

## PENDEKATAN DATA MINING UNTUK OPTIMASI PENERIMA ZAKAT PADA BAZNAS KOTA CIREBON MENGGUNAKAN ALGORITMA C4.5

Argyanti Ayu Putri Agung<sup>1</sup>, Bambang Irawan<sup>2</sup>, Willy Prihartono<sup>3</sup>, Aris Pratama Putra<sup>4</sup>.

Program Studi Teknik Informatika<sup>12</sup>  
Program Studi Komputerisasi Akuntansi<sup>3</sup>  
Program Studi Sistem Informasi<sup>4</sup>

STMIK IKMI Cirebon  
<https://ikmi.ac.id/page/18/?lang=de>  
ayuputriagung02@gmail.com

(\*) Corresponding Author : ayuputriagung02@gmail.com  
Published : 30 Januari 2026

**Abstract**—Zakat is one of the fundamental pillars of Islam and serves as an essential instrument for promoting social justice and reducing economic disparities. The accurate distribution of zakat funds is an essential factor in improving the effectiveness of BAZNAS Cirebon City's poverty alleviation programs. However, the current process of selecting zakat recipients is still conducted manually and subjectively, which may result in inaccuracies in determining eligible beneficiaries (mustahik). This study aims to optimize the zakat recipient selection process by applying a data mining approach using the C4.5 algorithm. The method classifies the eligibility of zakat recipients based on several socio-economic attributes such as income, dependents, occupation, education, and housing condition. The dataset was obtained from BAZNAS Cirebon City and processed using RapidMiner software. The results show that the C4.5 algorithm can produce a classification model with an accuracy rate of 0,6167, enabling BAZNAS to make decisions more objectively, efficiently, and accurately. Therefore, the implementation of data mining using the C4.5 algorithm is proven to be effective in optimizing zakat distribution at BAZNAS Cirebon City. The comparison includes an assessment of criteria weighting, ranking accuracy, and decision transparency

**Keywords** : Data Mining, C4.5 Algorithm, Zakat, BAZNAS, Classification, Mustahik

**Abstrak**—Zakat merupakan salah satu rukun Islam yang memiliki fungsi penting dalam mengurangi kesenjangan sosial dan meningkatkan kesejahteraan masyarakat. Penyaluran zakat yang tepat sasaran merupakan faktor penting dalam meningkatkan efektivitas program BAZNAS Kota Cirebon dalam mengentaskan kemiskinan. Namun, proses seleksi penerima zakat selama ini masih dilakukan secara manual dan subjektif, sehingga berpotensi menimbulkan ketidaktepatan dalam penentuan mustahik. Penelitian ini bertujuan untuk mengoptimalkan proses penentuan penerima zakat dengan menerapkan pendekatan *data mining* menggunakan algoritma C4.5. Metode ini digunakan untuk mengklasifikasikan kelayakan calon penerima zakat berdasarkan sejumlah atribut sosial-ekonomi seperti penghasilan, tanggungan, pekerjaan, pendidikan, dan kondisi tempat tinggal. Data penelitian diperoleh dari BAZNAS Kota Cirebon dan diolah menggunakan perangkat lunak *RapidMiner*. Hasil penelitian menunjukkan bahwa algoritma C4.5 mampu menghasilkan model klasifikasi dengan tingkat akurasi sebesar 0,6167, sehingga dapat membantu BAZNAS dalam mengambil keputusan secara lebih objektif, cepat, dan tepat sasaran. Dengan demikian, penerapan *data mining* menggunakan algoritma C4.5 terbukti efektif dalam mengoptimalkan penyaluran zakat di BAZNAS Kota Cirebon. Hasil dari penelitian ini diharapkan dapat menjadi acuan bagi BAZNAS dalam mengembangkan sistem berbasis teknologi guna meningkatkan kualitas pengelolaan zakat. Dengan demikian, zakat dapat lebih tepat sasaran, akuntabel, dan memberikan dampak yang lebih signifikan dalam pengentasan kemiskinan dan peningkatan kesejahteraan umat.

