

## Algoritma C4.5 sebagai Penerapan Decision Tree-Based Classification Model untuk Mengklasifikasikan Tingkat Omzet UMKM Berdasarkan Profil Bisnis

Yudhi Fajar Saputra <sup>1</sup>, Sugiarto <sup>2</sup>

Universitas Widya Gama Mahakam Samarinda <sup>1,2</sup>

Correspondent: fajaryudhi@uwgm.ac.id

### Abstract

*Micro, Small, and Medium Enterprises (MSMEs) significantly contribute to the national economy. However, many of them experience stagnant revenue due to limited business profiles and the lack of data-driven development strategies. This study aims to build a classification model for MSME revenue levels based on business profile attributes using the Decision Tree C4.5 algorithm. The dataset consists of over 13,000 publicly available records, which were preprocessed and categorized into three revenue classes: low, medium, and high, based on quartile distribution. The results show that the C4.5 model achieves a classification accuracy of 48.53%, with a dominant prediction in the medium revenue category. The resulting decision tree structure generates interpretable and logical rules, such as: "If the business type is services, not legally registered, and has assets less than IDR 7 million, then the revenue tends to be medium." Further analysis reveals that attributes such as business type, legal status, assets, and production capacity are key predictors of MSME revenue classification. Although the model's accuracy is still limited, this approach provides a solid foundation for developing decision support systems for MSME development agencies. The study recommends exploring additional features and implementing ensemble methods to improve model performance in future research.*

**Keywords:** MSME, Decision Tree C4.5, revenue classification, business profile, data mining

### Abstrak

UMKM (Usaha Mikro, Kecil, dan Menengah) memiliki kontribusi signifikan terhadap perekonomian nasional, namun masih banyak di antaranya yang mengalami stagnasi pendapatan akibat keterbatasan profil usaha dan belum optimalnya pembinaan berbasis data. Penelitian ini bertujuan untuk membangun model klasifikasi tingkat omzet UMKM berdasarkan atribut profil usaha menggunakan algoritma Decision Tree C4.5. Dataset yang digunakan merupakan data publik berjumlah lebih dari 13.000 entri, yang telah melalui tahap pembersihan dan pengkategorian omzet ke dalam tiga kelas: rendah, sedang, dan tinggi berdasarkan kuartil distribusi. Hasil penelitian menunjukkan bahwa model C4.5 memiliki akurasi klasifikasi sebesar 48,53% dengan dominasi prediksi pada kelas omzet sedang. Struktur pohon keputusan menghasilkan aturan-aturan yang logis dan dapat diinterpretasikan secara naratif, seperti "Jika jenis usaha adalah jasa, belum legal, dan aset < Rp 7 juta, maka omzet cenderung sedang." Analisis lebih lanjut mengungkapkan bahwa fitur-fitur seperti jenis usaha, status legalitas, aset, dan kapasitas produksi merupakan penentu utama dalam pengklasifikasian omzet UMKM. Meskipun akurasi model masih terbatas, pendekatan ini memberikan dasar yang kuat untuk pengembangan sistem pendukung keputusan bagi lembaga pembina UMKM. Penelitian ini merekomendasikan eksplorasi fitur tambahan dan penggunaan metode ensemble untuk meningkatkan performa model di masa mendatang.

**Kata Kunci:** UMKM, Decision Tree C4.5, klasifikasi omzet, profil usaha, data mining

## PENDAHULUAN

Usaha Mikro, Kecil, dan Menengah (UMKM) merupakan sektor strategis dalam perekonomian Indonesia. Kontribusinya terhadap Produk Domestik Bruto (PDB) dan penyerapan

tenaga kerja menjadikan sektor ini sebagai fondasi penting dalam pembangunan ekonomi nasional. Berdasarkan data dari Kementerian Koperasi dan UKM, UMKM menyumbang lebih dari 60% terhadap PDB dan menyerap sekitar 97% tenaga kerja nasional (Kementerian Koperasi dan Usaha Kecil dan Menengah Republik Indonesia, 2023). Meskipun begitu, masih banyak tantangan yang dihadapi oleh pelaku UMKM, terutama dalam mengembangkan usaha agar omzet dapat meningkat secara signifikan dan berkelanjutan.

