

# Kinerja Metode C4.5 dalam Penyaluran Bantuan Dana Bencana

1<sup>st</sup>\* Ahmad Akhlaqul karim  
Universitas Muslim Indonesia  
Fakultas Ilmu Komputer  
Makassar, Indonesia  
Ahmad.fikom29@gmail.com

2<sup>nd</sup> Yulita Salim  
Universitas Muslim Indonesia  
Fakultas Ilmu Komputer  
Makassar, Indonesia  
yulita.salim@umi.ac.id

3<sup>rd</sup> Huzain Azis  
Universitas Muslim Indonesia  
Fakultas Ilmu Komputer  
Makassar, Indonesia  
huzain.azis@umi.ac.id

**Abstrak**—Secara geografis Indonesia membentang dari 6<sup>o</sup> LU sampai 11<sup>o</sup> LS dan 92<sup>o</sup> sampai 142<sup>o</sup> BT, terdiri dari pulau-pulau besar dan kecil yang jumlahnya kurang lebih 17.504 pulau. Tiga per-empat wilayahnya adalah laut (5,9 juta km<sup>2</sup>), dengan panjang garis pantai 95.161 km, perairan terpanjang kedua setelah Kanada. Bencana alam kadang terjadi di suatu wilayah yang kadangkala terjadi di suatu wilayah. Seringnya kesalahan pembagian penyaluran bantuan dana bagi wilayah yang sangat membutuhkan ke lokasi bencana alam yang terjadi di berbagai wilayah. Dalam Penelitian ini bantuan yang disalurkan berdasarkan faktor kebutuhan sembako dan kebutuhan obat-obatan, pemimpin. Oleh karena itu faktor-faktor yang terkait dengan penyaluran bantuan akan dianalisa. Dalam penelitian ini metode C4.5 digunakan untuk menganalisa jumlah bantuan dana yang akan disalurkan ke wilayah bencana. Sebanyak 20 data sampel korban bencana alam telah di analisa pada wilayah Makassar. Berdasarkan Hasil eksperimen terdapat beberapa faktor utama yaitu pasokan sembako dan kebutuhan obat-obatan dalam menentukan jumlah penyaluran bantuan bencana ke wilayah tertentu. Perbaikan kinerja dapat menggunakan beberapa metode, salah satunya dengan menggunakan Metode Best-fit.

**Kata Kunci**-C4.5; Bencana Alam; Penyaluran Bantuan

## I. PENDAHULUAN

Bencana alam bisa saja terjadi di suatu wilayah di bagian Indonesia. Seringnya kesalahan pembagian atau kurang tepatnya wilayah yang sangat membutuhkan penyaluran bantuan dana ke lokasi bencana alam yang terjadi di berbagai wilayah.

Pendistribusian bantuan biasanya hanya menyalurkan bantuan kepada wilayah yang terkena bencana dimana wilayah tersebut terdiri dari beberapa kelurahan/desa atau terdiri dari beberapa kecamatan yang terkena bencana. Maka dari itu pendistribusian biasanya hanya pemberian langsung tanpa mengetahui wilayah mana yang sangat membutuhkan bantuan tersebut. Maka dibutuhkan sebuah sistem untuk menentukan wilayah yang sangat membutuhkan bantuan terutama bagi oknum pemerintah yang memiliki laporan pertanggung jawaban dari setiap dana yang dikeluarkan oleh pemerintah.

Algoritma C4.5 merupakan algoritma yang digunakan untuk membentuk pohon keputusan (Decision Tree). Klasifikasi

merupakan salah satu proses pada data mining yang bertujuan untuk menemukan pola yang berharga dari data yang berukuran relatif besar hingga sangat besar. Algoritma C4.5 sendiri merupakan pengembangan dari algoritma ID3. Cara kerja algoritma C4.5 ialah memberikan keputusan berdasarkan beberapa aspek atau beberapa poin yang memiliki nilai sesuai tingkat dari aspek tersebut.

## II. METODOLOGI

### A. Pohon Keputusan

Manusia selalu dihadapkan dengan berbagai macam masalah dari berbagai bidang kehidupan. Masalah ini juga memiliki variasi tingkat kesulitannya. Untuk menghadapi masalah ini manusia mulai mengembangkan sebuah sistem untuk membantu mereka menyelesaikan masalah-masalah ini, salah satu sistem tersebut adalah pohon keputusan [2].

