

Klasifikasi Strategi Penjualan Produk UMKM dengan Penerapan fitur seleksi Forward Selection pada Algoritma C4.5

M. Rudi Fanani^{*1}, Elvinda Bendra Agustina²

^{*1}Program Studi Teknologi Informasi, Fakultas Sains dan Teknologi,
Institut Teknologi dan Sains Nahdlatul Ulama Pekalongan

²Program Studi Fisika, Fakultas Sains dan Teknologi,

Institut Teknologi dan Sains Nahdlatul Ulama Pekalongan

Email: *idurinanaf@gmail.com, elvindabendra89@gmail.com

(Naskah masuk: 27 Juli 2024, diterima untuk diterbitkan: 13 Oktober 2025)

Abstrak: UMKM atau Usaha mikro, kecil dan menengah telah menjadi sebuah hal penting untuk menciptakan perubahan perekonomian, pengurangan kemiskinan, dan peningkatan kesejahteraan masyarakat di Indonesia. Selain memainkan peran dengan baik dalam hal menghasilkan lapangan kerja dan mempromosikan ekonomi lokal, UMKM juga merupakan pusat inspirasi untuk kreativitas dan inovasi. Kelurahan Kedungwuni Timur terletak di distrik Kedungwuni di Kabupaten Pekalongan, Jawa Tengah. Mayoritas warga Kelurahan Kedungwuni Timur bekerja di UMKM Fashion. Konsumen sekarang dapat dengan mudah membandingkan produk. Oleh karena itu, sebuah strategi penjualan yang menggunakan konsep teknologi informasi harus dibuat. Salah satu metode dalam ilmu teknologi informasi adalah data mining. Ilmu tersebut berfungsi sebagai penambangan data guna menghasilkan sebuah keputusan. Beberapa algoritma yang tersedia dalam data mining adalah algoritma C4.5, algoritma tersebut dapat difungsikan sebagai prediksi dalam menentukan strategi penjualan produk UMKM melalui pohon keputusan. Penelitian ini menghasilkan nilai akurasi yang didapat sebesar 82,78%. Hasil tersebut masih bisa lebih baik dengan menambahkan fitur forward selection, setelah menambahkan fitur tersebut nilai akurasi bertambah menjadi 86,11%.

Kata Kunci – Algoritma C4.5; Forward Selection; Data Mining; Penjualan Produk

Classification of MSME Product Sales Strategies by Applying the Forward Selection Feature in the C4.5 Algorithm

Abstract: MSMEs or Micro, Small and Medium Enterprises (MSMEs) have been instrumental in creating economic change, poverty reduction, and improved community welfare in Indonesia. Besides playing a good role in terms of generating employment and promoting the local economy, MSMEs are also a center of inspiration for creativity and innovation. Kelurahan Kedungwuni Timur is located in the Kedungwuni district of Pekalongan Regency, Central Java. The majority of residents of Kedungwuni Timur Village work in Fashion MSMEs. Consumers can now easily compare products. Therefore, a sales strategy that uses the concept of information technology must be created. One of the methods in information technology is data mining. This science functions as data mining to produce a decision. Some of the algorithms available in data mining are the C4.5 algorithm, the algorithm can function as a prediction in determining the sales strategy of MSME products through a decision tree. This research resulted in an accuracy value obtained of 82.78%. These results can still be better by adding the forward selection feature, after adding this feature the accuracy value increases to 86.11%.