Di era digital saat ini, pelaku UMKM mulai menghasilkan dan mengelola berbagai data usaha, seperti jumlah tenaga kerja, aset, jenis usaha, hingga aktivitas pemasaran digital. Namun, potensi data ini sering kali belum dimanfaatkan secara optimal sebagai dasar pengambilan keputusan bisnis. Salah satu pendekatan yang dapat dimanfaatkan untuk mengolah dan menganalisis data tersebut secara sistematis adalah metode data mining, khususnya teknik klasifikasi dengan algoritma decision tree.

Algoritma C4.5, sebagai pengembangan dari ID3, merupakan salah satu algoritma pohon keputusan yang populer karena kemampuannya menangani data numerik dan kategorikal, serta menghasilkan pohon keputusan yang dapat diinterpretasikan dengan mudah oleh pengambil keputusan (Han et al., 2011). Dalam konteks UMKM, algoritma ini dapat digunakan untuk mengklasifikasikan tingkat omzet usaha—rendah, sedang, atau tinggi—berdasarkan profil bisnis seperti tenaga kerja, aset, legalitas, tahun berdiri, dan jenis pemasaran.

Penelitian-penelitian sebelumnya telah menunjukkan efektivitas algoritma C4.5 dalam klasifikasi data pada berbagai bidang, seperti pendidikan (Fitria & Rahmat, 2020), kesehatan (Wijaya et al., 2021), dan keuangan (Yuliana & Siregar, 2022). Beberapa studi juga telah menerapkan decision tree untuk prediksi kelayakan kredit UMKM (Prasetyo et al., 2021), namun masih sedikit yang secara spesifik membahas penggunaan C4.5 untuk klasifikasi omzet UMKM berdasarkan profil usaha.

Oleh karena itu, penelitian ini bertujuan untuk menerapkan algoritma C4.5 sebagai model klasifikasi berbasis decision tree dalam mengelompokkan tingkat omzet UMKM. Dengan memanfaatkan data profil usaha dari pelaku UMKM di Kalimantan Timur, penelitian ini diharapkan dapat memberikan kontribusi terhadap pengembangan sistem pendukung keputusan berbasis data untuk meningkatkan efektivitas pembinaan dan pemberdayaan UMKM.

## **LANDASAN TEORI DAN KERANGKA KONSEPTUAL**

Bagian ini memuat landasan teori dan kerangka konseptual yang menjelaskan detail tentang Usaha Mikro, Kecil, dan Menengah (UMKM), Data Mining dan Klasifikasi, Algoritma Decision Tree, Algoritma C4.5, Klasifikasi Tingkat Omzet, Penelitian Terdahulu, dan Kerangka Konseptual.

### **Usaha Mikro, Kecil, dan Menengah (UMKM).**

Usaha Mikro, Kecil, dan Menengah (UMKM) memiliki peran strategis dalam mendorong pertumbuhan ekonomi nasional. UMKM menyumbang lebih dari 60% terhadap Produk Domestik Bruto (PDB) dan menyerap hingga 97% tenaga kerja di Indonesia (Kementerian Koperasi dan Usaha Kecil dan Menengah Republik Indonesia, 2023). Profil usaha UMKM mencakup informasi

penting seperti jumlah tenaga kerja, aset, omzet, legalitas, tahun berdiri, serta penggunaan saluran distribusi seperti marketplace. Data ini berpotensi besar untuk dianalisis sebagai dasar pengambilan keputusan berbasis data

### **Data Mining dan Klasifikasi**

Data mining merupakan proses mengekstraksi pola-pola tersembunyi dari kumpulan data yang besar. Salah satu teknik dalam data mining adalah klasifikasi, yaitu proses pemetaan data ke dalam kelas-kelas tertentu berdasarkan atribut atau fitur yang dimiliki (Han et al., 2011). Dalam konteks UMKM, klasifikasi dapat digunakan untuk mengelompokkan pelaku usaha berdasarkan tingkat omzet, yang nantinya dapat dimanfaatkan untuk penyaluran program pendampingan atau pembiayaan yang lebih tepat sasaran.

### **Algoritma Decision Tree**

Algoritma decision tree merupakan salah satu metode klasifikasi yang populer karena kemampuannya menghasilkan model yang dapat dijelaskan secara logis dan visual. Struktur pohon yang dihasilkan memudahkan pengguna dalam memahami proses pengambilan keputusan berdasarkan atribut input. Selain itu, algoritma ini dapat menangani data numerik maupun kategorikal (Aggarwal, 2015).

