

Klasifikasi Kerusakan Komputer Menggunakan Support Vector Machine pada Orange Data Mining

Alvino Tunjung Bagaskara¹, Feri Sulianta²

^{1,2} Fakultas Teknik, Universitas Widyatama, Bandung

E-mail korespondensi: ¹ alvino.tunjung@widyatama.ac.id

Keywords: *classification, data mining, Orange Data Mining, service repair, support vector machine*

Abstract

Computer repair services at small and medium enterprises (SMEs) often struggle to estimate damage categories before physically inspecting devices, reducing service efficiency and customer transparency. This study develops a Support Vector Machine (SVM) classification model implemented through Orange Data Mining to predict device damage categories from initial customer-reported information. The dataset comprises 58 service transactions from a Bandung-based repair business. Preprocessing included keyword matching to convert free-text complaints into structured categories, yielding five input features. Evaluated via 10-fold stratified cross-validation, the model achieved accuracy, AUC, and F1-Score above 0.89, with physical condition and complaint type as the most influential features. Results demonstrate that SVM via Orange Data Mining is effectively applicable to SME-scale operational data, enabling earlier damage estimation and improved service workflow.

Kata kunci: *klasifikasi, data mining, Orange Data Mining, servis komputer, support vector machine*

Abstrak

Layanan servis perangkat komputer pada usaha kecil dan menengah (UKM) sering kesulitan memberikan estimasi kategori kerusakan sebelum perangkat diperiksa langsung, sehingga proses servis menjadi kurang efisien dan kurang transparan bagi pelanggan. Penelitian ini membangun model klasifikasi berbasis Support Vector Machine (SVM) melalui Orange Data Mining untuk memprediksi kategori kerusakan berdasarkan informasi awal yang diberikan pelanggan. Dataset terdiri dari 58 transaksi servis pada sebuah usaha servis komputer di Bandung. Preprocessing meliputi keyword matching untuk mengubah keluhan teks bebas menjadi kategori terstruktur, menghasilkan lima fitur input. Model dievaluasi menggunakan 10-fold stratified cross-validation dan mencapai akurasi, AUC, serta F1-Score di atas 0,89, dengan fitur kondisi fisik dan tipe keluhan sebagai yang paling berpengaruh. Hasil penelitian menunjukkan bahwa SVM melalui Orange Data Mining efektif diterapkan pada data operasional servis komputer skala UKM untuk mendukung estimasi kerusakan lebih awal dan meningkatkan efisiensi layanan.

PENDAHULUAN

Layanan servis perangkat komputer pada skala usaha kecil dan menengah (UKM) sering menghadapi kendala dalam memberikan estimasi kategori kerusakan secara akurat sebelum perangkat diperiksa langsung. Pada praktiknya di Srikaton Inovasi Teknologi, sebuah usaha servis

laptop dan PC di Bandung, teknisi harus melakukan pemeriksaan fisik terlebih dahulu sebelum menentukan jenis kerusakan dan menyiapkan komponen yang dibutuhkan. Kondisi ini menyebabkan proses servis kurang efisien, persiapan spare part sering terlambat, dan pelanggan tidak memperoleh informasi yang memadai di awal kunjungan. Penerapan teknik data mining dapat menjadi solusi dengan membangun model klasifikasi yang mampu memprediksi kategori kerusakan berdasarkan informasi yang tersedia saat pelanggan pertama kali datang. Orange Data Mining dipilih sebagai platform implementasi karena antarmuka berbasis widget yang memudahkan seluruh tahapan klasifikasi, mulai dari persiapan data hingga evaluasi model (Dobesova, 2024).

