

PENERAPAN DATA MINING MENGGUNAKAN METODE C4.5 PADA DATA NASABAH KREDIT MACET

Tri Haryanti
Informatika
triharyanti@gmail.com

Abstrak

Pada tahun 2015 jumlah total keseluruhan kredit macet yang ada pada BPR Tataarta Swadaya Rumbia Lampung Tengah yaitu mencapai 10 % dari 129 data nasabah, total keseluruhan dana kredit yang telah dikeluarkan selama tahun 2015, sehingga akan membebankan pihak bank dan pendapatan bank tersebut akan berkurang secara berkala, untuk itu pihak bank harus lebih teliti lagi dalam pemilihan calon nasabah yang akan mengajukan kredit kepada bank. Faktor – faktor yang mempengaruhi kredit macet menjadi aspek penting untuk mengetahui potensi data kredit macet yang akan datang. Faktor yang mempengaruhi kredit macet pada BPR Tataarta Swadaya Rumbia Lampung Tengah di antaranya faktor internal bank dan eksternal bank. Faktor internal yaitu terkait dengan analisis kredit, sedangkan faktor eksternal meliputi tunggakan pokok bunga, usaha, penghasilan. Perkiraan klasifikasi dapat dilakukan dengan berbagai teknik, salah satu model yang digunakan untuk pengklasifikasian adalah menggunakan teknik data mining dengan menerapkan algoritma C4.5 (Decision Tree). Algoritma C4.5 dipilih untuk membentuk pohon keputusan yang dapat diartikan suatu cara untuk mengklasifikasi data yang sangat kuat. Pohon keputusan dapat membagi kumpulan data yang besar menjadi himpunan-himpunan record yang lebih kecil dengan menerapkan serangkaian aturan keputusan.

Kata Kunci: Kredit Macet, Metode C4.5, Data Mining.

PENDAHULUAN

Kredit macet adalah ketika nasabah tidak mampu membayar sebagian atau seluruh kewajibannya kepada bank sesuai kesepakatan (Alfiah & Damayanti, 2020);(Wibowo & Priandika, 2021). Pada tahun 2015 jumlah total keseluruhan kredit macet yang ada pada BPR Tataarta Swadaya Rumbia Lampung Tengah yaitu mencapai 10 % dari 129 data nasabah, total keseluruhan dana kredit yang telah dikeluarkan selama tahun 2015. Bank perlu lebih berhati-hati dalam memilih orang yang tertarik untuk mengajukan pinjaman kepada mereka, karena ini membebani bank dan memastikan bahwa pendapatan mereka menurun secara teratur (Damayanti et al., 2019);(Nurkholis & Nurkholis, 2021);(Ahluwalia et al., 2021). Pihak bank harus lebih berhati-hati dan tepat ketika membuat keputusan tentang prospek untuk mengajukan kredit atau batas kredit, analisis harus mempertimbangkan beberapa faktor, seperti tunggakan pokok dan bunga, untuk dapat membuat keputusan tentang klasifikasi (Pajar et al., 2017);(Puspitasari & Budiman, 2021);(Rahman Isnain, Indra Sakti, et al., 2021). Pengalaman pelanggan kesepakatan buruk, bank perlu memperhatikan klasifikasi dan perhitungannya (Alita et al.,

2020);(Abidin et al., 2021);(Alita, Putra, et al., 2021). Pada proses pengklasifikasian terdapat potensi terjadinya kesalahan pengelompokkan, karena proses penentuan masih dilakukan secara manual dengan cara melihat beberapa atribut sebagai bahan acuan untuk mengelompokkan data kredit macet ke dalam tiga kategori yaitu : kredit macet rendah, sedang, dan tinggi (Aldino, Saputra, et al., 2021);(Neneng et al., 2016);(Neneng & Fernando, 2017).