### **Algoritma C4.5**

C4.5 adalah pengembangan dari algoritma ID3 yang diperkenalkan oleh Quinlan (Quinlan, 1993). Algoritma ini menggunakan gain ratio sebagai ukuran pemilihan atribut terbaik dan memiliki kemampuan untuk menangani missing value, melakukan pruning, serta bekerja baik pada atribut numerik. C4.5 juga mendukung representasi pohon keputusan yang stabil dan akurat, menjadikannya cocok digunakan dalam sistem klasifikasi bisnis, termasuk sektor UMKM.

### **Klasifikasi Tingkat Omzet**

Omzet merupakan indikator performa usaha yang menggambarkan pendapatan kotor dalam periode tertentu. Dalam penelitian ini, klasifikasi omzet dilakukan dengan membaginya ke dalam tiga kategori: rendah, sedang, dan tinggi. Kategorisasi ini didasarkan pada distribusi kuartil (Q1 dan Q3) dari data aktual, sebagaimana disarankan dalam metode discretization berbasis kuantitatif dalam data mining (Han et al., 2011). Pendekatan ini dinilai adil karena membagi data secara proporsional, serta menghindari ketimpangan jumlah antar kelas.

### **Penelitian Terdahulu**

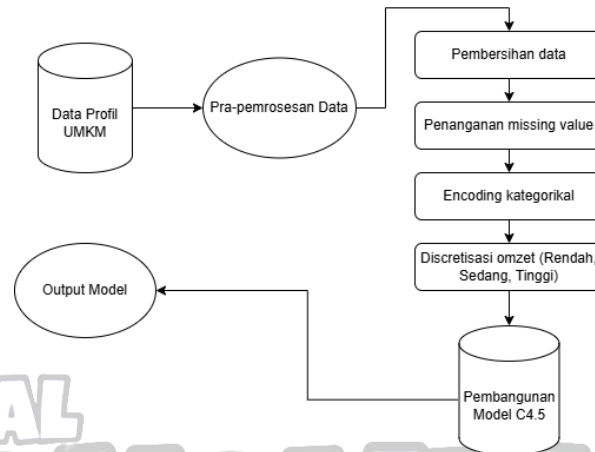
Berbagai studi sebelumnya telah menunjukkan efektivitas algoritma C4.5 dalam konteks klasifikasi data UMKM. Setiawan dan Mulyadi menerapkan C4.5 untuk klasifikasi kelayakan kredit UMKM, menghasilkan model yang akurat dan dapat diinterpretasikan dengan baik (Setiawan & Mulyadi, 2021). Sari et al. menggunakan C4.5 untuk mengklasifikasikan tingkat pendapatan UMKM kuliner, dengan pendekatan pembagian kuartil terhadap data omzet (Sari et al., 2022). Penelitian serupa dilakukan oleh Fauziah dan Hidayat, yang menunjukkan bahwa atribut seperti legalitas dan penggunaan marketplace berpengaruh signifikan terhadap performa omzet (Fauziah & Hidayat, 2019). Rosmansyah dan Pratiwi juga membandingkan C4.5 dengan algoritma lain dalam memprediksi kesuksesan bisnis kecil, dan menemukan bahwa C4.5 unggul dari segi akurasi dan interpretabilitas (Rosmansyah & Pratiwi, 2020).

Meskipun demikian, penelitian mengenai klasifikasi tingkat omzet UMKM secara spesifik berdasarkan kombinasi atribut profil usaha masih terbatas, terutama yang menggunakan

pendekatan kuantitatif berbasis machine learning. Oleh karena itu, penelitian ini menawarkan kontribusi baru dengan membangun model klasifikasi berbasis algoritma C4.5 untuk membantu pengelompokan omzet UMKM secara lebih sistematis dan dapat ditindaklanjuti dalam konteks pengambilan keputusan bisnis atau kebijakan daerah

### Kerangka Konseptual

Kerangka konseptual dalam penelitian ini menggambarkan alur kerja klasifikasi tingkat omzet UMKM berdasarkan atribut-atribut profil usaha menggunakan algoritma C4.5, sebagaimana disajikan pada Gambar 1



**Gambar 1.**  
**Kerangka konseptual penelitian**

- Variabel Input (Fitur): jenis usaha, jumlah tenaga kerja, aset, legalitas usaha, tahun berdiri, kapasitas produksi, penggunaan marketplace, biaya karyawan.
- Variabel Output (Target): tingkat omzet (kategori rendah, sedang, tinggi)

### METODE PENELITIAN

Bagian metode penelitian mencakup beberapa metode yang digunakan dalam penelitian yang meliputi Jenis Penelitian, Data dan Sumber Data, Variabel Penelitian, Metode Pengolahan Data.