Di antara berbagai algoritma klasifikasi, Support Vector Machine (SVM) dipilih karena kemampuannya menemukan hyperplane pemisah optimal antar kelas, sehingga efektif digunakan pada dataset berukuran kecil dengan campuran fitur kategorikal dan numerik. Beberapa penelitian sebelumnya telah membuktikan keunggulan SVM di berbagai domain: Nurqotimah dkk. (2024) mencapai akurasi 94,8% dalam klasifikasi penyakit stroke menggunakan Orange Data Mining dengan cross-validation; Pranadjaya dkk. (2024) menggunakan Orange untuk membandingkan beberapa algoritma termasuk SVM pada klasifikasi jenis kendaraan untuk sistem tilang digital; Syahri & Puspita (2025) mencapai akurasi 89,80% dalam klasifikasi mahasiswa berprestasi; serta Zulherry dkk. (2024) menerapkan SVM untuk klasifikasi data Tracer Study alumni sebagai bahan pengambilan keputusan institusi. Meskipun SVM terbukti konsisten menghasilkan performa yang baik di berbagai bidang, penerapannya untuk memprediksi kategori kerusakan pada layanan servis komputer skala UKM belum pernah diteliti sebelumnya.

Berdasarkan kesenjangan tersebut, penelitian ini bertujuan membangun model klasifikasi berbasis SVM menggunakan Orange Data Mining pada data transaksi servis Srikaton Inovasi Teknologi periode Oktober–Desember 2024. Model yang dibangun diharapkan mampu memprediksi kategori kerusakan — Hardware, Software, Peripheral, atau Maintenance — hanya dari informasi awal pelanggan, sehingga teknisi dapat memberikan estimasi lebih cepat dan proses servis menjadi lebih efisien.

METODE

Dataset yang digunakan merupakan data transaksi servis perangkat komputer di Srikaton Inovasi Teknologi, Bandung, periode Oktober–Desember 2024, yang bersumber dari sistem pencatatan operasional harian perusahaan. Total dataset berjumlah 58 transaksi dengan 12 atribut awal dalam format CSV, kelengkapan data 100% tanpa missing values. Variabel target yang diprediksi adalah kategori_kerusakan dengan empat kelas: Hardware, Software, Peripheral, dan Maintenance.

Tabel 1. Data Dictionary

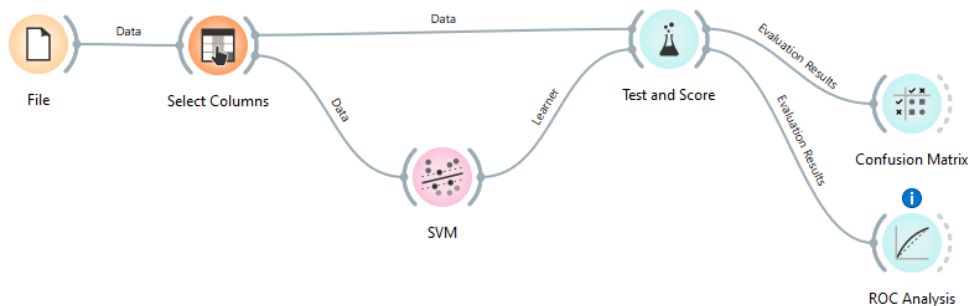
NO	Nama Kolom	Tipe Data	Deskripsi	Nilai
1	id_transaksi	Categorical	Kode transaksi	SRK107, SRK108, SRK109
2	tanggal	Date	Tanggal masuk servis	2024-10-02, 2024-10-07
3	tipe_device	Categorical	Jenis perangkat	Laptop, PC
4	brand	Categorical	Merk perangkat	Asus, Lenovo, HP, Acer, Dell, Rakit
5	usia_bulan	Numeric	Lama penggunaan perangkat	14, 28, 36, 52
6	keluhan_awal	Text	Keluhan user saat pertama datang	"Laptop lambat", "Engsel rusak", "Upgrade SSD"
7	kondisi_fisik	Categorical	Kondisi perangkat saat diterima	Bersih, Berdebu, Rusak Fisik
8	kategori_kerusakan	Categorical	Jenis kerusakan	Hardware, Software, Peripheral, Maintenance
9	detail_perbaikan	text	Rincian perbaikan	"Install ulang Windows 11", "Ganti engsel"
10	harga_total	Numeric	Total biaya servis dalam rupiah	50000, 75000, 200000, 1500000
11	lama_servis	Numeric	Durasi pengerjaan servis	1, 2, 3, 5, 7
12	teknisi	Categorical	Teknisi yang menangani	Teknisi A, Teknisi B, Teknisi C

