan rasa terima kasih yang sebesar-besarnya kepada semua pihak yang telah membantu, yaitu kepada:

1. Ir. Nuur Wachid Abdulmajid, M.Pd., IPM. ASEAN Eng. Selaku Ketua Program Studi Pendidikan Sistem dan Teknologi Informasi UPI Kampus Daerah Purwakarta serta selaku Dosen Penguji I.
2. Dr. H. Suprih Widodo, S.Si., M.T. dan Ibu Liptia Venica, S.T., M.T. selaku Dosen Pembimbing I dan Dosen Pembimbing II dalam pembuatan Skripsi ini yang telah menyediakan waktu, tenaga, dan pemikirannya dalam membantu, membimbing, dan memberikan masukan serta yang selalu memberi masukan dan motivasi kepada penulis dalam proses menyusun Skripsi.
3. Bapak Rizki Hikmawan, S.Pd., M.Pd. selaku Dosen Penguji II dalam pembuatan Skripsi ini yang telah bersedia menguji penulis.
4. Ibu Dian Permata Sari, S.Kom., M.Kom selaku Dosen Penguji III dalam pembuatan Skripsi ini yang telah bersedia menguji penulis.
5. Assoc. Prof. Dr. Mohamed Nor Azhari bin Azman selaku Dosen Penguji IV dalam pembuatan Skripsi ini yang telah bersedia menguji penulis.
6. Seluruh Dosen UPI Kampus Purwakarta dan Civitas akademik yang telah memberikan pengetahuan, arahan, dan membantu dalam proses penyelesaian skripsi ini.
7. Bapak dan Ibu Dosen Program Studi Pendidikan Sistem dan Teknologi Informasi UPI Kampus Daerah Purwakarta yang telah memberikan ilmu, arahan, dan dukungan kepada penulis untuk menyelesaikan Skripsi ini.
8. Bapak (Muhyidin), Ibu (Nengsih), Kakak (Nadiya Fitri Ningsih), dan Adik (M. Zam-Zam Firdaus MHN) yang telah memberikan doa, cinta, kasih, dukungan dan motivasi terbaik untuk menyelesaikan Skripsi ini.

9. Teman-teman seperjuangan, rekan-rekan mahasiswa/I jurusan PSTI, khususnya Angkatan 2020 yang telah memberikan dukungan, bantuan dan kerjasamanya selama proses penyusunan Skripsi ini.
10. Serta seluruh pihak yang terlibat secara langsung maupun tidak langsung yang tidak bisa penulis sebutkan satu persatu selama proses Penyusunan Skripsi ini, terimakasih atas segala bantuan dan dukungannya.

**IMPLEMENTASI ALGORITMA *LOGISTIC REGRESSION***

**UNTUK DETEKSI *LINK PHISHING***

Hilya Anbiyani Fitri Muhyidin

NIM: 2005517

## ABSTRAK

Peningkatan pengguna *internet* yang massif memicu peningkatan tindak kejahatan dunia maya melalui jaringan *internet* salah satunya *phishing* yang mana berpotensi merugikan pihak yang terkena dampak dalam hal finansial. Tujuan dari penelitian ini melengkapi penelitian terdahulu yaitu membuat *website* aplikasi deteksi *link phsihing* dengan menerapkan algoritma *logistic regression* sehingga diharapkan dapat mengurangi resiko dari kejahatan *phishing*. Metode dan desain penelitian menerapkan *Framework* AI *Project Cycle* untuk proses pengembangan *website* aplikasi. *Dataset* yang digunakan merupakan data URL yang diperoleh dari berbagai sumber. Berdasarkan hasil penelitian dapat disimpulkan: 1) Model *logistic regression* diimplementasikan ke dalam bentuk *website* aplikasi deteksi *link phishing* melalui tahapan *Framework* AI *Project Cycle* dengan memanfaatkan teknologi HTML, CSS, *JavaScript*, *Flask framework*, dan ngrok; 2) Algoritma model *logistic regression* yang dilatih menggunakan fitur *top*-6 dan fitur tambahan *path_len* menghasilkan skor akurasi sebesar 90,21% setelah dilakukan 10-*fold cross validation*, serta memiliki skor rata-rata *precision* sebesar 90%, *recall* sebesar 91%, dan f1-*score* sebesar 90%.