**Kata Kunci** : Data Mining, Algoritma C4.5, Zakat, BAZNAS, Klasifikasi, Mustahik

### INTRODUCTION

Zakat merupakan rukun Islam yang berfungsi tidak hanya sebagai ibadah, tetapi juga

sebagai instrumen distribusi ekonomi untuk mengurangi kemiskinan dan kesenjangan sosial. Berdasarkan Undang-Undang Nomor 23 Tahun

2011 tentang Pengelolaan Zakat, zakat diarahkan sebagai bagian dari upaya penanggulangan kemiskinan melalui pendistribusian dana kepada mustahik. Di Indonesia, pengelolaan zakat dilaksanakan oleh BAZNAS sebagai lembaga resmi yang bertanggung jawab secara profesional, transparan, dan akuntabel[1], [2].

Namun, proses seleksi penerima zakat di BAZNAS Kota Cirebon masih didominasi metode manual, seperti observasi lapangan dan penilaian subjektif petugas. Kondisi ini menimbulkan berbagai permasalahan, antara lain potensi subjektivitas, ketidakkonsistenan penilaian, serta keterlambatan proses seleksi ketika jumlah pemohon meningkat. Di sisi lain, BAZNAS Kota Cirebon memiliki data historis mustahik dengan berbagai atribut sosial-ekonomi yang belum dimanfaatkan secara optimal sebagai dasar pengambilan keputusan berbasis data, sehingga diperlukan pendekatan analitis yang lebih sistematis dan objektif[3], [4].

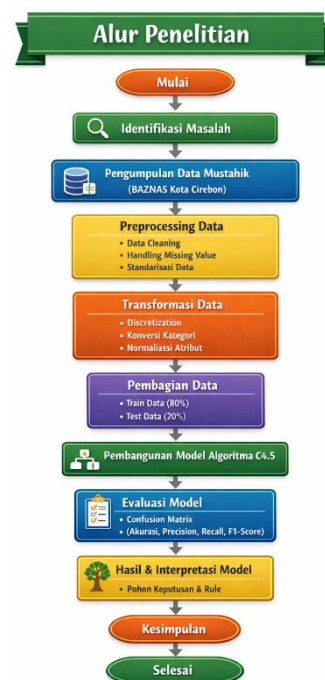
Di sinilah peran data mining menjadi relevan. Data mining merupakan proses penggalian informasi dan pola dari kumpulan data dalam jumlah besar menggunakan teknik statistik dan kecerdasan buatan. Dalam konteks penyaluran bantuan sosial, data mining telah terbukti mampu meningkatkan objektivitas dan efisiensi dalam proses seleksi penerima, karena keputusan dihasilkan berdasarkan pola data, bukan semata-mata penilaian subjektif manusia. Salah satu algoritma klasifikasi yang banyak digunakan dalam data mining adalah algoritma C4.5. Algoritma ini mampu membentuk model berupa pohon keputusan (decision tree) yang mengelompokkan data ke dalam kelas tertentu berdasarkan atribut- atribut yang dimilikinya. Kelebihan utama algoritma C4.5 adalah kemampuannya dalam menangani data numerik dan kategorikal, mengatasi missing value, serta menghasilkan aturan keputusan (rules) yang mudah dipahami oleh pengguna non-teknis, seperti petugas lapangan[5], [6].

Penelitian-penelitian sebelumnya menunjukkan bahwa algoritma C4.5 mampu memberikan tingkat akurasi tinggi dalam berbagai kasus klasifikasi penerima bantuan sosial. Namun, sebagian besar penelitian tersebut dilakukan di daerah dengan karakteristik sosial yang berbeda. Setiap wilayah memiliki kondisi ekonomi, pola pendapatan, jumlah tanggungan, dan tingkat kemiskinan yang unik. Oleh sebab itu, model yang efektif di suatu daerah belum tentu memberikan hasil yang sama di daerah lain[7], [8], [9].