KAJIAN PUSTAKA

Algoritma C.45

Algoritma C.45 merupakan algoritma untuk membentuk pohon keputusan (Nurkholis et al., 2021);(Sari & Isnaini, 2021). Pohon keputusan dapat diartikan sebagai metode prediksi atau klasifikasi yang sangat kuat (Neneng et al., 2021);(Nurkholis & Sitanggang, 2020);(Hendrastuty, 2021). Pohon keputusan dapat membagi kumpulan data besar menjadi kumpulan data yang lebih kecil dengan menerapkan seperangkat aturan keputusan (Anestiviya et al., 2021);(Sulistiani & Aldino, 2020);(Rahmanto et al., 2021). Algoritma C4.5 adalah nilai atribut dapat diprediksi dengan menggunakan algoritma berupa aturan-aturan yang umum diketahui dan digunakan untuk klasifikasi (Aldino & Sulistiani, 2020);(Febriani & Sulistiani, 2021);(Alita, Setiawansyah, et al., 2021). Pencarian Entropy dan Gain menggunakan rumus, untuk Gain digunakan rumus sebagai berikut :

$$Gain(S,A) = Entropy(S) - \sum_{i=1}^n \left| \frac{S_i}{S} \right| * Entropy(S_i) \quad (1)$$

Keterangan :

S : Himpunan Kasus

A : Atribut

N : Jumlah Partisi Atribut

|Si| : Jumlah Kasus Pada Partisi Ke -I

|S| : Jumlah Kasus Dalam S

Penerapan

Penerapan adalah tindakan mempraktikkan teori, metode, dan hal-hal lain untuk mencapai tujuan tertentu, dan tindakan merencanakan dan melaksanakan terlebih dahulu untuk keuntungan yang diinginkan oleh kelompok atau kelompok (Borman et al., 2020);(Nurkholis et al., 2022);(Abidin et al., 2022). Proses tindakan yang dilakukan baik

individu maupun kelompok untuk mencapai tujuan yang telah dirumuskan (Septilia et al., 2020);(Wantoro & Nata Prawira, n.d.). Adapun unsur-unsur penerapan meliputi :

- Adanya program yang dilaksanakan
- Adanya kelompok target, yaitu masyarakat yang menjadi sasaran dan diharapkan akan menerima manfaat dari program tersebut.

Adanya pelaksanaan, baik organisasi atau perorangan yang bertanggung jawab dalam pengelolaan, pelaksanaan maupun pengawasan dari proses penerapan tersebut.

Data Mining

Data mining adalah proses penambangan data untuk menemukan pola penting yang dapat menghasilkan informasi yang berguna (Qomariah & Sucipto, 2021);(Aldino, Darwis, et al., 2021). Istilah yang digunakan untuk menemukan pengetahuan yang tersembunyi dalam database (Rahman Isnain, Pasha, et al., 2021);(*Comparison of Support Vector Machine and Naïve Bayes on Twitter Data Sentiment Analysis*, 2021);(Sanger et al., 2021). Data mining adalah proses semi-otomatis yang menggunakan statistik, matematika, kecerdasan buatan, dan teknik pembelajaran mesin untuk mengekstrak dan mengidentifikasi informasi pengetahuan yang berpotensi berharga dan berguna yang disimpan dalam database besar (Ahmad et al., 2021);(Oktaviani, 2021);(Styawati & Mustofa, 2019).

METODE

Tahapan Pengumpulan Data

Pengumpulan data adalah kegiatan mencari data dalam suatu bidang yang akan digunakan untuk menjawab suatu pertanyaan penelitian (Aldino, Hendra, et al., 2021). Memperoleh data berkualitas tinggi memerlukan efektivitas pengumpulan data dan kualifikasi pengumpul data. Tahapan penelitian merupakan kegiatan penelitian yang direncanakan dan dilakukan secara teratur dan sistematis untuk mencapai suatu tujuan tertentu (Styawati et al., 2021). Tahap penelitian ini juga merupakan pengembangan dari kerangka penelitian, yang selanjutnya dibagi menjadi beberapa submenu.