**Kata kunci**: AI *Project Cycle*, *Logistic Regression*, *Machine Learning*, *Phishing Link Detection*.

***IMPLEMENTATION OF LOGISTIC REGRESSION ALGORITHM***

***FOR PHISHING LINK DETECTION***

Hilya Anbiyani Fitri Muhyidin

NIM: 2005517

## *ABSTRACT*

*The massive increase in internet users has triggered a rise in cybercrime through the internet, one of which is phishing, which has the potential to cause financial losses to affected parties. The aim of this research is to build upon previous studies by developing a website application for phishing link detection using the logistic regression algorithm, with the hope of reducing the risks of phishing crimes. The research method and design utilize the AI Project Cycle Framework for the website application development process. The dataset used comprises URL data obtained from various sources. Based on the research results, the following conclusions can be drawn: 1) The logistic regression model was implemented in a website application for phishing link detection through the stages of the AI Project Cycle Framework, utilizing HTML, CSS, JavaScript, Flask framework, and ngrok technologies; 2) The logistic regression model algorithm, trained using the top-6 features and an additional path_len feature, achieved an accuracy score of 90.21% after 10-fold cross-validation, with an average precision score of 90%, recall score of 91%, and f1-score of 90%.*

***Keywords****: AI Project Cycle*, *Logistic Regression*, *Machine Learning, Phishing Link Detection.*

# DAFTAR ISI

## DAFTAR GAMBAR

# DAFTAR TABEL

## DAFTAR LAMPIRAN

# DAFTAR PUSTAKA


Ajani, Y., Mangalorkar, K., Nadar, Y., Mehra, M., & Kalbande, D. (2021). College Project Preservation and Emulation Using Containerization Over Private Cloud. *Lecture Notes in Networks and Systems*, *190*(Ictcs 2020), 535–544. https://doi.org/10.1007/978-981-16-0882-7_46

Ajhari, A. A., Priambodo, D. F., Paradisa, R. H., & Yulianti, H. (2023). PROCTOR: A Robust URL Protection System Against Fraudulent, Phishing, and Scam Activities. *International Journal of Computing and Digital Systems*, *14*(1), 1013–1021. https://doi.org/10.12785/IJCDS/140179

Akshaya, J. (2020). *Phishing Websites Detection*. Kaggle. https://www.kaggle.com/datasets/akshaya1508/phishing-websites-detection

Al-Khater, W. A., Al-Maadeed, S., Ahmed, A. A., Sadiq, A. S., & Khan, M. K. (2020). Comprehensive review of cybercrime detection techniques. *IEEE Access*, *8*, 137293–137311. https://doi.org/10.1109/ACCESS.2020.3011259

Aleedy, M., Shaiba, H., & Bezbradica, M. (2019). Generating and analyzing Chatbot responses using natural language processing. *International Journal of Advanced Computer Science and Applications*, *10*(9), 60–68. https://doi.org/10.14569/ijacsa.2019.0100910

Alkhalil, Z., Hewage, C., Nawaf, L., & Khan, I. (2021). Phishing Attacks: A Recent Comprehensive Study and a New Anatomy. *Frontiers in Computer Science*, *3*(March), 1–23. https://doi.org/10.3389/fcomp.2021.563060

Anam, C., & Santoso, H. B. (2018). Perbandingan Kinerja Algoritma C4.5 dan Naive Bayes untuk Klasifikasi Penerima Beasiswa. *Energy - Jurnal Ilmiah Ilmu-Ilmu Teknik*, *8*(1), 13–19. https://ejournal.upm.ac.id/index.php/energy/article/view/111

Anti Phishing Working Group. (2023). APWG Phishing Activity Trend Report, 1st Quarter 2023. *Safe Places*, *June*, 80–88.