Tidak seluruh atribut dapat digunakan langsung untuk pemodelan, sehingga dilakukan dua tahap preprocessing menggunakan Python dan Excel. Pertama, kolom keluhan_awal yang berisi teks bebas ditransformasi menjadi tipe_keluhan melalui keyword matching, menghasilkan tujuh kategori terstruktur: Performa, Request_Upgrade, Kerusakan_Fisik, Thermal, System_Error, Request_Software, dan Maintenance_Request. Transformasi ini diperlukan karena SVM tidak dapat memproses teks bebas secara langsung dan variasi ekspresi keluhan perlu distandarisasi agar pola dapat dipelajari secara konsisten. Kedua, dilakukan seleksi atribut sehingga dataset final berisi 6 kolom — kolom id_transaksi, tanggal, teknisi dieliminasi karena tidak memiliki nilai prediktif, dan keluhan_awal dieliminasi karena telah tergantikan oleh tipe_keluhan. Di dalam Orange, widget Select Columns digunakan untuk menetapkan lima kolom sebagai features (tipe_device, brand, usia_bulan, kondisi_fisik, tipe_keluhan) dan kategori_kerusakan sebagai target. Orange secara otomatis melakukan encoding data kategorikal dan normalisasi fitur numerik sebelum training.

type_device	brand	usia_bulan	kondisi_fisik	tipe_keluhan	kategori_kerusakan
Laptop	Asus	28	Berdebu	Performa	Software
Laptop	Lenovo	36	Berdebu	Request_Upgrade	Hardware
PC	Rakit	18	Berdebu	Maintenance_Request	Maintenance
Laptop	HP	42	Rusak Fisik	Kerusakan_Fisik	Peripheral
Laptop	Asus	24	Rusak Fisik	Kerusakan_Fisik	Peripheral
Laptop	Lenovo	31	Berdebu	Kerusakan_Fisik	Peripheral
PC	Rakit	22	Bersih	Request_Upgrade	Hardware
Laptop	Asus	19	Berdebu	System_Error	Software
Laptop	HP	48	Berdebu	Thermal	Maintenance

Gambar 1. Data Dictionary setelah Preprocessing Data

Proses klasifikasi dilakukan menggunakan SVM melalui Orange Data Mining. SVM bekerja dengan mencari hyperplane optimal yang memisahkan data antar kelas dengan margin maksimum, sehingga efektif untuk dataset kecil dengan kombinasi fitur kategorikal dan numerik seperti pada penelitian ini. Workflow Orange yang dibangun terdiri dari lima widget: File → Select Columns → SVM dan Test and Score (paralel) → Confusion Matrix dan ROC Analysis. Model dievaluasi menggunakan 10-fold stratified cross-validation, di mana setiap data point diuji tepat satu kali sambil memaksimalkan data untuk training. Keseluruhan alur workflow dapat dilihat pada Gambar 3.