Pohon keputusan merupakan metode klasifikasi dan prediksi yang sangat kuat dan terkenal. Pohon-keputusan bekerja dengan melibatkan/menggunakan satu set data pelatihan untuk menghasilkan pohon keputusan dengan mengklasifikasikan data pelatihan yang merepresentasikan aturan-aturan[3].

Pohon keputusan adalah metode klasifikasi dan prediksi yang sudah terbukti powerful dan sangat terkenal. Metode ini berfungsi untuk mengubah fakta menjadi pohon keputusan yang merepresentasikan aturan yang dapat mudah dimengerti dengan bahasa alami. Proses dari pohon keputusan ini dimulai dari *node* akar hingga *node* daun yang dilakukan secara rekursif dimana setiap percabangan menyatakan kondisi dan setiap ujung pohon akan menyatakan keputusan.

Arsitektur pohon keputusan dibuat sedemikian rupa agar menyerupai pohon asli, dimana terdapat beberapa bagian yaitu:

a. *Root Node*: Node ini terletak pada bagian paling atas dari pohon keputusan.

b. *Leaf Node*: Node ini merupakan node yang terletak pada ujung pohon. Node ini hanya memiliki satu input dan tidak memiliki output. Menurut Triisant, pohon keputusan ini memiliki beberapa kekurangan dan kelebihan, yakni:

1) Kelebihan Pohon Keputusan:

*Daerah pengambilan keputusan yang kompleks dapat diubah menjadi sederhana.* Dapat menghilangkan perhitungan yang tidak penting karena proses pengujian hanya berdasarkan kriteria yang diperlukan saja. Proses pemilihan fitur dari internal node yang berbeda lebih fleksibel. Fitur yang telah dipilih ini akan menjadi pembeda antara kriteria yang satu dengan kriteria lainnya.

Metode ini dapat menghindari munculnya permasalahan dengan cara menggunakan kriteria dengan jumlah yang sedikit pada node internal tanpa mengurangi kualitas keputusan yang dihasilkan [2].

2) Kekurangan Pohon Keputusan :

Dapat terjadi overlap apabila hasil keputusan dan kriteria yang digunakan jumlahnya sangat banyak. Hal ini juga dapat berakibat bertambahnya waktu yang digunakan untuk pengambilan keputusan dan jumlah memori yang dibutuhkan semakin tinggi. Akumulasi jumlah error dari setiap tingkat pohon keputusan besar. Mendesain pohon keputusan yang optimal sulit. Kualitas keputusan yang didapatkan sangat tergantung dengan bagaimana pohon tersebut didesain.

B. Data Mining

Data mining adalah proses yang menggunakan statistik, matematika, kecerdasan buatan, dan *machine learning* untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai database besar[4].

Data mining adalah suatu proses menemukan hubungan yang berarti, pola, dan kecenderungan dengan memeriksa dalam sekumpulan besar data yang tersimpan dalam penyimpanan dengan menggunakan teknik pengenalan pola seperti teknik statistik dan matematika.

Beberapa solusi yang bisa diselesaikan dengan data mining diantaranya :

1) Menembak target pasar

Data mining dapat melakukan pengelompokan (*clustering*) dari model-model pembeli dan melakukan klasifikasi terhadap setiap pembeli sesuai dengan karakteristik yang diinginkan seperti kesukaan yang sama, tingkat penghasilan yang sama, kebiasaan membeli dan karakteristik lainnya.

2) Melihat pola beli pemakai dari waktu ke waktu

Data mining dapat digunakan untuk melihat pola beli seseorang dari waktu ke waktu. Sebagai contoh, ketika seseorang menikah bisa saja dia kemudian memutuskan pindah dari single account ke joint account (rekening bersama) dan kemudian setelah itu pola belinya berbeda dengan ketika dia masih bujangan.

3) Cross-Market Analysis

dapat memanfaatkan data mining untuk melihat hubungan antara penjualan satu produk dengan produk lainnya.

4) Profil Customer

Data mining dapat membantu untuk melihat profil customer/pembeli/nasabah sehingga kita dapat mengetahui kelompok customer tertentu suka membeli produk apa saja [5].

C. Metode C4.5

Algoritma C4.5 merupakan algoritma yang digunakan untuk melakukan proses klasifikasi data dengan menggunakan teknik pohon keputusan. Algoritma C4.5 merupakan ekstensi dari algoritma ID3 dan menggunakan prinsip *decision tree* yang mirip [2].