Arvind Prasad & Shalini Chandra. (2023). *PhiUSIIL: A diverse security profile empowered phishing URL detection framework based on similarity index and incremental learning*. UCI Machine Learning Repository. https://doi.org/https://doi.org/10.1016/j.cose.2023.103545

Assaidi, S. A., & Amin, F. (2022). Analisis Sentimen Evaluasi Pembelajaran Tatap Muka 100 Persen pada Pengguna Twitter menggunakan Metode Logistic Regression. *Jurnal Pendidikan Tambusai*, *6*(2), 13217–13227.

Ayu, A., Permata, N., & Ratnawati, E. (2023). *Tinjauan Kasus Cyber Phishing 16 Shop Berdasarkan UU ITE Nomor 19 Tahun 2016*. *6*(1), 2771–2779.

Badawi, A. (2021). the Effectiveness of Natural Language Processing (Nlp) As a Processing Solution and Semantic Improvement. *International Journal of Economic, Technology and Social Sciences (Injects)*, *2*(1), 36–44. https://doi.org/10.53695/injects.v2i1.194

Bathla, Y., Verma, C., & Kumar, N. (2020). Smart approach for real-time gender prediction of European school's principal using machine learning. *Lecture Notes in Electrical Engineering*, *597*, 159–175. https://doi.org/10.1007/978-3-030-29407-6_14

Bhatt, Y., & Pahade, P. (2021). Application of Python Programming and Its Future. *Lecture Notes in Networks and Systems*, *190*(Ictcs), 849–857. https://doi.org/10.1007/978-981-16-0882-7_76

Bimantara, A., & Dina, T. A. (2019). Klasifikasi Web Berbahaya Menggunakan Metode Logistic Regression. *Annual Research Seminar (ARS)*, *4*(1), 173–177. https://seminar.ilkom.unsri.ac.id/index.php/ars/article/view/1932

BSSN. (2024). *Lanskap Keamanan Siber Indonesia*. *70*.

Cahyanti, D., Rahmayani, A., & Husniar, S. A. (2020). Analisis performa metode Knn pada Dataset pasien pengidap Kanker Payudara. *Indonesian Journal of Data and Science*, *1*(2), 39–43. https://doi.org/10.33096/ijodas.v1i2.13

Carneiro, T., Da Nobrega, R. V. M., Nepomuceno, T., Bian, G. Bin, De Albuquerque, V. H. C., & Filho, P. P. R. (2018). Performance Analysis of Google Colaboratory as a Tool for Accelerating Deep Learning Applications. *IEEE Access*, *6*, 61677–61685. https://doi.org/10.1109/ACCESS.2018.2874767

Chandra, A. A., Nathaniel, V., Satura, F. R., & Adhinata, F. D. (2022). Pengembangan Chatbot Informasi Mahasiswa Berbasis Telegram dengan Metode Natural Language Processing. *Journal ICTEE*, *3*(1), 20. https://doi.org/10.33365/jictee.v3i1.1886

CISA, C. and I. S. A. (2021). *Phishing (General Security Postcard)*. https://www.cisa.gov/sites/default/files/publications/Phishing General Security Postcard_6.24.2021_508cV2.pdf

Das Guptta, S., Shahriar, K. T., Alqahtani, H., Alsalman, D., & Sarker, I. H. (2024). Modeling Hybrid Feature-Based Phishing Websites Detection Using Machine Learning Techniques. *Annals of Data Science*, *11*(1), 217–242. https://doi.org/10.1007/s40745-022-00379-8

Dhall, D., Kaur, R., & Juneja, M. (2020). Machine learning: A review of the algorithms and its applications. *Lecture Notes in Electrical Engineering*, *597*, 47–63. https://doi.org/10.1007/978-3-030-29407-6_5

Dicoding. (2024a). *Belajar Dasar Pemrograman JavaScript*. https://www.dicoding.com/academies/256/tutorials/13857?from=15292

Dicoding. (2024b). *Belajar Dasar Pemrograman Web*. https://www.dicoding.com/academies/123/corridor

Dicoding. (2024c). *Memulai Pemrograman dengan Python*. https://www.dicoding.com/academies/86/corridor

Elkan, C. (2014). *Maximum Likelihood , Logistic Regression , and Stochastic Gradient Training Examples of maximizing likelihood*.

