p-ISSN:2527-5321
e-ISSN:2527-5941

PREDIKSI EMAIL PHISING MENGGUNAKAN SUPPORT VECTOR MACHINE

Chaerul Umam[1], L. Budi Handoko[2]

Universitas Dian Nuswantoro
Jl. Imam Bonjol No 207, Pendrikan Kidul, 50131, Semarang, Jawa Tengah, Indonesia
[1]chaerul@dsn.dinus.ac.id, [2]handoko@dsn.dinus.ac.id

ABSTRAK

Email phising merupakan salah satu bentuk kejahatan di internet yang dapat merugikan banyak orang. Ketika seseorang sudah terkena phising maka data data orang tersebut dapat hilang dan digunakan oleh orang yang tidak bertanggung jawab. Pada penelitian ini, akan melakukan proses klasiifkasi *email phising* dengan menggunakan bantuan *machine learning* yaitu algoritma SVM. Dataset yang digunakan pada penelitian ini yaiitu merupakan dataset yang berisi body email yang terdiri dari total 18650 data yang terdiri dari 11322 data safe email dan 7328 data *phising* email. Dari data tersebut, akan dibagi menjadi 70% data pelatihan dan 30% data pengujian. Setelah dilakukan proses pengujian pada penelitian ini, algoritma SVM yang digunakan mendapatkan akurasi pengujian sebesar 84.56%
Kata Kunci: *Email Phising, Machine Learning,* SVM, TF-IDF

ABSTRACT

Email phishing is a form of crime on the internet that can harm many people. When someone has been exposed to phishing, the person's data can be lost and used by irresponsible people. In this research, the process of classifying phishing emails will be carried out using the help of machine learning, namely the SVM algorithm. The dataset used in this research is a dataset containing body emails consisting of a total of 18650 data consisting of 11322 safe email data and 7328 phishing email data. From this data, it will be divided into 70% training data and 30% testing data. After the testing process in this study, the SVM algorithm used obtained a testing accuracy of 84.56%.
Key Word: *Phising Email, Machine Learning, SVM, TF-IDF*

PENDAHULUAN

Perkembangan teknologi saat ini semakin pesat. Tentunya hal ini menguntungkan kita karena berkat teknologi dapat mempermudah aktivitas kita. Bidang telekomunikasi juga berkembang pesat dengan berkembangnya teknologi internet, membantu masyarakat untuk berkomunikasi dimana saja dan kapan saja. Namun dampak *negative* yang ditimbulkan oleh kemajuan teknologi adalah meningkatnya penipuan yang dilakukan melalui internet. Salah satu kemungkinan penipuan adalah *phishing*. *Phishing* adalah cara yang mengeksploitasi pengguna Internet untuk mendapatkan informasi penting dan sensitif dari pengguna tersebut (Salloum et all, 2022). Oleh karena itu, ketika data diperoleh, data tersebut dapat menjadi ancaman dari pelakukepada pihak yang dirugikan agar pihak yang dirugikan dapat membayar uang tebusan agar data tersebut tidak diungkapkan. Data tahun 2020 menunjukkan bahwa 75% bisnis dan organisasi mengalami serangan phishing dan 96% serangan phishing terjadi melalui email (Alhogail et.all, 2021). Oleh karena itu, penting untuk dapat melakukan proses pengklasifikasian email yang diterima agar kami dapat memastikan bahwa email tersebut bukanlah email phishing yang dapat menimbulkan kerugian bagi pihak-pihak yang terlibat.