Data tingkat kemiskinan per kecamatan di Kota Cirebon menunjukkan bahwa seluruh wilayah penelitian memiliki persentase penduduk miskin yang cukup signifikan, yaitu berkisar antara 7,5% hingga 9,2%, dengan jumlah penduduk miskin yang relatif besar serta komposisi kategori fakir dan miskin yang bervariasi di setiap kecamatan. Kondisi ini mencerminkan karakteristik mustahik yang heterogen dan kompleks, sehingga berpotensi menimbulkan subjektivitas dan ketidakkonsistenan apabila proses penetapan penerima zakat dilakukan secara manual. Oleh karena itu, Kota Cirebon menjadi lokasi penelitian yang relevan untuk penerapan algoritma C4.5 karena ketersediaan data sosial- ekonomi yang beragam dan bersumber dari BPS serta data internal BAZNAS, sehingga memungkinkan pengembangan model klasifikasi kelayakan mustahik yang lebih objektif, terukur, dan mendukung pengambilan keputusan penyaluran zakat secara akuntabel.

## MATERIALS AND METHODS

Tahap ini merupakan awal dari proses penelitian. Pada tahap ini peneliti menentukan topik penelitian, menentukan objek penelitian, serta menyiapkan rencana penelitian yang akan dilakukan.



Gambar 1. Alur Penelitian

Alur penelitian pada gambar tersebut menggambarkan tahapan sistematis yang dilakukan dalam proses penelitian untuk membangun model klasifikasi penerima zakat menggunakan algoritma

C4.5. Penelitian dimulai dari tahap identifikasi masalah, yaitu menganalisis permasalahan yang terjadi pada proses penentuan penerima zakat di BAZNAS Kota Cirebon yang masih dilakukan secara manual sehingga berpotensi menimbulkan subjektivitas dan ketidakkonsistenan dalam pengambilan keputusan. Setelah masalah diidentifikasi, tahap selanjutnya adalah pengumpulan data, yaitu mengumpulkan data mustahik yang diperoleh dari BAZNAS Kota Cirebon yang berisi atribut sosial ekonomi seperti pendapatan, pekerjaan, jumlah tanggungan, kondisi rumah, status kepemilikan rumah, beban pengeluaran, serta riwayat bantuan sebelumnya.

Tahap berikutnya adalah preprocessing data, yaitu proses pembersihan data untuk meningkatkan kualitas dataset sebelum digunakan dalam proses analisis. Pada tahap ini dilakukan penghapusan data duplikat, penanganan nilai kosong (missing value), serta standarisasi format data agar data memiliki struktur yang seragam dan dapat diproses dengan baik oleh algoritma data mining. Setelah data dibersihkan, dilakukan transformasi data yang bertujuan untuk mengubah data mentah menjadi format yang lebih terstruktur dan sesuai dengan kebutuhan algoritma C4.5. Transformasi ini meliputi proses discretization atau pengelompokan nilai numerik menjadi kategori tertentu, konversi data teks menjadi kategori standar, serta normalisasi atribut agar data memiliki skala yang lebih konsisten.

Setelah proses transformasi selesai, dataset kemudian memasuki tahap pembagian data menggunakan metode train-test split. Pada penelitian ini data dibagi menjadi dua bagian, yaitu 80% data training yang digunakan untuk membangun model klasifikasi dan 20% data testing yang digunakan untuk menguji performa model terhadap data baru. Tahap selanjutnya adalah pembangunan model menggunakan algoritma C4.5, yaitu salah satu metode decision tree yang mampu menghasilkan model klasifikasi dalam bentuk pohon keputusan berdasarkan perhitungan entropy, information gain, dan gain ratio. Model ini akan menghasilkan aturan keputusan yang dapat digunakan untuk menentukan apakah seorang calon mustahik layak atau tidak layak menerima zakat.

Setelah model terbentuk, dilakukan tahap evaluasi model untuk mengetahui tingkat kinerja model klasifikasi yang dihasilkan. Evaluasi dilakukan menggunakan confusion matrix dengan beberapa metrik pengukuran seperti akurasi, precision, recall, dan F1-score.