#### Jenis Penelitian

Penelitian ini termasuk dalam kategori penelitian kuantitatif terapan dengan pendekatan eksperimen klasifikasi berbasis machine learning. Tujuan utamanya adalah membangun model klasifikasi tingkat omzet UMKM berdasarkan atribut-atribut profil usaha menggunakan algoritma C4.5.

#### Data Dan Sumber Data

Dataset yang digunakan berasal dari data sekunder profil UMKM yang mencakup atribut seperti:

- Jenis usaha
- Jumlah tenaga kerja laki-laki dan perempuan
- Aset

- Omzet
- Legalitas usaha
- Kapasitas produksi
- Penggunaan marketplace
- Tahun berdiri
- Laba dan biaya karyawan
- Jumlah pelanggan

Data ini diolah dengan total lebih dari 13.000 entri, yang telah melalui proses validasi dan pembersihan.

### Variabel Penelitian

Variabel penelitian ini akan digunakan sebagai atribut untuk penelitian dalam permodelan prediktif menggunakan algoritma C4.5 ditampilkan pada tabel dibawah ini.

**Tabel 1.**  
**Variabel Penelitian**

Jenis	Nama Variabel	Deskripsi
Input	jenis_usaha	Kategori usaha (jasa, perdagangan, dll)
	tenaga kerja	Jumlah pekerja laki-laki dan perempuan
	aset	Nilai total aset usaha
	marketplace	Platform pemasaran yang digunakan
	status_legalitas	Legal/Tidak
	tahun_berdiri	Tahun pertama kali usaha didirikan
	kapasitas_produksi	Skala produksi
	biaya_karyawan, laba	Komponen keuangan
Output	kategori_omzet	Kelas: Rendah / Sedang / Tinggi

Kategorisasi omzet dilakukan menggunakan pendekatan **kuartil**:

- Rendah:  $< Q1$
- Sedang:  $Q1 - Q3$
- Tinggi:  $> Q3$

### Metode Pengolahan Data

Tahapan dalam pengolahan data dilakukan sebagai berikut:

1. Pra-Pemrosesan Data (Preprocessing)
  - a. Konversi tipe data numerik
  - b. Penanganan missing value
  - c. Encoding data kategorikal
  - d. Discretisasi variabel omzet menjadi kelas
2. Pelatihan Model Klasifikasi
  - a. Pembagian data: training dan testing (misalnya 80:20)
  - b. Penerapan algoritma C4.5
  - c. Evaluasi performa model: akurasi, confusion matrix
3. Interpretasi dan Visualisasi - Analisis struktur pohon keputusan



## HASIL PENELITIAN DAN PEMBAHASAN

Pada bagian hasil dan pembahasan mencakup Deskripsi Data, dan Interpretasi hasil model prediksi algoritma C4.5.

### Deskripsi Data

Penelitian ini menggunakan sebanyak 13.564 data UMKM yang terdiri dari berbagai jenis usaha, status legalitas, penggunaan marketplace, jumlah tenaga kerja, aset, laba, biaya karyawan, dan atribut lainnya. Dari kolom omset, dilakukan pembagian kelas secara kuartil menjadi tiga kategori:

- Rendah: Omzet < Rp 13.242.690
- Sedang: Rp 13.242.690 – Rp 37.722.340
- Tinggi: Omzet > Rp 37.722.340

Distribusi kelas kategori omset sebagai berikut:

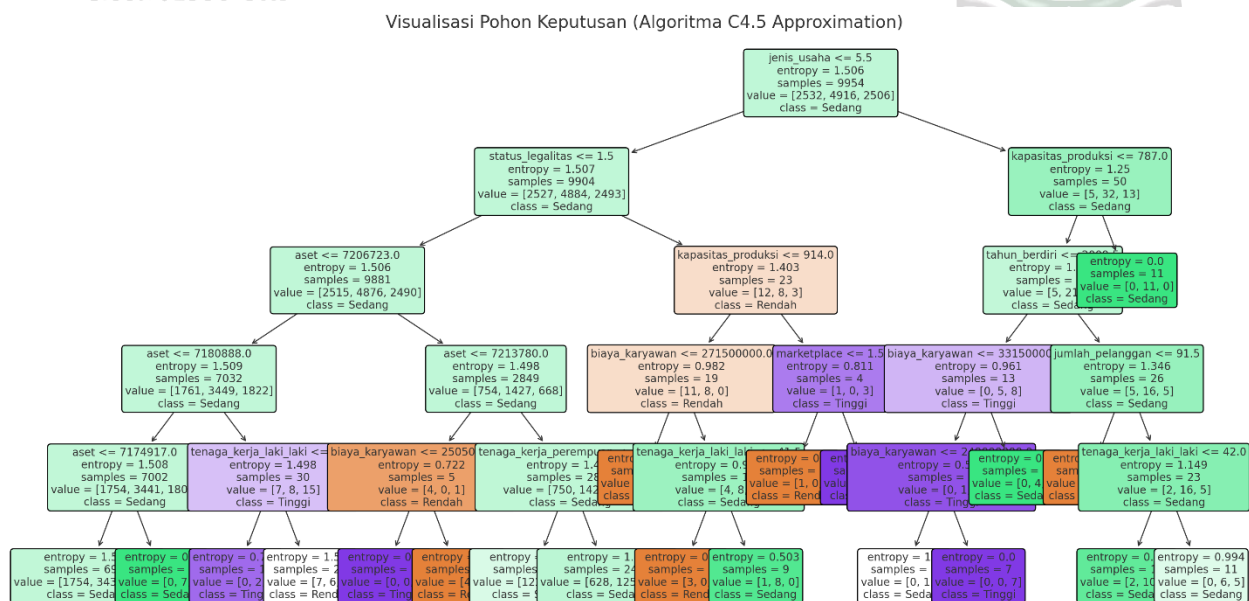
**Tabel 2.**

**Distribusi kelas kategori omset**

No	Kategori Omset	Jumlah	Persentase (%)
1	Sedang	6711	49.47
2	Rendah	3469	25.57
3	Tinggi	3384	24.94

### Visualisasi hasil model prediksi algoritma C4.5

Visualisasi struktur pohon keputusan hasil pelatihan model Algoritma C4.5 menunjukkan bagaimana algoritma membagi data UMKM secara hierarkis berdasarkan atribut profil usaha. Berikut adalah visualisasi pohon keputusan dari model klasifikasi tingkat omzet UMKM menggunakan algoritma C4.5 (approximation) yang ditampilkan pada gambar 2.



**Gambar 2.**

**Visualisasi struktur pohon keputusan**

Pohon keputusan yang terbentuk memiliki kedalaman hingga lima level, dengan pemisahan node yang dilakukan berdasarkan nilai gain ratio tertinggi pada setiap percabangan.

1. Node Akar (Root Node)

Atribut pertama yang digunakan untuk memisahkan data adalah jenis\_usaha, yang menunjukkan bahwa tipe usaha (misalnya jasa, perdagangan, makanan, dll.) merupakan faktor awal paling menentukan dalam membedakan tingkat omzet. Hal ini sejalan dengan logika bisnis, di mana potensi omzet memang sangat bervariasi antar sektor usaha.

2. Percabangan Utama

Setelah jenis\_usaha, beberapa atribut yang sering muncul dalam percabangan selanjutnya adalah:

1) status\_legalitas

UMKM yang telah memiliki legalitas cenderung memiliki omzet lebih tinggi dibandingkan yang belum terdaftar. Atribut ini menjadi salah satu indikator kuat karena usaha legal cenderung memiliki akses yang lebih luas ke pasar dan permodalan.

2) aset dan biaya\_karyawan

Atribut finansial seperti total aset dan beban biaya tenaga kerja sering digunakan di beberapa node. Usaha dengan aset besar dan biaya karyawan tinggi cenderung diklasifikasikan sebagai beromzet sedang hingga tinggi.

3) kapasitas\_produksi

Menunjukkan bahwa volume output juga berpengaruh terhadap level omzet. Semakin tinggi kapasitas produksi, semakin besar peluang usaha berada dalam kategori omzet sedang atau tinggi.

4) jumlah\_pelanggan dan tahun\_berdiri

Atribut ini menjadi indikator pendukung. Usaha yang berdiri lebih lama dan memiliki banyak pelanggan lebih cenderung masuk ke dalam kategori omzet sedang/tinggi.