Machine learning merupakan salah satu bagian dari teknologi kecerdasan buatan (AI), dimana komputer dapat melakukan proses pembelajaran secara mandiri tanpa harus diprogram terlebih dahulu (Bi et.al, 2019) untuk dapat melakukan proses pembelajaran secara pertunjukan. Dalam proses pembelajaran data, pembelajaran mesin dibagi menjadi 3 metode, yaitu pembelajaran yang diawasi atau pembelajaran dari data yang ditargetkan, pembelajaran tanpa pengawasan atau model yang dibangun yang dapat belajar dari data dan menentukan Identifikasi tujuan atau hasil yang terkandung dalam data.berdasarkan kesamaan antara data, data lain dan juga model pembelajaran penguatan atau tindakan untuk melakukan sesuatu dan menerima imbalan berupa

p-ISSN:2527-5321
e-ISSN:2527-5941

imbalan atau hukuman berupa umpan balik terhadap angka kinerja pembelajaran model tersebut (Sah et. al, 2020). *Support vector machine* merupakan algoritma pembelajaran mesin pencari hyperplane, dimana *hyperplane* digunakan untuk membagi kelas data dan merupakan *hyperplane* yang paling optimal serta mempunyai amplitudo tertinggi di setiap kelasnya (Boateng et. al, 2020), sehingga hyperplane tersebut dekat dengan sampel data kedua kelas. Oleh karena itu, SVM merupakan metode yang memberikan kinerja klasifikasi data yang baik dibandingkan dengan algoritma klasifikasi lainnya (Muthukrishnan et. al, 2020).

Vectorizer TF-IDF adalah metode yang digunakan untuk melakukan konversi teks ke vektor (Kumar et. al, 2020). Selama proses ini, Term Frekuensi Invers Dokumen Frekuensi atau TF-IDF juga digunakan untuk menghitung pentingnya kata-kata yang muncul dalam dokumen menggunakan metode statistik (Akuma et. al, 2020). Melalui ini, data yang diperoleh dapat digunakan untuk pelatihan dan pengujian model. Matriks konfusi adalah matriks yang dapat digunakan untuk menghitung kinerja suatu model (Krstinić et. al, 2020) menggunakan presisi, akurasi, recall serta poin f1. Dengan nilai-nilai tersebut maka dapat dilakukan perhitungan kinerja dan dapat diketahui kualitas model yang dibangun yang dapat melakukan proses klasifikasi.

Tujuan dilakukannya proses penelitian ini yaitu agar dapat membangun model yang dapat membantu melakukan proses klasifikasi email phising sehingga orang tidak terkena phising melalui email. Sedangkan tujuan melakukan proses klasifikasi dengan menggunakan algoritma SVM yaitu karena algoritma ini merupakan algoritma yang baik untuk melakukan proses klasifikasi data.

Dalam sebuah penelitian yang dilakukan oleh M.O.Khairandish a, M. Sharma b, V. Jain c, J.M. Chatterjee d dan N.Z. Jhanjhi e, membahas tentang klasifikasi dan deteksi tumor otak menggunakan kombinasi metode CNN dan SVM. Tujuan dari penelitian ini adalah membangun model yang mampu melakukan proses klasifikasi tumor otak ganas dan jinak berdasarkan citra MRI otak dengan ekstraksi ciri menggunakan CNN dan proses klasifikasi berbantuan SVM. Hasil yang diperoleh setelah dilakukan pengujian penelitian ini mempunyai akurasi sebesar 98,4959%. Dalam penelitian yang dilakukan oleh R.Vijayarajeswari, P.Parthasarathy, S.Vivekanandan dan A.Alavudeen Basha membahas tentang membangun model menggunakan kombinasi SVM dan Hough Transform untuk melakukan proses klasifikasi jenis kanker payudara. Tujuan dari penelitian ini adalah membangun model untuk mendeteksi dan mengklasifikasikan kanker payudara secara otomatis berdasarkan gambar mamografi. Hasil yang diperoleh setelah selesai mencapai akurasi pengujian sebesar 94%.

METODE PENELITIAN

Dataset yang akan digunakan pada penelitian ini yaitu merupakan dataset email phising yang didapatkan dari website kaggle.com dengan judul Phising Email Detection. Dengan total data yang digunakan yaitu sebanyak 18650 data email phising yang terdiri dari 11322 data safe email dan 7328 data phising email. Untuk visualisasi dataset diberikan pada Gambar 1.