Hasil evaluasi ini memberikan gambaran mengenai kemampuan model dalam mengklasifikasikan data mustahik secara tepat. Tahap terakhir adalah interpretasi hasil dan penarikan kesimpulan, di mana hasil berupa pohon keputusan dan aturan klasifikasi dianalisis untuk mengetahui faktor-faktor yang paling berpengaruh dalam menentukan kelayakan penerima zakat. Kesimpulan penelitian kemudian dirumuskan untuk menjelaskan efektivitas penerapan algoritma C4.5 dalam membantu BAZNAS Kota Cirebon menentukan penerima zakat secara lebih objektif, cepat, dan tepat sasaran.

## RESULTS AND DISCUSSION

Tahapan ini membangun model C 4.5 dari data yang telah dibersihkan dan ditransformasikan, dalam tahapan ini dimulai dari menginisiasi Model, menghitung Information Gain, Implisit Pruning untuk mengatur membatasi kompleksitas, dan melatih untuk membangun struktur pohon keputusan menggunakan data latih.

```
# Inisialisasi Model Decision Tree Classifier (mirip C4.5 menggunakan kriteria 'entropy')
# Kriteria 'entropy' di scikit-learn menghitung Information Gain, yang merupakan prinsip utama C4.5.
# max_depth diatur untuk membatasi kompleksitas (implisit Pruning).
# min_samples_leaf diatur untuk pruning.
c45_model = DecisionTreeClassifier(
    criterion='entropy', # Menggunakan Entropy/Information Gain (inti dari C4.5)
    max_depth=5, # Pruning: Batasi kedalaman pohon untuk mencegah Overfitting
    min_samples_leaf=5, # Pruning: Minimal 5 sampel di daun
    random_state=42
)

# Membangun struktur pohon keputusan menggunakan data latih
c45_model.fit(X_train, y_train)
print("Model Decision Tree (C4.5) berhasil dilatih.")

# Uji Model: Prediksi pada data uji
y_pred = c45_model.predict(X_test)
```

Gambar 2 Pembangunan model C4.5

Evaluasi model C4.5 adalah proses untuk memvalidasi sejauh mana pohon keputusan yang dibentuk dapat melakukan tugas klasifikasi secara andal dan akurat pada data baru, menggunakan metrik standar seperti Akurasi, Presisi, Recall, dan Confusion Matrix.

```
Akurasi Model: 0.6167
Presisi Rata-rata: 0.6299
Recall Rata-rata: 0.6167
F1-Score Rata-rata: 0.6218
```

```
Confusion Matrix:
|-----| Layak | Tidak Layak |
|:-----:|-----:|
| Layak | 26 | 13 |
| Tidak Layak | 10 | 11 |
```

```
Classification Report (Presisi, Recall, F1-Score per Kelas):
precision recall f1-score support
```

```
Layak 0.72 0.67 0.69 39
Tidak Layak 0.46 0.52 0.49 21

accuracy 0.62 60
macro avg 0.59 0.60 0.59 60
weighted avg 0.63 0.62 0.62 60
```

Gambar 3 Evaluasi kinerja model

Hasil Akurasi, Presisi, Recall, dan Confusion Matrix.



Gambar 4 Confusion Matrix

Secara kritis, algoritma C4.5 menawarkan keunggulan dalam mengekstrak aturan keputusan yang mudah dipahami, namun memiliki keterbatasan seperti sensitivitas terhadap data noise dan overfitting jika dataset tidak seimbang, yang dapat mengurangi akurasi di dunia nyata (Han et al., 2012). Dalam realitas BAZNAS Kota Cirebon,

Dimana data mustahik sering kali tidak lengkap atau bervariasi kualitasnya, model ini mungkin menghasilkan aturan yang kurang generalizable, sehingga perlu validasi berkala untuk menghindari kesalahan klasifikasi yang memengaruhi keadilan distribusi zakat. Kritik lain adalah bahwa C4.5 mengasumsikan atribut independen, padahal dalam konteks sosial-ekonomi mustahik, faktor seperti kesehatan dan pendidikan saling terkait, yang bisa menyebabkan bias jika tidak ditangani dengan preprocessing data yang tepat (Quinlan, 1993).