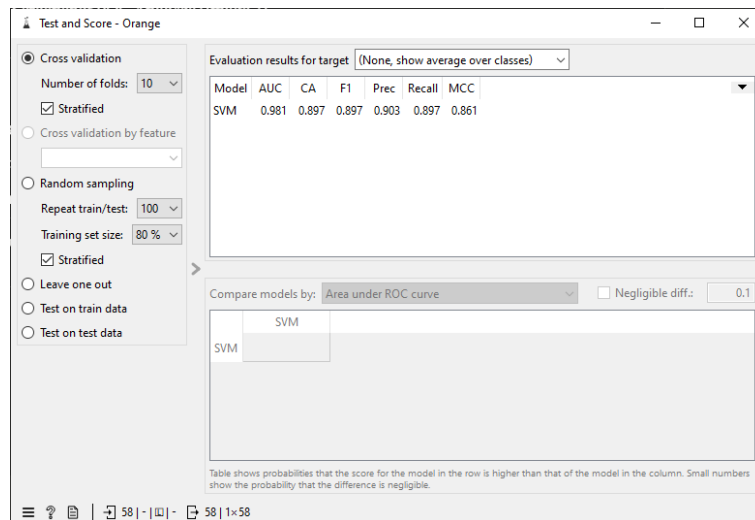


Gambar 2. Implementasi pada Orange Data Mining

HASIL

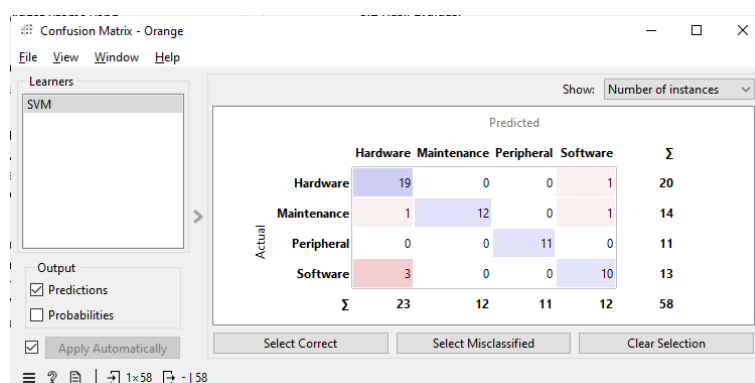
Model SVM dievaluasi menggunakan metode 10-fold stratified cross-validation. Metode ini dipilih karena cocok untuk dataset berukuran kecil seperti yang digunakan dalam penelitian ini (58 transaksi). Pada 10-fold cross-validation, dataset dibagi menjadi 10 bagian (folds) yang sama besar. Setiap fold berisi sekitar 6 transaksi. Proses evaluasi dilakukan sebanyak 10 kali, dimana pada setiap iterasi, 9 folds digunakan untuk training (52 transaksi) dan 1 fold digunakan untuk testing (6 transaksi). Setiap data point di-test tepat satu kali, sehingga semua data berkontribusi

dalam evaluasi namun tidak ada data yang digunakan untuk training dan testing secara bersamaan. Stratified sampling diterapkan untuk memastikan proporsi setiap kategori kerusakan (Hardware, Software, Peripheral, Maintenance) konsisten di setiap fold. Ini penting agar tidak ada fold yang kekurangan atau kelebihan kategori tertentu, yang bisa menyebabkan bias dalam evaluasi.



Gambar 3. Hasil Test and Score

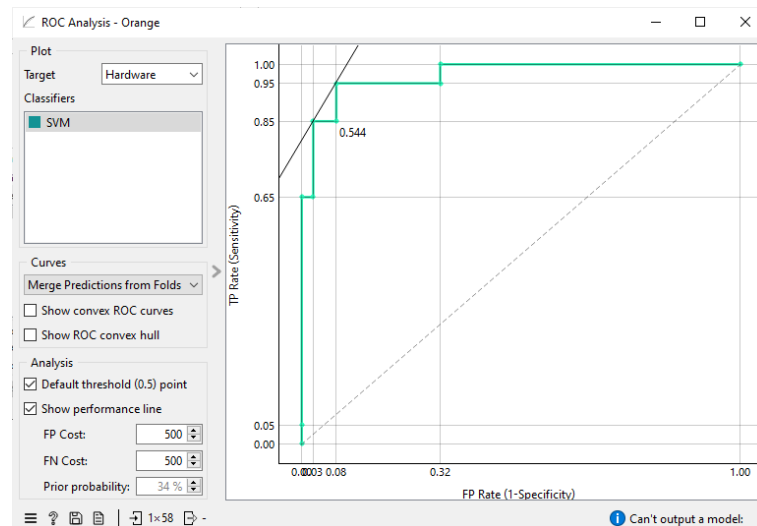
Model mencapai akurasi sebesar 89,7%, yang berarti dari 58 transaksi yang dievaluasi, model berhasil memprediksi kategori kerusakan dengan benar sebanyak 52 transaksi dan salah pada 6 transaksi. Nilai AUC sebesar 0,981 menunjukkan kemampuan model yang sangat baik dalam membedakan antar kategori — nilai yang mendekati 1,0 mengindikasikan model tidak hanya akurat tapi juga cukup yakin dengan prediksinya. F1-Score sebesar 0,897 mencerminkan keseimbangan yang baik antara Precision (0,903) dan Recall (0,897).