	Email Text	Email Type
0	re : 6 . 1100 , disc : uniformitarianism , re ...	Safe Email
1	the other side of * galicismos * * galicismo *...	Safe Email
2	re : equistar deal tickets are you still avail...	Safe Email
3	\nHello I am your hot lil horny toy.\n I am...	Phishing Email
4	software at incredibly low prices (86 % lower...	Phishing Email

Gambar 1. Dataset penelitian

Dari total 18650 data citra tersebut, akan dibagi menjadi 70% data pelatihan dan 30% data pengujian. Yang dimana data pelatihan digunakan agar model dapat berlatih mengenali pola dari data. Sedangkan, data pengujian berguna untuk melakukan proses evaluasi performa model dalam melakukan proses klasifikasi.

Support Vector Machine merupakan algoritma yang akan mencari *hyperplane* dimana *hyperplane* tersebut merupakan hyperplane yang paling optimal dan memiliki amplitudo tertinggi di setiap kelasnya (Boateng et. al, 2020) untuk membagi data ke dalam kelas-kelas yang ada. Untuk rumus SVM diberikan pada poin 1, 2 dan 3.

$$\omega * x + b \quad (1)$$

Dimana:
ω = nilai vector bobot (untuk menentukan orientasi arah dari hyperplane)
x = nilai fitur yang digunakan dalam bentuk vector
b = nilai bias (nilai ruang jarak hyperplane menuju titik asal (0,0) pada lingkup ruang fitur)

$$Margin = \frac{2}{||\omega||} \quad (2)$$

Dimana ||ω|| = nilai Panjang dari vector bobot pada SVM

$$Maksimal \ \frac{2}{||\omega||} \ dengan \ batas \ z_i(\omega * x_i + b) \geq 1 \ untuk \ data \ pelatihan \ (x_i, z_i) \quad (3)$$

Dimana dengan memaksimalkan margin dengan Batasan setidaknya jarak 1 dari hyperplane dengan nilai z_i merupakan nilai target pada data yang diberikan dan x_i merupakan nilai fitur dari data. Vectorizer TF-IDF (*Term Frekuensi-Inverse Document* Frekuensi) adalah metode yang digunakan untuk melakukan proses konversi teks ke format vektor dan juga digunakan untuk melakukan proses penghitungan kemunculan kata-kata penting dalam dokumen dengan menggunakan metode statistik. Rumus TF-IDF diberikan pada poin 6. Pada proses ini kita akan mencari nilai TF yang digunakan untuk mengukur frekuensi kemunculan katadalam dokumen. Rumus TF diberikan pada poin 5. Kemudian, nilai IDF digunakan untuk menentukan kepentingan kata secara keseluruhan dalam keseluruhan isi dokumen. Rumus IDF diberikan pada poin 4

$$TF(kata, dokumen) = \frac{jumlah \ muncul \ kata \ dalam \ dokumen}{total \ kata \ dalam \ dokumen} \quad (4)$$

$$IDF(kata, korpus) = \log\left(\frac{total \ dokumen \ dalam \ korpus}{jumlah \ dokumen \ yang \ mengandung \ kata + 1}\right) + 1 \quad (5)$$

$$TFIDF(kata, dokumen, korpus) = TF(kata, dokumen) * IDF(kata, korpus) \quad (6)$$

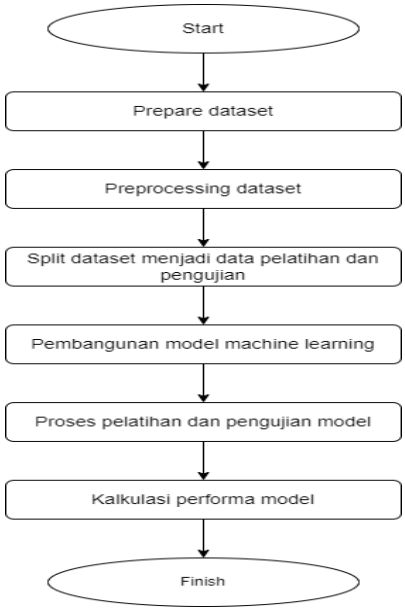
Confusion matrix adalah matriks yang dapat digunakan untuk menghitung performa suatu model menggunakan nilai presisi, akurasi, recall, serta nilai f1-score. Dengan demikian kita dapat mengetahui kualitas model yang dibangun untuk dapat melakukan proses klasifikasi. Pada poin 7 disajikan rumus perhitungan keakuratan prediksi model. Poin 8 menunjukkan rumus yang menentukan performa model dalam membuat prediksi akurat untuk setiap data atau kelas yang ada. Sedangkan poin 9 menyajikan rumus untuk mencari nilai keselarasan atau keseimbangan antara nilai presisi dan nilai recall yang diperoleh sebelumnya.