Akurasi merupakan salah satu metrik utama dalam mengevaluasi performa sebuah model klasifikasi. Dalam konteks ini, model yang digunakan menghasilkan akurasi sebesar 0.6167 atau 61.67%, yang menunjukkan bahwa hampir dua pertiga dari seluruh prediksi yang dilakukan oleh model tersebut sesuai dengan kelas sebenarnya. Secara umum, nilai ini dapat dikategorikan sebagai akurasi sedang, bukan buruk tetapi juga belum mencapai tingkat akurasi yang ideal, terutama jika model digunakan untuk pengambilan keputusan penting seperti klasifikasi kelayakan, identifikasi risiko, atau penentuan status tertentu.

Nilai akurasi yang berada pada kisaran 60% sering kali mencerminkan bahwa model masih memiliki keterbatasan dalam membedakan karakteristik antara kelas yang ada. Hal ini dapat terjadi karena beberapa faktor, seperti kualitas data, ketidakseimbangan kelas (class imbalance), pemilihan fitur yang kurang optimal, atau penggunaan algoritma yang belum

sepenuhnya sesuai dengan pola dalam dataset. Dari hasil yang ditampilkan, terlihat adanya ketidakseimbangan jumlah antara kelas Layak dan Tidak Layak, yang secara otomatis memengaruhi performa model. Kelas Layak tampak memiliki data lebih banyak sehingga model lebih sering memprediksi kelas tersebut, baik benar maupun salah.

Akurasi sebesar 61,67% juga menunjukkan adanya kelemahan dalam proses klasifikasi yang dapat disebabkan oleh beberapa faktor. Pertama, kondisi data yang digunakan cenderung tidak homogen, di mana terdapat data yang tidak lengkap serta format nilai yang tidak seragam pada beberapa atribut sosial-ekonomi. Ketidakteraturan ini berpotensi menghambat algoritma dalam menemukan pola yang konsisten sehingga proses pemilihan atribut terbaik menjadi kurang optimal.

Kedua, ketidakseimbangan kelas menjadi penyebab utama turunnya akurasi. Jumlah data mustahik kategori "Layak" lebih dominan dibandingkan kategori "Tidak Layak" sehingga model cenderung melakukan prediksi mayoritas class. Kondisi ini menyebabkan peningkatan False Positive (layak diprediksi layak tetapi sebenarnya tidak layak) dan False Negative (tidak layak tetapi sebenarnya layak), sehingga menurunkan kualitas generalisasi model terhadap kasus baru.

Selain itu, atribut yang tersedia masih terbatas pada variabel dasar seperti pendapatan, pekerjaan, jumlah tanggungan, dan kondisi rumah. Absennya atribut tambahan seperti riwayat bantuan sebelumnya, kondisi kesehatan secara terukur, kepemilikan aset produktif, atau pengeluaran bulanan secara detail membuat model memiliki ruang informasi yang sempit dalam menghasilkan keputusan.

Algoritma C4.5 juga memiliki kecenderungan membangun struktur pohon terlalu spesifik pada pola data yang ada (overfitting) apabila tidak dilakukan pruning secara optimal. Hal ini memungkinkan model tampak cukup baik terhadap data pelatihan, namun menurun performanya ketika diuji pada data baru.

## CONCLUSION

Zakat merupakan rukun Islam yang berfungsi tidak hanya sebagai ibadah, tetapi juga sebagai instrumen distribusi ekonomi untuk mengurangi kemiskinan dan kesenjangan sosial. Berdasarkan Undang-Undang Nomor 23 Tahun 2011 tentang Pengelolaan Zakat, zakat diarahkan sebagai bagian dari upaya penanggulangan kemiskinan melalui pendistribusian dana kepada mustahik. Di Indonesia, pengelolaan zakat dilaksanakan oleh BAZNAS sebagai lembaga resmi yang bertanggung jawab secara profesional, transparan, dan

akuntabel. Berdasarkan hasil penelitian dan pembahasan mengenai penerapan algoritma C4.5 dalam optimasi penentuan penerima zakat pada BAZNAS Kota Cirebon, maka dapat diambil beberapa kesimpulan sebagai berikut:

Model klasifikasi kelayakan penerima zakat berhasil dibangun menggunakan algoritma C4.5 dengan memanfaatkan atribut sosial-ekonomi mustahik, seperti pendapatan, pekerjaan, jumlah tanggungan, dan kondisi tempat tinggal. Proses pembangunan model dilakukan melalui tahapan data mining yang sistematis, mulai dari preprocessing data hingga pembentukan pohon keputusan.