Tinjauan Pustaka

Metode pengumpulan data atau informasi yang dilakukan dengan cara mempelajari berbagai laporan-laporan ilmiah, jurnal penelitian dan dokumen atau sumber bacaan serta buku-buku referensi yang berkaitan atau berhubungan dengan topik usulan penelitian yang sedang diteliti.

Analisis dan Pemetaan

Pada tahapan ini menganalisis tentang faktor data nasabah kredit macet menggunakan metode Algoritma C4.5

Pengembangan

Tahapan ini adalah pengembangan terhadap data nasabah kredit macet.

Validasi

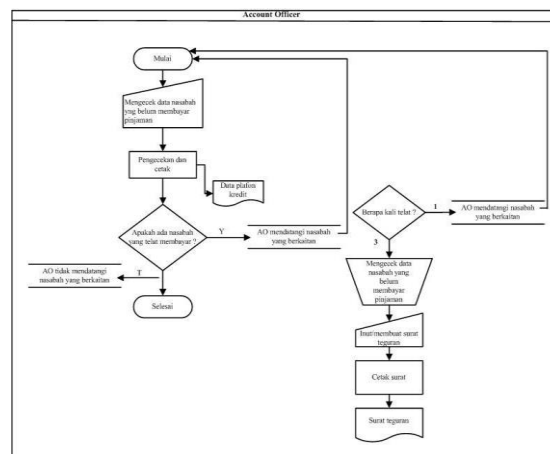
Pada tahap ini dilakukan pengujian aplikasi menggunakan pengujian manual dan mape.

Penutup

Pada tahap ini berisi tentang kesimpulan dan saran terhadap Penerapan Data Mining Menggunakan Metode C4.5 pada Data Nasabah Kredit Macet.

HASIL DAN PEMBAHASAN

Flowchart pengecekan tunggakan nasabah

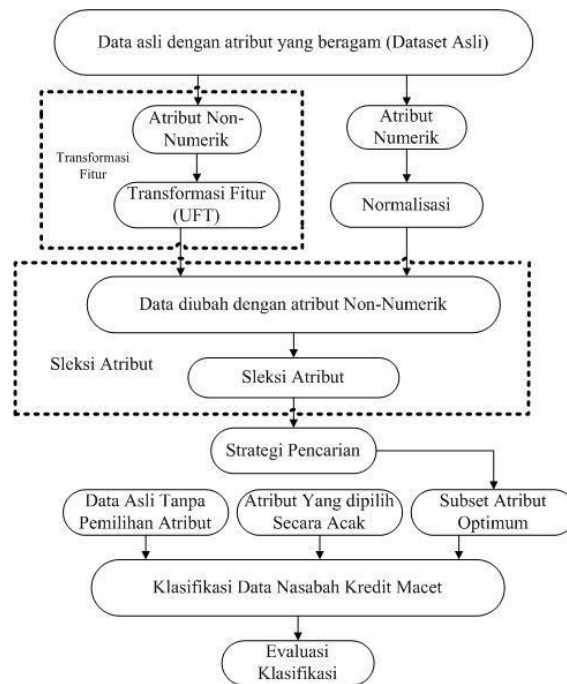


Gambar 1. Flowchart pengecekan tunggakan

Klasifikasi Data Nasabah Kredit Macet

Pengujian sistem dilakukan untuk mengetahui tingkat kelayakan penggunaan aplikasi. Pengujian didasarkan pada model kotak hitam dan metode pengujian penerimaan pengguna.

Kerangka Pengujian



Gambar 2 Kerangka Pengujian

Berdasarkan latar belakang yang telah dipaparkan sebelumnya, penelitian ini menerapkan model pohon keputusan untuk menganalisis data nasabah kredit macet yang bergerak cepat dengan menambahkan transformasi atribut dan pemilihan atribut guna akurasi kinerja yang lebih baik. Pohon keputusan adalah diagram alur seperti pohon di mana setiap simpul bagian dalam mewakili atribut atau subset fitur yang akan diuji, setiap cabang mewakili hasil dari atribut pengujian, dan simpul daun mewakili kelas atau distribusi kelas tertentu.