Keywords – C4.5 Algorithm; Forward Selection; Data Mining; Product Sales

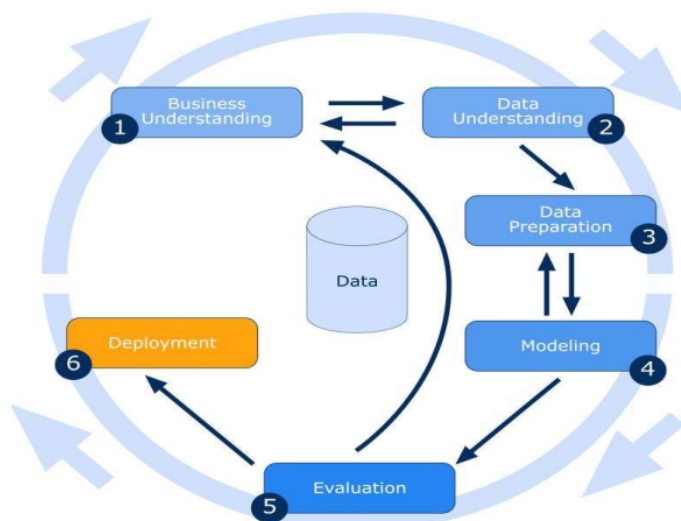
1. PENDAHULUAN

UMKM atau Usaha mikro, kecil dan menengah menjadi sangat berguna dalam mengurangi kemiskinan, memajukan ekonomi, dan meningkatkan kesejahteraan masyarakat. UMKM memiliki peran strategis untuk menghasilkan lapangan kerja, mendorong ekonomi warga lokal, dan sebagai pusat inspirasi untuk kreativitas dan inovasi. Selama bertahun-tahun, UMKM telah menjadi komponen usaha yang penting dalam ekonomi Indonesia [1]. Saat ini, jumlah UMKM dalam setiap tahunnya bertambah, dan berita bagus ini akan menjadikan suatu hal yang baik bagi perekonomian

dalam Indonesia. Kementerian Koperasi dan UMKM menghimpun data yang dimana para pelaku UMKM membuat efek yang baik di angka 60,5% untuk PDB Nasional. Dalam hal ini memberikan gambaran nyata bahwa para pelaku UMKM dalam negeri bisa menjadi lebih baik karena potensi maupun efek baik yang timbul terhadap roda perekonomian [2]. Kabupaten Pekalongan berada di provinsi Jawa Tengah. Terletak di pesisir utara laut Jawa, terdapat kecamatan dengan jumlah 19 desa dengan jumlah 270 dan kelurahan dengan jumlah 13 [3]. Dengan 3.795 UMKM di Kabupaten Pekalongan, sektor fashion masih menjadi yang paling banyak, dengan mayoritas penduduknya bekerja di sektor fashion di Kelurahan Kedungwuni Timur [4]. Hasil wawancara dengan pengusaha UMKM menunjukkan bahwa konsumen saat ini menghadapi masalah dengan mudah membandingkan kualitas, harga, layanan, dan kecepatan bisnis, yang menyebabkan penurunan permintaan atau penjualan. Oleh karena itu, untuk meningkatkan minat pasar, daya saing, dan hubungan baik dengan konsumen, perlu dilakukan inovasi atau pendekatan baru dalam penggunaan konsep teknologi informasi. Data mining merupakan upaya yang digunakan untuk mengekstraksi data yang bisa memberikan informasi juga memberikan pengetahuan baru yang sebelumnya belum diketahui [5]. C4.5 adalah salah satu dari banyak algoritma data mining [6]. Dalam memutuskan penjualan produk UMKM diperlukan sebuah strategi, algoritma C4.5 digunakan untuk mengklasifikasikan data statistik [7]. Selain itu, fitur seleksi maju ditambahkan untuk meningkatkan akurasi algoritma C4.5 [8].

2. METODE PENELITIAN

Dalam Penelitian ini model Cross-Industry Standard Process for Data Mining (CRISP-DM) akan digunakan, seluruh enam siklus dalam data mining akan termuat dalam penelitian ini. Fase-fase tersebut meliputi pemahaman tentang bisnis, pemahaman tentang data, persiapan tentang data, pembuatan model, evaluasi, dan tahap terakhir penerapan [9]:



Gambar 1. Metode CRISP-DM [10]

1. Business Understanding (Pemahaman Bisnis)

Fokus bagian pemahan tentang bisnis adalah untuk mendapatkan pemahaman menyeluruh tentang tujuan bisnis yang ingin dicapai melalui proses penambangan data. Salah satu cara untuk mendapatkan pemahaman tentang bisnis adalah dengan membuat pemetaan konsep berdasarkan data, informasi, dan proses bisnis yang diperoleh dari wawancara dan konfirmasi dari pihak-pihak yang terkait dengan perusahaan yang diteliti.