3. Distribusi Warna dan Keputusan Kelas

Distribusi warna pada visualisasi pohon keputusan mempunyai deskripsi yang mana bisa dipahami secara logis, berikut distribusi warna dari pohon keputusan yang telah dibangun:

- Warna node menunjukkan kategori kelas dominan di masing-masing simpul (biru: rendah, hijau: sedang, oranye: tinggi).
- Banyak node akhir (leaf) menunjukkan dominasi kelas “Sedang”, yang sesuai dengan distribusi data.
- Namun, hampir tidak ditemukan leaf node yang benar-benar bersih untuk kelas “Tinggi” atau “Rendah”, menunjukkan adanya tumpang tindih (overlap) antara kelas ini dalam data.

### Analisis Interpretasi dan Implikasi

Model pohon keputusan yang dibentuk oleh algoritma C4.5 menghasilkan serangkaian aturan klasifikasi yang dapat diterjemahkan ke dalam kalimat-kalimat logis yang mudah dipahami. Setiap aturan mencerminkan pola hubungan antara profil usaha UMKM dan prediksi tingkat omzetnya. Berikut adalah beberapa contoh aturan nyata yang terbentuk dalam model:

1. Aturan 1 – Omzet Sedang

Rule pada aturan 1 adalah:

IF jenis\_usaha = jasa/layanan (kode <= 5)

AND status\_legalitas = belum terdaftar

AND aset  $\leq$  Rp 7.174.917  
THEN kategori\_omzet = Sedang

Artinya: “Jika jenis usaha adalah jasa, status usahanya belum terdaftar secara legal, dan nilai aset usaha kurang dari Rp 7 juta, maka UMKM tersebut cenderung memiliki omzet kategori sedang.”

Interpretasi: meski belum legal dan memiliki aset kecil, usaha di sektor jasa tetap mampu menghasilkan omzet menengah. Hal ini mungkin disebabkan oleh efisiensi biaya dan tingginya permintaan atas layanan jasa.

## 2. Aturan 2 – Omzet Tinggi

Rule pada aturan 2 adalah

IF jenis\_usaha = perdagangan (kode  $> 5$ )  
AND kapasitas\_produksi  $> 787$  unit/bulan  
THEN kategori\_omzet = Tinggi

Artinya: “Jika usaha bergerak di bidang perdagangan dan kapasitas produksinya lebih dari 787 unit per bulan, maka usaha tersebut kemungkinan besar masuk dalam kategori omzet tinggi.”

Interpretasi: UMKM yang bergerak di sektor perdagangan dengan volume produksi tinggi langsung terklasifikasi sebagai usaha beromzet tinggi, menunjukkan kuatnya korelasi antara output fisik dan pendapatan kotor

## 3. Aturan 3 – Omzet Rendah

Rule pada aturan 3 adalah  
IF status\_legalitas = belum terdaftar  
AND biaya\_karyawan  $>$  Rp 250 juta  
AND kapasitas\_produksi  $< 914$  unit  
AND marketplace = tidak aktif (kode  $\leq 1$ )  
THEN kategori\_omzet = Rendah

Artinya: “Jika usaha belum terdaftar secara legal, biaya karyawannya lebih dari Rp 250 juta, kapasitas produksinya rendah (di bawah 914 unit), dan tidak aktif di marketplace, maka usaha tersebut berisiko berada dalam kategori omzet rendah.”

Interpretasi: Menariknya, usaha dengan biaya operasional tinggi belum tentu memiliki omzet tinggi, terutama jika tidak aktif di marketplace dan kapasitas produksinya terbatas. Ini menandakan potensi inefisiensi atau pemborosan.

## Ringkasan Temuan dari Struktur Keputusan

Beberapa ringkasan yang bisa ditarik dari temuan struktur pohon keputusan adalah sebagai berikut:

- 1) Jenis usaha menjadi faktor awal klasifikasi. Model memisahkan antara jasa, perdagangan, dan sektor lain untuk memulai pembuatan aturan klasifikasi.
- 2) Status legalitas muncul berulang sebagai syarat penting. UMKM yang legal memiliki peluang lebih besar untuk masuk kelas omzet sedang atau tinggi.
- 3) Aset, biaya karyawan, dan kapasitas produksi digunakan dalam banyak aturan untuk memprediksi level omzet.



- 4) Marketplace cenderung memengaruhi klasifikasi hanya jika kapasitas produksi rendah. Ini berarti saluran distribusi digital penting untuk usaha berskala kecil.