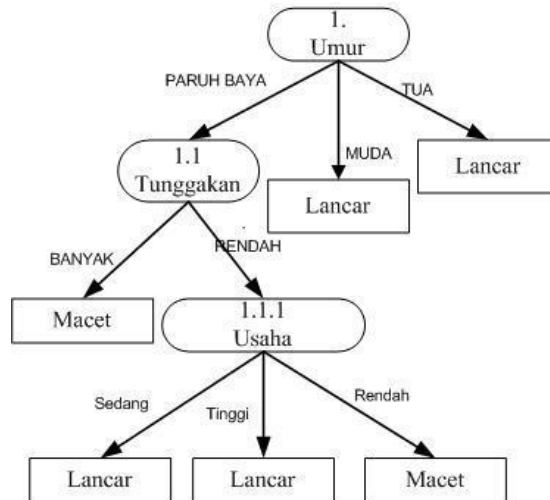
Skenario Uji Coba dan Analisis Hasil

Pada bagian ini akan dijelaskan mengenai skenario uji coba, skenario uji coba merupakan rencana uji coba. Sehingga analisis dari uji coba yang dilakukan dapat menjawab rumusan masalah dan tujuan yang telah ditetapkan sebelumnya. Penelitian ini berfokus pada pemilihan atribut pada tahap pra proses terhadap metode klasifikasi. Skenario ini adalah melakukan klasifikasi data nasabah kredit macet menggunakan metode pohon keputusan berdasarkan:

- Menggunakan seluruh atribut dari data aslinya (sebelum dilakukan pemilihan atribut)
- Menggunakan metode pemilihan atribut yang diusulkan;

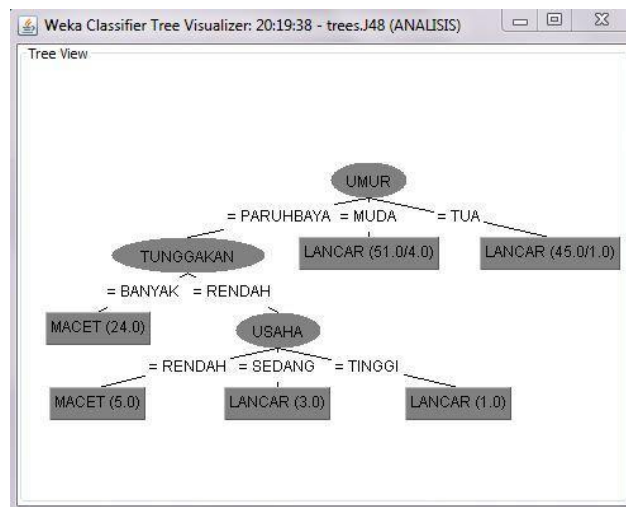
c. Menggunakan atribut yang dipilih berdasarkan perkiraan peneliti.

Hasil Pembahasan



Gambar 3. Hasil manual

Hasil menggunakan Weka



Gambar 4. Hasil Weka

SIMPULAN DAN SARAN

Kesimpulan

Kesimpulan yang dapat diambil dari penelitian yang dilakukan sebagai berikut.

Untuk pengklasifikasian nasabah kredit macet berdasarkan data karakteristik yang tersedia menggunakan algoritma C4.5 yang diklasifikasikan berdasarkan data training dengan atribut umur (paruh baya, muda, tua), tunggakan (banyak, rendah), usaha (rendah, sedang, tinggi), analisis kredit (sedang, rendah, tinggi), pekerjaan (petani, pedagang), penghasilan (rendah, sedang, tinggi). Dari hasil pengolahan dan uji coba menggunakan pohon keputusan dari 129 data dengan enam atribut terpilih maka didapatkan hasil 124 data (96,124%) dapat diklasifikasikan dengan benar, sedangkan 5 data (3,876%) salah diklasifikasikan.

Untuk menyajikan informasi klasifikasi kredit macet di BPR Tataarta Swadaya dilakukan dengan cara membuat pohon keputusan menggunakan data training yang akan menghasilkan pohon keputusan. Untuk mengidentifikasi dan mengklasifikasi data nasabah kredit macet berdasarkan karakter yang tersedia dengan algoritma C4.5.