Tingkat akurasi yang didapatkan menggunakan algoritma C4.5 ini sudah cukup tinggi, oleh karena itu dapat disimpulkan algoritma ini cocok untuk diimplementasikan pada penelitian yang melibatkan proses perekrutan [6].

Ide dasar dari algoritma ini adalah pembuatan pohon keputusan berdasarkan pemilihan atribut yang memiliki prioritas tertinggi atau dapat disebut memiliki nilai gain tertinggi berdasarkan nilai entropy atribut tersebut sebagai poros atribut klasifikasi [7].

Algoritma ini sudah sangat terkenal dan disukai karena memiliki banyak kelebihan. Kelebihan ini misalnya dapat mengolah data numerik dan diskret, dapat menangani nilai atribut yang hilang, menghasilkan aturan-aturan yang mudah diinterpretasikan dan performanya merupakan salah satu yang tercepat dibandingkan dengan algoritma lain. Ide dasar dari algoritma ini adalah pembuatan pohon keputusan berdasarkan pemilihan atribut yang memiliki prioritas tertinggi atau dapat disebut memiliki nilai *gain* tertinggi

berdasarkan nilai *entropy* atribut tersebut sebagai poros atribut klasifikasi. Kemudian secara rekursif cabang-cabang pohon diperluas sehingga seluruh pohon terbentuk. Menurut kamus IGI Global (*International Publisher of Progressive Academic*), *entropy* adalah jumlah data yang tidak relevan terhadap informasi dari suatu kumpulan data. *Gain* adalah informasi yang didapatkan dari perubahan *entropy* pada suatu kumpulan data, baik melalui observasi atau bisa juga disimpulkan dengan cara melakukan partisipasi terhadap suatu *set* data[2].

Berdasarkan apa yang ditulis oleh Jefri [10], terdapat empat langkah dalam proses pembuatan pohon keputusan pada algoritma C4.5, yaitu:

- 1) Memilih atribut sebagai akar
- 2) Membuat cabang untuk masing-masing nilai
- 3) Membagi setiap kasus dalam cabang
- 4) Mengulangi proses dalam setiap cabang sehingga semua kasus dalam cabang memiliki kelas yang sama[2].

Untuk memilih atribut sebagai akar harus dicari nilai *gain* tertinggi di antara atribut-atribut yang ada, ditunjukkan pada persamaan (1)

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

Keterangan :

S : Himpunan Kasus

A : Atribut

S<sub>i</sub> : Jumlah Kasus pada Partisi ke-i

n : Jumlah Partisi Atribut A

|S| : Jumlah Kasus dalam S

Adapun untuk mencari nilai *Entropy*, ditunjukkan pada persamaan (2)

$$Entropy(S) = \sum_{i=1}^n -pi * log_2 pi \quad (2)$$

Keterangan :

S : Himpunan Kasus

A : Fitur

N : Jumlah Partisi S

pi : Proporsi dari Si terhadap S[8]

#### D. Ukuran Kinerja C4.5

Peran metode C4.5 pada aplikasi ini adalah menentukan jumlah dana yang disalurkan pada wilayah yang terkena bencana alam. Dengan adanya pohon keputusan sebagai output akhir dalam metode C4.5 membuat penyaluran bantuan dana sesuai dengan yang diperlukan oleh wilayah yang terkena bencana.

### III. HASIL DAN PEMBAHASAN

#### A. Hasil Implementasi

Dari hasil implementasi maka dapat disimpulkan bahwa penggunaan metode C4.5 untuk penentuan penyaluran bantuan bencana sangat dibutuhkan dalam proses penentuan keputusan. Sebagaimana mana hasil akhir dari Kinerja C4.5 dalam aplikasi ini. Sesuai dengan output yang akan di bahas pada poin pengujian metode C4.5.