Hasil klasifikasi yang dihasilkan oleh algoritma C4.5 dapat diinterpretasikan secara jelas melalui aturan keputusan (IF-THEN), sehingga mampu mendukung pengambilan keputusan penyaluran zakat secara lebih objektif, konsisten, dan terukur. Model ini dapat digunakan sebagai sistem pendukung keputusan bagi BAZNAS Kota Cirebon dalam menentukan kelayakan mustahik.

Tingkat akurasi model klasifikasi yang dihasilkan sebesar 61,67%, yang termasuk dalam kategori akurasi sedang. Meskipun belum mencapai tingkat akurasi yang optimal, model berbasis C4.5 telah memberikan pendekatan yang lebih sistematis dan berbasis data dibandingkan metode seleksi manual yang selama ini digunakan, yang cenderung bersifat subjektif dan tidak terukur secara kuantitatif.

#### REFERENCE

- [1] Washilaturrizqi And Nurdiawan Odi, "Implementasi Algoritma C45 Untuk Menentukan Penerima Bantuan Sosial," *Jurnal Mahasiswa Teknik Informatika*, Vol. 7, No. 1, Pp. 373-380, 2023.
- [2] G. Taufik And D. Jatmika, "Penerapan Algoritma C45 Untuk Klasifikasi Keberhasilan Pengiriman Barang," *INOVTEK Polbeng - Seri Informatika*, Vol. 6, No. 1, P. 12, 2021, Doi: 10.35314/Isi.V6i1.1446.
- [3] M. Nonsi Tentua And K. Nurbagja, "Komparasi Metode Naïve Bayes Dan C45 Pada Prediksi Pelanggan Deposito Berjangka," *Jurnal Dinamika Informatika*, Vol. 11, No. 1, 2022, [Online]. Available: <https://www.kaggle.com/Sonujha090/Bank-Marketing>.
- [4] H. Hasanah, "Perbandingan Tingkat Akurasi Algoritma Support Vector Machines (SVM) Dan C45 Dalam Prediksi Penyakit Jantung," 2023.
- [5] Dhamayanti, "Penentuan Pemberian Reward Bagi Karyawan Berprestasi Di Lingkungan Universitas Indo Global Mandiri Dengan Algoritma C45," *Jurnal Ilmiah Informatika Global*, Vol. 9, No. 1, 2018.
- [6] J. Amarda, N. Suarna, And O. Nurdiawan, "KLASIFIKASI INDUSTRI KERAJINAN BERDASARKAN PROVINSI MENGGUNAKAN METODE DECISION TREE UNTUK MEMBANDINGKAN PENGUJIAN ALGORITMA ID3 ALGORITMA C45 DAN ALGORITMA CART," 2022.
- [7] B. Santoso, "Penerapan Decision Tree Dalam Pengelompokan Data Penjualan," *Jurnal Sains Komputer Dan Informatika*, 2022.
- [8] I. Arfyanti, M. Fahmi, And P. Adytia, "Penerapan Algoritma Decision Tree Untuk Penentuan Pola Penerima," *Building Of Informatics Technology And Science*, Vol. 4, No. 3, Pp. 1196-1201, 2022, Doi: 10.47065/Bits.V4i3.2275.
- [9] S. A. Arnomo, A. A. Fajrin, Y. Siyamto, S. Fairuz, And N. Sadikin, "Evaluasi Model Decision Tree Pada Keputusan Kelayakan Kredit," 2022. [Online]. Available: <http://journal.aptikomkepri.org/index.php/JDDAT200JURNALDESAINANANALISIS TEKNOLOGI>