REFERENSI

- Abidin, Z., Amelia, D., & Aguss, R. M. (2022). *PELATIHAN GOOGLE APPS UNTUK MENAMBAH KEAHLIAN TEKNOLOGI INFORMASI BAGI GURU SMK PGRI 1 LIMAU*. 3(1), 43–48.
- Abidin, Z., Wijaya, A., & Pasha, D. (2021). Aplikasi Stemming Kata Bahasa Lampung Dialek Api Menggunakan Pendekatan Brute-Force dan Pemograman C. *JURNAL MEDIA INFORMATIKA BUDIDARMA*, 5(1), 1–8.
- Ahluwalia, L., Permatasari, B., Husna, N., & Novita, D. (2021). *Penguatan Sumber Daya Manusia Melalui Peningkatan Keterampilan Pada Komunitas ODAPUS Lampung*. 2(1), 73–80. <https://doi.org/10.23960/jpkmt.v2i1.32>
- Ahmad, I., Borman, R. I., Caksana, G. G., & Fakhrurozi, J. (2021). IMPLEMENTASI STRING MATCHING DENGAN ALGORITMA BOYER-MOORE UNTUK MENENTUKAN TINGKAT KEMIRIPAN PADA PENGAJUAN JUDUL SKRIPSI/TA MAHASISWA (STUDI KASUS: UNIVERSITAS XYZ). *SINTECH (Science and Information Technology) Journal*, 4(1), 53–58.
- Aldino, A. A., Darwis, D., Prastowo, A. T., & Sujana, C. (2021). Implementation of K-means algorithm for clustering corn planting feasibility area in south lampung

regency. *Journal of Physics: Conference Series*, 1751(1), 12038.

Aldino, A. A., Hendra, V., & Darwis, D. (2021). Pelatihan Spada Sebagai Optimalisasi Lms Pada Pembelajaran Di Masa Pandemi Covid 19. *Journal of Social Sciences and Technology for Community Service (JSSTCS)*, 2(2), 72. <https://doi.org/10.33365/jsstcs.v2i2.1330>

Aldino, A. A., Saputra, A., & Nurkholis, A. (2021). *Application of Support Vector Machine (SVM) Algorithm in Classification of Low-Cape Communities in Lampung Timur*. 3(3), 325–330. <https://doi.org/10.47065/bits.v3i3.1041>

Aldino, A. A., & Sulistiani, H. (2020). Decision Tree C4. 5 Algorithm For Tuition Aid Grant Program Classification (Case Study: Department Of Information System, Universitas Teknokrat Indonesia). *Edutic-Scientific Journal of Informatics Education*, 7(1).

Alfiah, & Damayanti. (2020). Aplikasi E-Marketplace Penjualan Hasil Panen Ikan Lele (Studi Kasus: Kabupaten Pringsewu Kecamatan Pagelaran). *Jurnal Teknologi Dan Sistem Informasi (JTSI)*, 1(1), 111–117. <http://jim.teknokrat.ac.id/index.php/sisteminformasi>

Alita, D., Fernando, Y., & Sulistiani, H. (2020). Implementasi Algoritma Multiclass SVM pada Opini Publik Berbahasa Indonesia di Twitter. *Jurnal Tekno Kompak*, 14(2), 86–91.

Alita, D., Putra, A. D., & Darwis, D. (2021). Analysis of classic assumption test and multiple linear regression coefficient test for employee structural office recommendation. *IJCCS (Indonesian Journal of Computing and Cybernetics Systems)*, 15(3), 1–5.