2. Data Understanding (Pemahaman Data)

Pada tahap pemahaman data, analisis dilakukan dengan cara eksplorasi awal terhadap data yang tersedia untuk proses penambangan data. Informasi ini diperoleh dengan menggunakan data dari hasil pemetaan konsep dan dengan mewawancarai subjek penelitian untuk memastikan bahwa data dan informasinya tepat.

3. Data Preparation (Persiapan Data)

Data akan dimasukkan ke dalam bentuk digital, kemudian dibersihkan dan diproses untuk menjadi siap digunakan dalam proses pemodelan dengan algoritma C4.5. Hasil dari tahap persiapan data ini dapat dipresentasikan pada lembar digital (*spreadsheet*) menggunakan Microsoft Excel dalam penelitian ini.

4. Modeling (Pemodelan)

Model deskriptif dan prediktif dibuat menggunakan teknik penambangan data yang sesuai berdasarkan data yang telah disiapkan sebelumnya. Tahap pemodelan ini merupakan bagian penting dari proses penambangan data. Dalam proses penambangan data ini, wawasan dan pengetahuan penting akan diperoleh, dimana dalam hal ini bisa digunakan dalam pemutusan sesuatu hal lebih tepat dan strategis. Penelitian ini menggunakan Rapidminer sebagai alat bantu penambangan data dan visualisasi. Ini menggunakan algoritma C4.5 dengan fitur seleksi *Forward Selection*.

5. Evaluation (Evaluasi)

Untuk mengukur kualitas dan relevansinya terhadap hal sebelumnya, model dan hasil penambangan data dievaluasi secara menyeluruh. Tahap evaluasi memastikan bahwa hasil penambangan data bermanfaat, dapat diandalkan, dan sesuai dengan kebutuhan bisnis. Ini memastikan kesuksesan implementasi dan hasil dari informasi didapatkan yang lebih baik bisa digunakan dalam penentuan pengambilan keputusan.

6. Deployment (Penerapan)

Hasil model dan analisis penambangan data diterapkan ke lingkungan bisnis nyata. Analisis bekerja sama dengan tim bisnis untuk memastikan bahwa model dan wawasan diterapkan dengan baik dan terintegrasi dengan proses operasional.

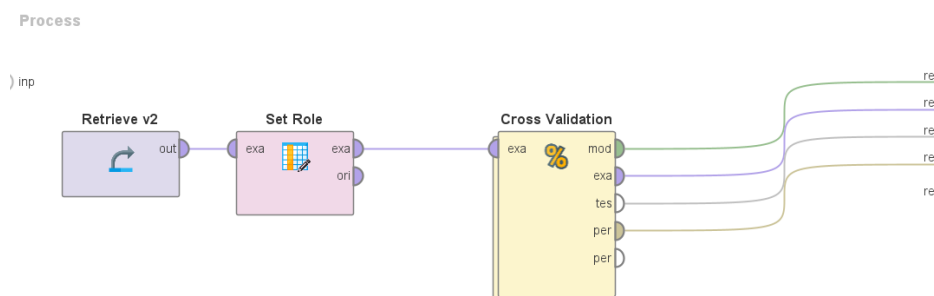
3. HASIL DAN PEMBAHASAN

3.1. Pembahasan

Data-data relevan yang disajikan merupakan hasil yang akan diberikan dalam penelitian ini yaitu berupa sebuah pohon Keputusan dan faktor-faktor yang berpengaruh untuk meningkatkan minat dan daya saing pasar serta hubungan dengan konsumen, data yang digunakan adalah data privat hasil dari wawancara terhadap para pelaku umkm di Kelurahan Kedungwuni Timur Kecamatan Kedungwuni Kabupaten Pekalongan Jawa Tengah.