$$Presisi = \frac{Pos \ benar}{Pos \ benar + Pos \ salah} \quad (7)$$

$$Recall = \frac{Pos \ benar}{Pos \ benar + Neg \ salah} \quad (8)$$

$$F1 - Score = 2 * \frac{Presisi * Recall}{Presisi + Recall} \quad (9)$$

Pada penelitian ini, akan menggunakan IDE Jupyter Notebook dan Bahasa pemrograman python untuk melakukan proses klasifikasi *email phising*. Untuk alur proses klasifikasi diberikan pada Gambar 2.



Gambar 2. Tahapan penelitian

Gambar 2 menunjukkan urutan penelitian yang akan dilakukan. Pertama-tama, dalam penelitian ini, data akan dibaca. Kemudian setelah data dibaca maka dilakukan preprocessing data yaitu proses

menghilangkan stopword dan proses mengubah data teks menjadi vektor. Setelah data diolah terlebih dahulu, makadata tersebut dibagi menjadi 70% data untuk pelatihan dan 30% data untuk pengujian. Anda kemudian dapat membuat model SVM. Setelah data dan model penggunaan telah disiapkan, pelatihan dan pengujian model dapat dilakukan dan akurasi serta kinerja model saat melakukan proses klasifikasi dapat dihitung menggunakan matriks konfusi setelah pengujian.

HASIL DAN PEMBAHASAN

Pada penelitian ini, proses pengujian akan menggunakan IDE Jupyter notebook dan Bahasa pemrograman python untuk dapat melakukan proses klasifikasi terumbu karang. Setelah data citra terumbu karang disiapkan dan di proses, maka selanjutnya akan melakukan proses pelatihan model.

Setelah model dilakukan proses pelatihan untuk dapat mengenali pola dari data, maka model tersebut dilakukan proses pengujian dengan menggunakan data pengujian agar dapat diketahui performa model yang dibangun. Untuk hasil pengujian untuk proses klasifikasi email phising pada penelitian ini mendapatkan akurasi sebesar 97% dan dengan nilai nilai perform jika dihitung dengan confusion matrix yang sudah didapatkan diberikan pada Tabel 1.

Tabel 1. Nilai confusion matrix

Keterangan	Nilai
Akurasi	97%
Presisi	97%
Recall	94%
F1-Score	95%

Tabel 1 menunjukkan hasil klasifikasi dengan menggunakan algoritma SVM. Dapat terlihat dari tabel 1 bahwa model yang dibangun memiliki nilai keakuratan prediksi yang sangat baik yaitu sebesar 97%, nilai ketepatan prediksi pada semua kelas yang sangat baik juga yaitu sebesar 94% dan nilai harmonic yang baik antara presisi dan recall yaitu sebesar 95%. Dari nilai nilai tersebut, maka dapat dilihat bahwa model yang dibangun dapat melakukan proses klasifikasi dengan baik dan dapat menghasilkan keakuratan prediksi sehingga dapat dengan baik membantu dalam melakukan proses klasifikasi *email phising.*

SIMPULAN DAN SARAN

Setelah dilakukan proses pelatihan dan pengujian dengan menggunakan algoritma SVM untuk melakukan proses klasifikasi email phising maka dapat disimpulkan bahwa metode SVM dapat dengan baik melakukan proses klasifikasi data terutama email phising sehingga dapat membantu Masyarakat untuk dapat tidak tertipu karena phising lagi.