Alita, D., Setiawansyah, S., & ... (2021). C45 Algorithm for Motorcycle Sales Prediction On CV Moka Rawajitu. *Jurnal Sisfotek ...*, 11(2), 127–134. <http://journal.stmikglobal.ac.id/index.php/sisfotek/article/view/392>

Anestiviya, V., Ferico, A., & Pasaribu, O. (2021). Analisis Pola Menggunakan Metode C4.5 Untuk Peminatan Jurusan Siswa Berdasarkan Kurikulum (Studi Kasus : Sman 1

- Natar). *Jurnal Teknologi Dan Sistem Informasi (JTISI)*, 2(1), 80–85.
<http://jim.teknokrat.ac.id/index.php/JTISI>
- Borman, R. I., Megawaty, D. A., & Attohiroh, A. (2020). Implementasi Metode TOPSIS Pada Sistem Pendukung Keputusan Pemilihan Biji Kopi Robusta Yang Bernilai Mutu Ekspor (Studi Kasus: PT. Indo Cafco Fajar Bulan Lampung). *Fountain of Informatics Journal*, 5(1), 14–20.
- Damayanti, Warsito, Meilinda, Manurung, P., & Sembiring, S. (2019). E-crm Information System for Tapis Lampung SMEs. *Journal of Physics: Conference Series*, 1338(1).
<https://doi.org/10.1088/1742-6596/1338/1/012051>
- Febriani, S., & Sulistiani, H. (2021). Analisis Data Hasil Diagnosa Untuk Klasifikasi Gangguan Kepribadian Menggunakan Algoritma C4. 5. *Jurnal Teknologi Dan Sistem Informasi*, 2(4), 89–95.
- Hendrastuty, N. (2021). *Text Summarization in Multi Document Using Genetic Algorithm*. 15(4), 327–338.
- Comparison of Support Vector Machine and Naïve Bayes on Twitter Data Sentiment Analysis, (2021).
- Neneng, N., Adi, K., & Isnanto, R. (2016). Support Vector Machine Untuk Klasifikasi Citra Jenis Daging Berdasarkan Tekstur Menggunakan Ekstraksi Ciri Gray Level Co-Occurrence Matrices (GLCM). *JSINBIS (Jurnal Sistem Informasi Bisnis)*, 6(1), 1–10.
- Neneng, N., & Fernando, Y. (2017). Klasifikasi Jenis Daging Berdasarkan Analisis Citra Tekstur Gray Level Co-Occurrence Matrices (GlcM) Dan Warna. *Prosiding Semnastek*.
- Neneng, N., Puspaningrum, A. S., & Aldino, A. A. (2021). Perbandingan Hasil Klasifikasi Jenis Daging Menggunakan Ekstraksi Ciri Tekstur Gray Level Co-occurrence Matrices (GLCM) Dan Local Binary Pattern (LBP). *SMATIKA JURNAL*, 11(01), 48–52.
- Nurkholis, A., Anggela, Y., & Octaviansyah P, A. F. (2022). Web-Based Geographic Information System for Lampung Gift Store. *Jurnal Teknoinfo*, 16(1), 34.