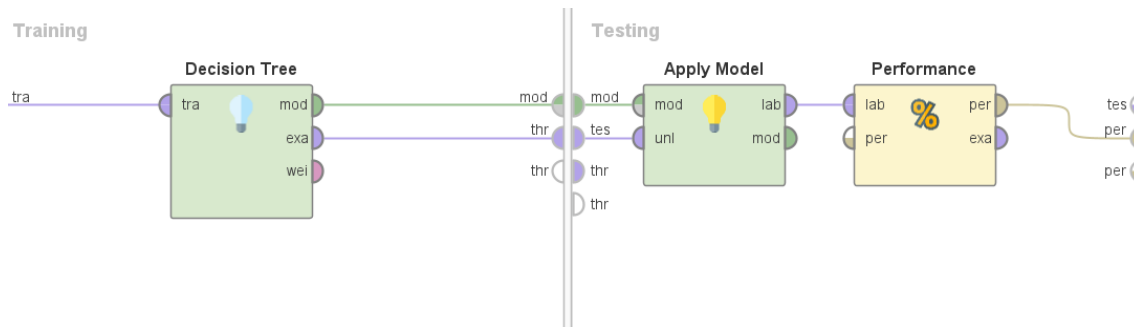
3.2. Pengujian dengan tools Rapidminer

Penelitian ini menggunakan *tools rapidminer* untuk menguji nilai akurasi, presisi dan recall serta hasil pohon Keputusan dari algoritma C4.5 saja dan penambahan fitur seleksi *forward selection* dalam algoritma C4.5.



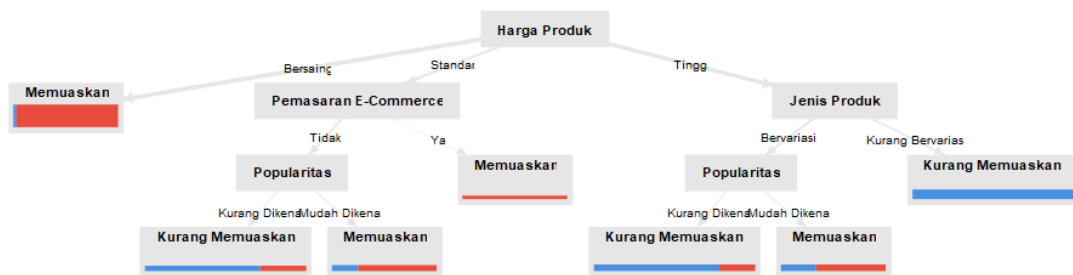
Gambar 2. Model Pengujian Algoritma C4.5

Pengujian ini menggunakan *Croos Validation* yang digunakan untuk membagi data secara acak (data *testing* dan data *training*) menggunakan 10 fold *Validation* dan *sampling type automatic*.



Gambar 3. Hasil Pengujian Menggunakan Algoritma C4.5 Saja

X-Validation menghasilkan sebuah proses pemodelan menggunakan algoritma C4.5 yang akan menghasilkan sebuah nilai akurasi dan pohon Keputusan.



Gambar 4. Hasil Pohon Keputusan

Hasil dari pohon Keputusan sebagai berikut:

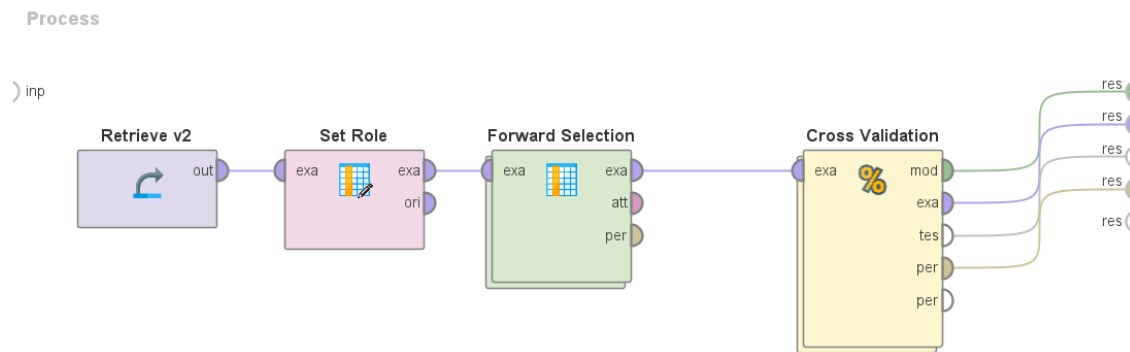
1. Jika Harga Produk Bersaing maka hasilnya adalah memuaskan
2. Jika harga produk standar maka dilihat pemasaran e-commerce jika ya maka hasilnya memuaskan
3. Jika harga produk standar maka dilihat pemasaran e-commerce jika tidak dilihat popularitas jika mudah dikenal maka hasilnya memuaskan jika kurang dikenal hasilnya kurang memuaskan
4. Jika harga produk tinggi maka dilihat jenis produknya jika kurang bervariasi maka hasilnya memuaskan
5. Jika harga produk tinggi maka dilihat jenis produknya jika bervariasi maka dilihat popularitasnya jika mudah dikenal hasilnya memuaskan jika kurang dikenal maka hasilnya kurang memuaskan.

accuracy: 82.78% +/- 13.41% (micro average: 82.56%)

	true Kurang Memuaskan	true Memuaskan	class precision
pred. Kurang Memuaskan	26	9	74.29%
pred. Memuaskan	6	45	88.24%
class recall	81.25%	83.33%	

Gambar 5. Hasil nilai Akurasi

Hasil akurasi menggunakan algoritma C4.5 adalah 82.78%.



Gambar 6. Proses Forward Selection dan Algoritma C4.5

Proses selanjutnya akan ditambahkan sebuah seleksi fitur *forward selection* guna menghasilkan peningkatan nilai akurasi dalam algoritma C4.5 yaitu akan menyeleksi fitur-fitur yang berpengaruh terhadap dataset.

accuracy: 86.11% +/- 10.46% (micro average: 86.05%)

	true Kurang Memuaskan	true Memuaskan	class precision
pred. Kurang Memuaskan	27	7	79.41%
pred. Memuaskan	5	47	90.38%
class recall	84.38%	87.04%	

Gambar 7. Hasil Akurasi dari penambahan fitur seleksi *forward selection* pada algoritma C4.5

Berdasarkan pengujian menggunakan tools *rapidminer* penambahan fitur seleksi *forward selection* bisa meningkatkan akurasi menjadi 86.11%.

Row No.	Tingkat Penj...	Harga Produk	Promosi Me...	Popularitas	Jenis Produk
1	Kurang Mem...	Bersaing	Ya	Mudah Dikenal	Kurang Berv...
2	Memuaskan	Standar	Ya	Mudah Dikenal	Kurang Berv...
3	Kurang Mem...	Tinggi	Ya	Mudah Dikenal	Bervariasi
4	Memuaskan	Bersaing	Tidak	Mudah Dikenal	Bervariasi
5	Memuaskan	Bersaing	Tidak	Mudah Dikenal	Bervariasi
6	Kurang Mem...	Tinggi	Tidak	Kurang Dikenal	Kurang Berv...
7	Memuaskan	Standar	Ya	Kurang Dikenal	Bervariasi
8	Memuaskan	Bersaing	Ya	Mudah Dikenal	Bervariasi
9	Memuaskan	Bersaing	Ya	Mudah Dikenal	Kurang Berv...
10	Kurang Mem...	Tinggi	Ya	Kurang Dikenal	Kurang Berv...

Gambar 8. Fitur seleksi yang berpengaruh terhadap dataset

Berikut hasil perbandingan akurasi algoritma c4.5 saja serta adanya fitur seleksi *forward selection* yang ditambahkan dalam algoritma C4.5.