Fandru, M., Rifqi, A., Dina, M., Nababan, M. N. K., & Aisyah, S. (2022). Comparative Analysis of Phishing Website Prediction Classification Algorithm Using Logistic Regression , Decision Tree , and Random Forest. *Infor.Seaninstitute.Org*, *10*(2), 859–869. http://infor.seaninstitute.org/index.php/infokum/article/view/425

Fatkhurohman, A., & Pujastuti, E. (2019). Penerapan Algoritma Naïve Bayes Classifier Untuk Meningkatkan Keamanan Data Dari Website Phising. *Respati*, *14*(1), 115–124. https://doi.org/10.35842/jtir.v14i1.279

Ferelestian, V. J., Susanto, B., & Senapartha, I. K. D. (2023). *Pengembangan Telegram Chatbot Informasi Mahasiswa Menggunakan Wit.ai*. 89–97. https://doi.org/10.21460/jutei.72.257

Gani, A. G. (2014). Cybercrime (Kejahatan Berbasis Komputer). *Jurnal Sistem Informasi Universitas Suryadarma*, *5*(1), 16–29. https://doi.org/10.35968/jsi.v5i1.18

Ghimire, D. (2020). Comparative study on Python web frameworks: Flask and Django. *Metropolia University of Applied Sciences*, *May*, 13–33. https://urn.fi/URN:NBN:fi:amk-2020052513398

Harahap, A. H., Difa Andani, C., Christie, A., Nurhaliza, D., & Fauzi, A. (2023). Pentingnya Peranan CIA Triad Dalam Keamanan Informasi dan Data Untuk Pemangku Kepentingan atau Stakholder. *Jurnal Manajemen Dan Pemasaran Digital*, *1*(2), 73–83.

Ichi.pro. (2020). *Apa itu algoritma Regresi Logistik dan bagaimana cara kerjanya?* https://ichi.pro/id/apa-itu-algoritma-regresi-logistik-dan-bagaimana-cara-kerjanya-161590515918689

Internet World State. (2022). *Asia Internet Use, Population Statistics Data and Facebook Data MidYear 2022*. Miniwatts Marketing Group. https://apwg.org/trendsreports/#:~:text=Summary – 1st Quarter 2022,t

Jared, D. (2014). *Big data, data mining, and machine learning [internet resource]: value creation for business leaders and practitioners*.

Johri, P., Khatri, S. K., Al-Taani, A. T., Sabharwal, M., Suvanov, S., & Kumar, A. (2021). Natural Language Processing: History, Evolution, Application, and Future Work. *Lecture Notes in Networks and Systems*, *167*, 365–375. https://doi.org/10.1007/978-981-15-9712-1_31

Joshi, A. P., & Patel, B. V. (2021). Data Preprocessing: The Techniques for Preparing Clean and Quality Data for Data Analytics Process. *Oriental Journal of Computer Science and Technology*, *13*(0203), 78–81. https://doi.org/10.13005/ojcst13.0203.03

Kharisma Putra, I. K. O., Darmawan, I. M. A., Juliana, I. P. G., & Indriyani. (2023). Tindakan Kejahatan Pada Dunia Digital Dalam Bentuk Phising. *Cyber Security Dan Forensik Digital*, *5*(2), 77–82. https://doi.org/10.14421/csecurity.2022.5.2.3797

Kirasich, K. ;, Smith, T. ;, & Sadler, B. (2018). Random Forest vs Logistic Regression: Binary Classification for Heterogeneous Datasets. *SMU Data Science Review*, *1*(3), 9. https://scholar.smu.edu/datasciencereviewAvailableat:https://scholar.smu.edu/datasciencereview/vol1/iss3/9http://digitalrepository.smu.edu.

Kurnia, R. P., Atma, Y. A., & Padang, P. N. (2022). *ANALISIS REKOMENDASI FILM DARI DATA IMDB MENGGUNAKAN PYTHON ANALYSIS OF FILM RECOMMENDATIONS FROM IMDB DATA*. *4*(2).