Gambar 4. Confusion Matrix

Dari Confusion Matrix terlihat bahwa model menghasilkan 52 prediksi benar yang ditunjukkan oleh angka di diagonal, dan 6 prediksi salah di luar diagonal. Kategori Peripheral menunjukkan performa sempurna dengan 11 dari 11 transaksi diprediksi benar. Kategori Hardware juga sangat

baik dengan 19 dari 20 transaksi benar, hanya 1 transaksi yang salah diprediksi sebagai Software. Kategori Maintenance memiliki 12 dari 14 transaksi benar, dengan 2 kesalahan yaitu 1 diprediksi sebagai Hardware dan 1 sebagai Software. Kategori Software memiliki akurasi paling rendah dengan 10 dari 13 transaksi benar, di mana 3 transaksi Software salah diprediksi sebagai Hardware.



Gambar 5. ROC Analysis

Kurva ROC menunjukkan bahwa garis model berada jauh di atas garis diagonal yang merepresentasikan random classifier, dengan nilai AUC 0,981. Threshold optimal berada di titik 0,544 di mana model mencapai keseimbangan terbaik antara True Positive Rate dan False Positive Rate.

PEMBAHASAN

Model SVM yang dibangun berhasil mencapai akurasi 89,7% dengan AUC 0,981, F1-Score 0,897, Precision 0,903, dan Recall 0,897. Hasil ini menunjukkan bahwa SVM mampu memprediksi kategori kerusakan perangkat komputer hanya dari informasi awal pelanggan dengan cukup baik. Capaian ini sebanding dengan penelitian sebelumnya yang juga menggunakan SVM, akurasi 89,80% dicapai untuk klasifikasi mahasiswa berprestasi (Syahri & Puspita, 2025), dan akurasi 94,8% dicapai untuk klasifikasi penyakit stroke menggunakan SVM melalui Orange Data Mining (Nurqotimah dkk., 2024). Perbedaan nilai akurasi antar penelitian ini wajar karena karakteristik dan ukuran dataset yang berbeda-beda.

Yang menarik adalah perbedaan performa antar kategori yang cukup mencolok. Kategori Peripheral mencapai akurasi 100% karena karakteristiknya tidak ambigu ketika kondisi fisik perangkat tercatat "Rusak Fisik", hampir pasti masuk Peripheral. Sebaliknya, kategori Software hanya mencapai 76,9% dan menjadi yang paling rendah. Ini terjadi karena keluhan seperti "laptop

lemot" bisa masuk ke Software maupun Hardware tergantung solusi yang dipilih, sehingga model kesulitan membedakan keduanya tanpa informasi tambahan. Tiga dari enam total kesalahan prediksi adalah transaksi Software yang salah diprediksi sebagai Hardware.