#### B. Pengujian Metode C4.5

TABEL I. KEPUTUSAN PENYALURAN BANTUAN

No	Wilayah	Sembako	Obat	Penyaluran
1	Antang	Lebih	Lengkap	YA
2	Tamalanrea	Kurang	Kurang Lengkap	YA

No	Wilayah	Sembako	Obat	Penyaluran
3	Antang	Kurang	Lengkap	YA
4	Batua	Cukup	Kurang Lengkap	TIDAK
5	Batua	Kurang	Lengkap	TIDAK
6	Tamalanrea	Lebih	Lengkap	YA
7	Antang	Kurang	Lengkap	YA
8	Tamalanrea	Lebih	Kurang Lengkap	TIDAK
9	Tamalanrea	Lebih	Kurang Lengkap	TIDAK
10	Batua	Lebih	Lengkap	YA
11	Antang	Lebih	Kurang Lengkap	YA
12	Antang	Kurang	Lengkap	YA
13	Tamalanrea	Kurang	Kurang Lengkap	YA
14	Tamalanrea	Cukup	Lengkap	TIDAK
15	Batua	Kurang	Lengkap	TIDAK
16	Batua	Lebih	Lengkap	YA
17	Antang	Kurang	Kurang Lengkap	YA
18	Tamalanrea	Lebih	Kurang Lengkap	TIDAK
19	Batua	Lebih	Lengkap	TIDAK
20	Antang	Lebih	Lengkap	YA

Dengan adanya Tabel I sebagai data uji maka akan dilakukan perhitungan entropy dan gain untuk menentukan pohon keputusan secara berulang kali dengan setiap data di eliminasi setelah terseleksi digunakan. Adapun perhitungan gain dan entropy pertama, dapat dilihat pada Tabel II.

TABEL II. PERHITUNGAN ENTROPY DAN GAIN PERTAMA

	Jumlah kasus(s)	Tidak (s1)	Ya (s2)	Entropy	Gain
TOTAL	20	8	12	0.970950594	
Wilayah					
Antang	7	0	7	0	0.393369645
Tamalanrea	7	5	2	0.863120569	
Batua	6	4	2	0.918295834	
Sembako					
Lebih	10	4	6	0.970950594	0.160964047
Cukup	2	2	0	0	
Kurang	8	2	6	0.811278124	
Obat					
Lengkap	12	2	10	0.650022422	0.256425892
Tidak Lengkap	8	6	2	0.811278124	

Adapun contoh perhitungan *Entropy* dan gain adalah sebagai berikut:

$$Entropy(S) = \sum_{i=1}^n -pi * log_2 pi \quad (3)$$

$$Entropy(Total) = -\frac{8}{20} * log_2(\frac{8}{20}) + -\frac{12}{20} * log_2(\frac{12}{20}) = 0.970950594 \quad (4)$$

$$Gain(S, A) = Entropy(S) - \sum_{i=1}^n \frac{|S_i|}{|S|} * Entropy(S_i) \quad (5)$$

$$Gain(Total, Wilayah) = 10 - \left( \left( \frac{7}{20} \right) * 0.970950594 \right) - \left( \left( \frac{7}{20} \right) * 0.970950594 \right) - \left( \left( \frac{6}{20} \right) * 0.970950594 \right) = 0.393369645 \quad (6)$$

Untuk perhitungan yang pertama telah ditentukan bahwa yang menjadi sebagai akar dari pohon keputusan adalah Wilayah, dikarenakan Wilayah memiliki nilai gain tertinggi dari poin penilaian yang lainnya yaitu 0.295461844.

Perhitungan selanjutnya Wilayah tidak akan dicek nilai Gain-nya, namun akan dijadikan sebagai acuan perhitungan untuk poin sembako dan obat. Adapun proses perhitungan selanjutnya akan ditunjukkan pada tabel III berikut :

TABEL III. PERHITUNGAN ENTROPY DAN GAIN KEDUA

Node		Jumlah kasus(s)	Tidak(s1)	Ya(S2)	Entropy	Gain
Wilayah						
1 Tamalanrea		14	4	10	0.863120569	
Sembako						
	Lebih	7	3	4	0.985228136	0.370506501
	Cukup	1	1	0	0	
	Kurang	6	0	6	0	
Obat						
	Lengkap	7	1	6	0.591672779	0.074670111
	Kurang Lengkap	7	3	4	0.985228136	

Untuk perhitungan yang kedua telah ditentukan bahwa yang menjadi sebagai kelanjutan dari akar pohon keputusan adalah sembako, dikarenakan Wilayah memiliki nilai gain tertinggi dari poin penilaian yang lainnya yaitu 0.370506501.