Madhumitha, C., Likhitha, P. S., & Sai, C. P. (2022). *Phishing Websites Detection Using Machine Learning*. 1–4.

Mahesh, B. (2019). *Machine Learning Algorithms - A Review | Enhanced Reader*. *October*. https://doi.org/10.21275/ART20203995

Malau, T., & Joseph, T. (2023). Analisis Metode Logistik Regresi Ensemble untuk Klasifikasi dengan Pra-Pemrosesan Menggunakan Principal Component Analysis. *IJM: Indonesian Journal of Multidisciplinary*, *1*, 707–722. https://journal.csspublishing/index.php/ijm

Moz. (2021). *The Moz Top 500 Websites*. Moz. https://moz.com/top500

Nasteski, V. (2017). An overview of the supervised machine learning methods. *Horizons.B*, *4*(December), 51–62. https://doi.org/10.20544/horizons.b.04.1.17.p05

Nur Latifah, F., Mawardi, I., & Wardhana, B. (2022). Threat of Data Theft (Phishing) Amid Trends in Fintech Users During the Covid-19 Pandemic (Study Phishing In Indonesia). *Perisai : Islamic Banking and Finance Journal*, *6*(1), 74–86. https://doi.org/10.21070/perisai.v6i1.1598

PhishTank. (2024). *PhishTank*. PhishTank. https://phishtank.com/developer_info.php

Pramakrisna, F. D., Adhinata, F. D., & Tanjung, N. A. F. (2022). Aplikasi Klasifikasi SMS Berbasis Web Menggunakan Algoritma Logistic Regression. *Teknika*, *11*(2), 90–97. https://doi.org/10.34148/teknika.v11i2.466

Pratama, A., Nurcahyo, A. C., & Firgia, L. (2023). Penerapan Machine Learning dengan Algoritma Logistik Regresi untuk Memprediksi Diabetes. *Prosiding CORISINDO 2023*, 116–121. https://stmikpontianak.org/ojs/index.php/corisindo/article/view/30%0Ahttps://stmikpontianak.org/ojs/index.php/corisindo/article/download/30/22

Rian Handoko, & Tata Sutabri. (2023). Analisa Machine Learning Dengan Algoritma Multi-Layer Perceptron Untuk Penanganan Kejahatan Phishing. *Jurnal Informatika Teknologi Dan Sains*, *5*(1), 13–17. https://doi.org/10.51401/jinteks.v5i1.2221

Rifa'i, H., Ryan Hamonangan, Dian Ade Kurnia, Kaslani, & Mulyawan. (2022). Implementasi Algoritma Decision Tree Dalam Klasifikasi Kompetensi Siswa. *KOPERTIP : Jurnal Ilmiah Manajemen Informatika Dan Komputer*, *6*(1), 15–20. https://doi.org/10.32485/kopertip.v6i1.131

Rina Noviana. (2022). Pembuatan Aplikasi Penjualan Berbasis Web Monja Store Menggunakan Php Dan Mysql. *Jurnal Teknik Dan Science*, *1*(2), 112–124. https://doi.org/10.56127/jts.v1i2.128

Roihan, A., Sunarya, P. A., & Rafika, A. S. (2020). Pemanfaatan Machine Learning dalam Berbagai Bidang: Review paper. *IJCIT (Indonesian Journal on Computer and Information Technology)*, *5*(1), 75–82. https://doi.org/10.31294/ijcit.v5i1.7951

Rosid, M. A., Fitrani, A. S., Astutik, I. R. I., Mulloh, N. I., & Gozali, H. A. (2020). Improving Text Preprocessing for Student Complaint Document Classification Using Sastrawi. *IOP Conference Series: Materials Science and Engineering*, *874*(1). https://doi.org/10.1088/1757-899X/874/1/012017

Salmu, S., & Solichin, A. (2017). Prediksi Tingkat Kelulusan Mahasiswa Tepat Waktu Menggunakan Naïve Bayes : Studi Kasus UIN Syarif Hidayatullah Jakarta. *Seminar Nasional Multidisiplin Ilmu (SENMI) 2017*, *April*, 701–709.