Dari lima fitur yang digunakan, kondisi_fisik dan tipe_keluhan terbukti paling berpengaruh. kondisi_fisik langsung bisa dilihat saat pelanggan datang dan bersifat sangat decisive, sedangkan tipe_keluhan yang sudah diubah dari teks bebas menjadi kategori terstruktur membantu model memahami maksud pelanggan dengan lebih konsisten. Fitur usia_bulan juga membantu sebagai pendukung perangkat tua dengan keluhan performa lebih cenderung butuh upgrade Hardware, sedangkan perangkat baru dengan keluhan serupa lebih cenderung cukup install ulang Software. Dibandingkan dengan penelitian yang menggunakan Orange untuk klasifikasi gambar kendaraan (Pranadjaya dkk., 2024) dan data akademik (Zulherry dkk., 2024), penelitian ini menunjukkan bahwa SVM melalui Orange Data Mining juga bisa digunakan untuk data transaksi operasional bisnis seperti data servis komputer di UKM — sesuatu yang belum pernah diteliti sebelumnya. Untuk pengembangan ke depan, penambahan fitur budget pelanggan dan pengumpulan data minimal 200 transaksi direkomendasikan untuk meningkatkan akurasi terutama pada kategori Software.

KESIMPULAN

Penelitian ini berhasil membuktikan bahwa algoritma SVM yang diimplementasikan melalui Orange Data Mining mampu mengklasifikasikan kategori kerusakan perangkat komputer hanya berdasarkan informasi awal pelanggan dengan akurasi 89,7%, AUC 0,981, dan F1-Score 0,897. Model ini praktis digunakan karena hanya mengandalkan informasi yang sudah tersedia saat pelanggan pertama kali datang tanpa perlu diagnosis terlebih dahulu. Fitur kondisi_fisik dan tipe_keluhan terbukti paling berpengaruh dalam menentukan prediksi, sementara kategori Software masih menjadi tantangan karena ambiguitas keluhan pelanggan yang bisa mengarah ke dua solusi berbeda. Secara keseluruhan, pendekatan ini menunjukkan bahwa data mining berbasis SVM melalui platform visual seperti Orange dapat diterapkan secara efektif pada data operasional bisnis servis komputer skala UKM, membuka peluang pemanfaatan machine learning di domain yang sebelumnya belum pernah diteliti.

Untuk penelitian selanjutnya, disarankan menambahkan fitur budget pelanggan ke dalam model karena informasi ini berpotensi mengurangi ambiguitas antara kategori Software dan Hardware secara signifikan. Pengumpulan data minimal 200 transaksi selama satu tahun penuh juga direkomendasikan agar model dapat mengenali pola yang lebih beragam. Selain itu, perbandingan performa SVM dengan algoritma lain seperti Random Forest atau Naive Bayes pada dataset serupa dapat menjadi topik penelitian lanjutan yang menarik untuk dieksplorasi.

REFERENSI

- Dobesova, Z. (2024). *Orange data mining tool in geoinformatics education*. *Computer Applications in Engineering Education*, 32, e22735. <https://doi.org/10.1002/cae.22735>
- Nurqotimah, D. R., Khudori, A. N., & Pradini, R. S. (2024). Implementasi algoritma *Support Vector Machine* (SVM) untuk klasifikasi penyakit stroke. *Journal of Applied Computer Science and Technology (JACOST)*, 5(2), 179–185. <https://doi.org/10.52158/jacost.v5i2.817>
- Pranadjaya, E., Pangestu, E. S., Sereati, C. O., Octaviani, S., & Darmawan, M. (2024). Perbandingan algoritma *machine learning* menggunakan *Orange Data Mining* untuk klasifikasi jenis kendaraan pada sistem tilang digital. *Jurnal Elektro*, 17(1), 41–47. <https://doi.org/10.25170/jurnalelektro.v17i1.5429>
- Syahri, R., & Puspita, D. (2025). *Classification of outstanding students using Support Vector Machine (SVM) based on data mining*. *JITE (Journal of Informatics and Telecommunication Engineering)*, 9(1), 13–23. <https://doi.org/10.31289/jite.v9i1.13191>
- Zulherry, A., Ramadhani, F., & Satria, A. (2024). Klasifikasi data *Tracer Study* dengan pemanfaatan data mining menggunakan algoritma *Support Vector Machine* dan *Neural Network*. *PRINSIP: Portal Riset dan Inovasi Sistem Perangkat Lunak*, 2(1), 45–54. <https://doi.org/10.59696/prinsip.v2i1.73>