Berdasarkan tabel III maka perhitungan ke tiga tidak diperlukan lagi, dikarenakan sudah dapat terlihat jelas hasil perhitungan gain dan entropynya maka telah di tentukan bahwa obat yang menjadi cabang selanjutnya dari pohon keputusan tersebut dimana Lengkap adalah YES dan Kurang lengkap adalah NO.

Dari hasil beberapa perhitungna pada table II dan III maka akan dihasilkan sebuah pohon keputusan yang dimana akan menjadi keputusan atau hasil akhir dari metode atau algoritma C4.5 tersebut. Adapun pohon keputusannya ditunjukkan pada Gambar 1

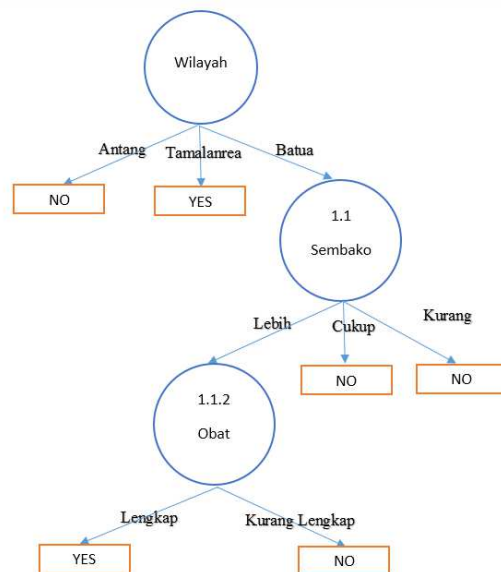
#### IV. KESIMPULAN

Berdasarkan hasil dari pengujian yang menggunakan data uji maka dapat ditarik sebuah kesimpulan bahwa kinerja metode C4.5 dalam proses perhitungan aplikasi tersebut membawakan sebuah hasil yang memuaskan yaitu mendukung

keputusan dalam penyaluran bantuan bencana. Dengan bantuan metode C4.5 dapat membantu proses perhitungan pada aplikasi yang digunakan untuk mempermudah pengambilan keputusan khususnya kepada aplikasi tersebut.

Dari gambar IV dapat dilihat bahwa kita akan menyalurkan bantuan kepada wilayah batua dengan nilai Entropy 0.918295834, dengan sembako yang lebih dikarenakan nilai Entropy yang tinggi yaitu 0.985228136 dan obat yang lengkap dengan nilai Entropy yaitu 0.811278124.

Gambar. 1. Pohon Keputusan



#### DAFTAR PUSTAKA

- [1] R. Lasabuda, "Pembangunan wilayah pesisir dan lautan dalam perspektif Negara Kepulauan Republik Indonesia," *J. Ilm. Platax*, vol. 1, no. 2, pp. 92–101, 2013.
- [2] F. F. Harryanto and S. Hansun, "Penerapan Algoritma C4 . 5 untuk Memprediksi Penerimaan Calon Pegawai Baru di PT WISE," *Jatiji*, vol. 3, no. 2, pp. 95–103, 2017.
- [3] T. H. Pudjiantoro, U. Jenderal, and A. Yani, "Penentuan Penanganan Kasus Terhadap Penyakit Berdasarkan Gejala Menggunakan Case Base Reasoning dan Algoritma Nearest Neighbor (Studi kasus : Klinik Citra Medika Cianjur)," no. February, 2018.
- [4] L. Swastina, "Penerapan Algoritma C4 . 5 Untuk Penentuan Jurusan Mahasiswa," vol. 2, no. 1, 2013.
- [5] T. B. Santoso, "Analisa Dan Penerapan Metode C4.5 Untuk Prediksi Loyalitas Pelanggan," *CEUR Workshop Proc.*, vol. 1542, no. 1, pp. 33–36, 2015.
- [6] Jefri and Kusriani, "Implementasi Algoritma C4.5 Dalam Aplikasi Untuk Memprediksi Jumlah Mahasiswa Yang Mengulang Mata Kuliah Di STMIK AMIKOM Yogyakarta," *Naskah Publ.*, pp. 1–14, 2013.
- [7] Marwana, "Algoritma c4.5 untuk simulasi prediksi kemenangan dalam pertandingan sepakbola," pp. 53–58.
- [8] A. Zulkifli, "Metode C45 Untuk Mengklarifikasi Pelanggan Perusahaan Telekomunikasi Seluler," *Riau J. Comput. Sci.*, vol. 2, no. 1, pp. 65–76, 2016.