<https://doi.org/10.33365/jti.v16i1.1486>

- Nurkholis, A., & Nurkholis, A. (2021). *Prediction Model for Soybean Land Suitability Using C5.0 Algorithm*. 6(2), 163–171. <https://doi.org/10.15575/join.v6i2.711>
- Nurkholis, A., & Sitanggang, I. S. (2020). Optimization for prediction model of palm oil land suitability using spatial decision tree algorithm. *Jurnal Teknologi Dan Sistem Komputer*, 8(3), 192–200. <https://doi.org/10.14710/jtsiskom.2020.13657>
- Nurkholis, A., Sitanggang, I. S., Annisa, & Sobir. (2021). Spatial decision tree model for garlic land suitability evaluation. *IAES International Journal of Artificial Intelligence*, 10(3), 666–675. <https://doi.org/10.11591/ijai.v10.i3.pp666-675>
- Oktaviani, L. (2021). Penerapan Sistem Pembelajaran Dalam Jaringan Berbasis Web Pada Madrasah Aliyah Negeri 1 Pesawaran. *Jurnal WIDYA LAKSMI (Jurnal Pengabdian Kepada Masyarakat)*, 1(2), 68–75.
- Pajar, M., Setiawan, D., Rosandi, I. S., & Darmawan, S. (2017). *Deteksi Bola Multipola Pada Robot Krakatau FC*. 6–9.
- Puspitasari, M., & Budiman, A. (2021). Perancangan Sistem Informasi Manajemen Perpustakaan Menggunakan Metode Fast (Framework for the Application System Thinking) (Studi Kasus : Sman 1 Negeri Katon). *Jurnal Teknologi Dan Sistem Informasi (JTSI)*, 2(2), 69–77. <http://jim.teknokrat.ac.id/index.php/JTSI>
- Qomariah, L., & Sucipto, A. (2021). Sistem Infomasi Surat Perintah Tugas Menggunakan Pendekatan Web Engineering. *JTSI-Jurnal Teknologi Dan Sistem Informasi*, 2(1), 86–95.
- Rahman Isnain, A., Indra Sakti, A., Alita, D., & Satya Marga, N. (2021). Sentimen Analisis Publik Terhadap Kebijakan Lockdown Pemerintah Jakarta Menggunakan Algoritma Svm. *Jdmsi*, 2(1), 31–37. <https://t.co/NfhmfMjtXw>
- Rahman Isnain, A., Pasha, D., & Sintaro, S. (2021). Workshop Digital Marketing “Temukan Teknik Pemasaran Secara Daring.” *Journal of Social Sciences and Technology for Community Service (JSSTCS)*, 2(2), 113–120. <https://ejurnal.teknokrat.ac.id/index.php/JSSTCS/article/view/1365>

- Rahmanto, Y., Burlian, A., & Samsugi, S. (2021). SISTEM KENDALI OTOMATIS PADA AKUAPONIK BERBASIS MIKROKONTROLER ARDUINO UNO R3. *Jurnal Teknologi Dan Sistem Tertanam*, 2(1), 1–6.
- Sanger, J. B., Sitanayah, L., & Ahmad, I. (2021). A Sensor-based Garbage Gas Detection System. *2021 IEEE 11th Annual Computing and Communication Workshop and Conference, CCWC 2021, February*, 1347–1353. <https://doi.org/10.1109/CCWC51732.2021.9376147>
- Sari, R. K., & Isnaini, F. (2021). PERANCANGAN SISTEM MONITORING PERSEDIAAN STOK ES KRIM CAMPINA PADA PT YUNIKAR JAYA SAKTI. *Jurnal Informatika Dan Rekayasa Perangkat Lunak*, 2(1), 151–159.
- Septilia, H. A., Parjito, P., & Styawati, S. (2020). Sistem Pendukung Keputusan Pemberian Dana Bantuan menggunakan Metode AHP. *Jurnal Teknologi Dan Sistem Informasi*, 1(2), 34–41.
- Styawati, Andi Nurkholis, Zaenal Abidin, & Heni Sulistiani. (2021). Optimasi Parameter Support Vector Machine Berbasis Algoritma Firefly Pada Data Opini Film. *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, 5(5), 904–910. <https://doi.org/10.29207/resti.v5i5.3380>
- 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), 219–230.
- Sulistiani, H., & Aldino, A. A. (2020). Decision Tree C4.5 Algorithm for Tuition Aid Grant Program Classification (Case Study: Department of Information System, Universitas Teknokrat Indonesia). *Edutic - Scientific Journal of Informatics Education*, 7(1), 40–50. <https://doi.org/10.21107/edutic.v7i1.8849>
- Wantoro, A., & Nata Prawira, F. (n.d.). *Implementation of Simple Additive Weighting (SAW) Method for Determining Social Customer Relationship Management (SCRM) Model as Business Strategy in University*.
- Wibowo, D. O., & Priandika, A. T. (2021). SISTEM PENDUKUNG KEPUTUSAN

PEMILIHAN GEDUNG PERNIKAHAN PADA WILAYAH BANDAR LAMPUNG
MENGUNAKAN METODE TOPSIS. *Jurnal Informatika Dan Rekayasa Perangkat Lunak*, 2(1), 73–84.