Pada penelitian selanjutnya, diharapkan untuk dapat melakukan proses klasifikasi dengan model lainnya seperti naïve bayes classifier, random forest, KNN, decision tree, dan lainnya sehingga dapat dilakukan proses perbandingan metode mana yang terbaik untuk melakukan proses klasifikasi email phising.

DAFTAR PUSTAKA

Salloum, S., Gaber, T., Vadera, S., & Shaalan, K. (2022). *A Systematic Literature Review on Phishing Email Detection Using Natural Language Processing Techniques*. In IEEE Access (Vol. 10, pp. 65703–65727). Institute of Electrical and Electronics Engineers Inc. https://doi.org/10.1109/ACCESS.2022.3183083

Alhogail, A., & Alsabih, A. (2021). *Applying machine learning and natural language processing to detect phishing email*. *Computers and Security*, 110. https://doi.org/10.1016/j.cose.2021.102414

Bi, Q., Goodman, K. E., Kaminsky, J., & Lessler, J. (2019). *What is machine learning? A primer for the epidemiologist*. *American Journal of Epidemiology*, 188(12), 2222–2239. https://doi.org/10.1093/aje/kwz189

Sah, S. (2020). *Machine Learning: A Review of Learning Types*. https://doi.org/10.20944/preprints202007.0230.v1

Boateng, E. Y., Otoo, J., & Abaye, D. A. (2020). *Basic Tenets of Classification Algorithms K-Nearest-Neighbor, Support Vector Machine, Random Forest and Neural Network: A Review*. *Journal of Data Analysis and Information Processing*, 08(04), 341–357. https://doi.org/10.4236/jdaip.2020.84020

Muthukrishnan, S., Krishnaswamy, H., Thanikodi, S., Sundaresan, D., & Venkatraman, V. (2020). *Support vector machine for modelling and simulation of heat exchangers*. Thermal Science, 24(1PartB), 499–503. https://doi.org/10.2298/TSCI190419398M

V. Kumar and B. Subba, "*A TfidfVectorizer and SVM based sentiment analysis framework for text data corpus," 2020 National Conference on Communications (NCC),* Kharagpur, India, 2020, pp. 1-6, doi: 10.1109/NCC48643.2020.9056085.

Akuma, S., Lubem, T., & Adom, I. T. (2022). *Comparing Bag of Words and TF-IDF with different models for hate speech detection from live tweets*. International Journal of Information Technology (Singapore), 14(7), 3629–3635. https://doi.org/10.1007/s41870-022-01096-4

Krstinić, D., Braović, M., Šerić, L., & Božić-Štulić, D. (2020). *Multi-label Classifier Performance Evaluation with Confusion Matrix*. 01–14. https://doi.org/10.5121/csit.2020.100801

Khairandish, M. O., Sharma, M., Jain, V., Chatterjee, J. M., & Jhanjhi, N. Z. (2022). *A Hybrid CNN-SVM Threshold Segmentation Approach for Tumor Detection and Classification of MRI Brain Images*. IRBM, 43(4), 290–299. https://doi.org/10.1016/j.irbm.2021.06.003

Styawati, S., & Mustofa, K. (2019). *A Support Vector Machine-Firefly Algorithm for Movie Opinion Data Classification*. IJCCS (Indonesian Journal of Computing and Cybernetics Systems), 13(3), 219. https://doi.org/10.22146/ijccs.41302

Vijayarajeswari, R., Parthasarathy, P., Vivekanandan, S., & Basha, A. A. (2019*). Classification of mammogram for early detection of breast cancer using SVM classifier and Hough transform. Measurement: Journal of the International Measurement Confederation*, 146, 800–805. https://doi.org/10.1016/j.measurement.2019.05.083

Ma, T. M., Yamamori, K., & Thida, A. (2020). *A Comparative Approach to Naïve Bayes Classifier and Support Vector Machine for Email Spam Classification.* 2020 IEEE 9th Global Conference on Consumer Electronics, GCCE 2020, 324–326. https://doi.org/10.1109/GCCE50665.2020.9291921