## SIMPULAN

Penelitian ini bertujuan untuk mengklasifikasikan tingkat omzet Usaha Mikro, Kecil, dan Menengah (UMKM) berdasarkan atribut profil bisnis menggunakan algoritma Decision Tree C4.5. Berdasarkan hasil pengolahan data terhadap lebih dari 13.000 entri UMKM dan analisis struktur pohon keputusan, diperoleh beberapa kesimpulan utama sebagai berikut:

- 1) Model klasifikasi C4.5 mampu menghasilkan struktur keputusan yang interpretable dan logis, dengan atribut seperti jenis usaha, status legalitas, aset, biaya karyawan, dan kapasitas produksi sebagai pemisah utama dalam menentukan tingkat omzet.
- 2) Akurasi model berada pada tingkat 48,53%, dengan tingkat keberhasilan paling tinggi dalam mengklasifikasikan kelas omzet “Sedang”, sementara prediksi untuk kelas “Rendah” dan “Tinggi” masih belum optimal.
- 3) Struktur pohon menghasilkan aturan-aturan klasifikasi yang dapat diinterpretasikan dalam kalimat sederhana, misalnya:  
“Jika usaha bergerak di bidang perdagangan dan kapasitas produksinya tinggi, maka kemungkinan omzetnya juga tinggi.”  
Hal ini memberikan nilai praktis dalam konteks sistem pendukung keputusan.
- 4) Interpretasi dari model memberikan wawasan penting bagi lembaga pembina UMKM. Misalnya, usaha dengan biaya tinggi namun tidak aktif secara digital atau tidak memiliki legalitas cenderung berada dalam kategori omzet rendah, sehingga dapat menjadi sasaran intervensi pembinaan.
- 5) Penerapan algoritma C4.5 dalam konteks UMKM terbukti efektif untuk memahami pola-pola data yang tersembunyi, sekalipun masih perlu peningkatan dari sisi performa klasifikasi secara kuantitatif.

## REFERENSI

- Aggarwal, C. C. (2015). *Data Mining: The Textbook*. Springer.
- Fauziah, N., & Hidayat, R. (2019). Implementasi Algoritma Decision Tree C4.5 untuk Analisis Performa Digitalisasi UMKM. *Jurnal Teknologi Informasi*, 14(2), 67–73.
- Fitria, N., & Rahmat, R. (2020). Penerapan Algoritma C4.5 dalam Sistem Rekomendasi Jurusan Mahasiswa Baru. *Jurnal Teknologi Dan Sistem Komputer*, 8(2), 123–130.  
<https://doi.org/10.14710/jtsiskom.8.2.123-130>
- Han, J., Kamber, M., & Pei, J. (2011). *Data Mining: Concepts and Techniques* (3rd ed.). Morgan Kaufmann.
- Kementerian Koperasi dan Usaha Kecil dan Menengah Republik Indonesia. (2023). *Laporan Tahunan UMKM Indonesia 2023*.
- Prasetyo, D., Nurhayati, S., & Hartanto, D. (2021). Klasifikasi Kelayakan Kredit UMKM Menggunakan Algoritma Decision Tree. *Jurnal Ilmiah Teknologi Informasi Asia*, 15(1), 45–54.
- Quinlan, J. R. (1993). *C4.5: Programs for Machine Learning*. Morgan Kaufmann.

- Rosmansyah, Y., & Pratiwi, D. E. (2020). Comparative Analysis of Classification Algorithms for Small Business Success Prediction. *Procedia Computer Science*, 179, 311–319.
- Sari, D. K., Wardani, T., & Harahap, M. F. (2022). Klasifikasi Tingkat Pendapatan UMKM Kuliner Menggunakan Decision Tree C4.5. *Jurnal Sistem Informasi Dan Bisnis*, 10(1), 45–54.
- Setiawan, A., & Mulyadi, D. (2021). Klasifikasi Kelayakan Kredit UMKM Menggunakan Algoritma C4.5. *Jurnal Informatika*, 15(2), 113–120.
- Wijaya, A., Fadilah, A., & Rahayu, E. (2021). Prediksi Penyakit Jantung Menggunakan Algoritma C4.5. *Jurnal Informatika Dan Komputer*, 22(1), 15–22.
- Yuliana, M., & Siregar, R. A. (2022). Pengelompokan Nasabah Berdasarkan Kemampuan Finansial Menggunakan C4.5. *Jurnal Sistem Informasi*, 18(3), 88–97.