Sandag, G. A., Leopold, J., & Ong, V. F. (2018). Klasifikasi Malicious Websites Menggunakan Algoritma K-NN Berdasarkan Application Layers dan Network

Characteristics. *CogITo Smart Journal*, *4*(1), 37–45. https://doi.org/10.31154/cogito.v4i1.100.37-45

Santosa, F. (2022). Akurasi dalam Mengidentifikasi Tingkat Similarity pada Artikel Ilmiah Menggunakan Algorithma Jaro Winkler. *Jurnal Informasi Dan Teknologi*, *4*(3), 142–147. https://doi.org/10.37034/jidt.v4i3.217

Sen, P. C., Hajra, M., & Ghosh, M. (2020). Supervised Classification Algorithms in Machine Learning: A Survey and Review. In *Advances in Intelligent Systems and Computing* (Vol. 937). Springer Singapore. https://doi.org/10.1007/978-981-13-7403-6_11

Setiawan, A., Santoso, L. W., & Adipranata, R. (2020). Klasifikasi Artikel Berita Bahasa Indonesia Dengan Naive Bayes Classifier. *Jurnal Infra*, *8*(1), 146–151.

Siddhartha, M. (2021). *Malicious URLs dataset*. Kaggle. https://www.kaggle.com/datasets/sid321axn/malicious-urls-dataset

Sudrajat, W., & Cholid, I. (2023). K-Nearest Neighbor (K-Nn) Untuk Penanganan Missing Value Pada Data Umkm. *Jurnal Rekayasa Sistem Informasi Dan Teknologi*, *1*(2), 54–63. https://doi.org/10.59407/jrsit.v1i2.77

Tiwari, T. (2020). *Phishing Site URLs*. Kaggle. https://www.kaggle.com/datasets/taruntiwarihp/phishing-site-urls

Tyagi, V. (2023). *Phishtank Updated Dataset*. Kaggle. https://www.kaggle.com/datasets/vishalxtyagi/phishing-website-detector-based-on-url

Ubing, A. A., Jasmi, S. K. B., Abdullah, A., Jhanjhi, N. Z., & Supramaniam, M. (2019). Phishing website detection: An improved accuracy through feature selection and ensemble learning. *International Journal of Advanced Computer Science and Applications*, *10*(1), 252–257. https://doi.org/10.14569/IJACSA.2019.0100133

Vayansky, I., & Kumar, S. (2018). Phishing – challenges and solutions. *Computer Fraud and Security*, *2018*(1), 15–20. https://doi.org/10.1016/S1361-3723(18)30007-1

Widodo, S., Setiawan, D., Ridwan, T., & Ambari, R. (2022). Perancangan Deteksi Emosi Manusia berdasarkan Ekspresi Wajah Menggunakan Algoritma VGG16. *Syntax : Jurnal Informatika*, *11*(01), 01–12. https://doi.org/10.35706/syji.v11i01.6594

Winson. (2024). *Dataset for Link Phishing Detection*. Kaggle. https://www.kaggle.com/datasets/winson13/dataset-for-link-phishing-detection

Yenni, Y., Utnasari, I., & Rahmawati, M. (2021). Sosialisasi Pemanfaatan Teknologi Informasi Internet Berbasis Media Sosial Sebagai Usaha dan Transaksi. *Jurdimas (Jurnal Pengabdian Kepada Masyarakat) Royal*, *4*(1), 1–6. https://doi.org/10.33330/jurdimas.v4i1.543

Yuda, P. A. P. D., & Hendra Suputra, I. P. G. (2021). Implementation of the Support Vector Machine (SVM) Algorithm in Classifying Website Phishing. *JELIKU (Jurnal Elektronik Ilmu Komputer Udayana)*, *9*(4), 467. https://doi.org/10.24843/jlk.2021.v09.i04.p03

Yuniati, T., & Kresna A., I. (2020). Secure Electronic Payment Methods for Online Shopping Based on Visual Cryptography. *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, *4*(2), 319–328. https://doi.org/10.29207/resti.v4i2.1732