Tabel 1. Perbandingan Nilai Akurasi

Algoritma	Nilai Akurasi
C4.5	82.78%
Forward Selection dan C4.5	86.11%

4. KESIMPULAN

Berdasarkan hasil penelitian telah dilakukan pengumpulan data terhadap para pelaku UMKM di Kelurahan Kedungwuni Timur Kabupaten Pekalongan Jawa Tengah, selanjutnya dilakukan penggalian dan pemrosesan data menggunakan Teknik data mining, algoritma yang dipakai dalam penelitian ini adalah algoritma C4.5 menghasilkan nilai akurasi sebesar 82.78% serta penambahan fitur seleksi *forward selection* dalam Algoritma C4.5 memberikan nilai akurasi dengan jumlah 86.11%. Penambahan fitur seleksi *forward selection* terbukti bisa meningkatkan akurasi daripada menggunakan algoritma C4.5 saja.

UCAPAN TERIMA KASIH

Terima kasih kepada pihak Pemerintah Kelurahan Kedungwuni Timur Kabupaten Pekalongan, Para pelaku UMKM, tim peneliti dan pembantu peneliti serta LPPM ITSNU Pekalongan untuk memastikan kegiatan penelitian ini dapat selesai dengan baik dan lancar, semoga bisa memberikan manfaat untuk semua.

DAFTAR PUSTAKA

- [1] M. Farras Nasrida, A. Pandahang, and D. Febrian, "Perkembangan UMKM Di Indonesia Dan Potensi Di Kota Palangka Raya," *Jurnal Manajemen Bisnis Kewirausahaan*, vol. 2, no. 1, 2023.
- [2] Kementrian Keuangan RI, "Kontribusi UMKM dalam Perekonomian Indonesia," <https://djpb.kemenkeu.go.id/>.
- [3] DPRD Kabupaten Pekalongan, "Selayang Pandang Kabupaten Pekalongan," <https://dprd-pekalongankab.go.id/halaman/selayang-pandang>.
- [4] Dinas Perindustrian dan Perdagangan (Dinperindag), "Jumlah UMKM Menurut Bidang di Kabupaten Pekalongan," https://data.pekalongankab.go.id/zh_TW/dataset/jumlah-umkm-menurut-bidang-di-kabupaten-pekalongan#.
- [5] Pratomo, "Datamining untuk Menganalisis Pengaruh Nilai Pemrograman Dasar Kelas X terhadap Hasil Belajar pada Mata Pelajaran Produktif Rekayasa Perangkat Lunak Kelas XI," *Smart Comp*, vol. 11, no. 2, 2022.
- [6] M. R. Fanani and K. Fari, "Seleksi Fitur PSO untuk Klasifikasi Jenis Kekerasan dengan Algoritma C4.5," *Smart Comp*, 2022.
- [7] T. Widodo, D. Setiawan, A. Syahputri, S. Informasi, and S. Triguna Dharma, "Data Mining Menentukan Minat Konsumen Memilih Sepeda Motor Idaman Dengan Algoritma C4.5," *Jurnal Sistem Informasi TGD*, 2022, Available: <https://ojs.trigunadharma.ac.id/index.php/jsi>
- [8] P. Harapan and B. Tegal, "PEMILIHAN FITUR DENGAN FORWARD SELECTION PADA METODE K-NEAREST NEIGHBOR UNTUK PREDIKSI BIMBINGAN KONSELING SISWA Wildani Eko Nugroho," 2023.
- [9] N. Lundén, E. T. Bekar, A. Skoogh, and J. Bokrantz, "Domain Knowledge in CRISP-DM: An Application Case in Manufacturing," in *IFAC-PapersOnLine*, Elsevier B.V., Jul. 2023, pp. 7603–7608. doi: 10.1016/j.ifacol.2023.10.1156.
- [10] S. Huber, H. Wiemer, D. Schneider, and S. Ihlenfeldt, "DMME: Data mining methodology for engineering applications - A holistic extension to the CRISP-DM model," in *Procedia CIRP*, Elsevier B.V., 2019, pp. 403–408. doi: 10.1016/j.procir.2019.02.106.