

Penerapan Decision Tree Algoritma C4.5 dalam Penentuan Izin Pembongkaran Muatan Kapal

1) **Jaka Kusuma**

Magister Ilmu Komputer Universitas Potensi Utama, Sumatera Utara, Indonesia
E-Mail: jakakusuma41@gmail.com

2) **Abwabul Jinan**

Magister Ilmu Komputer Universitas Potensi Utama, Sumatera Utara, Indonesia
E-Mail: abwabul.jinan7@gmail.com

2) **Zakarias Situmorang**

Universitas Katolik Santo Thomas, Sumatera Utara, Indonesia
E-Mail: zakarias65@yahoo.com

ABSTRACT

Along with the increasing number of bulk cargoes that are dismantled every year at belawan port and for the creation of services in accordance with expectations, it is necessary to develop services in support of indonesia's logistics improvement readiness, especially in terms of demolition. Utilization of machine learning using the C4.5 algorithm can make it easier to conduct selection and classification of the feasibility of ships that get permission for demolition activities. The use of the C4.5 algorithm will produce a decision tree that can equalize the results of data mining, so that the information obtained from the data will be easier to identify in testing methods using the Orange Data Mining tool. The results obtained by the C4.5 algorithm in the form of a decision tree with an accuracy value of 84%, 90% precision and 84% recall.

Keyword : Data Mining, C4.5, Izin Pembongkaran Muatan Kapal

PENDAHULUAN

Dinamika perekonomian nasional dan global menuntut Indonesia untuk selalu siap menghadapi perubahan. Berbagai strategi akselerasi pertumbuhan harus diupayakan untuk memaksimalkan potensi Indonesia, termasuk di sektor maritim. Indonesia ialah negara kepulauan terbesar didunia yang mana mempunyai banyak pelabuhan yang seluruh layanannya saling terintegrasi. PT Pelabuhan Indonesia merupakan perusahaan yang berkecimpung pada bidang logistik, terutama pada hal pengelolaan serta pengembangan pelabuhan [1].

Transportasi laut merupakan urat nadi perekonomian negara kepulauan, khususnya Indonesia. Keberadaan transportasi laut di Indonesia dapat memperlancar kegiatan perekonomian di seluruh wilayah Indonesia, sehingga dapat mendorong pertumbuhan ekonomi di wilayah-wilayah Indonesia. Angkutan laut merupakan sarana angkutan yang efisien untuk perdagangan dunia. Barang dalam jumlah besar dalam dan luar negeri banyak menggunakan angkutan laut karena biaya pengangkutan barang jauh lebih murah, karena dapat diangkut dalam jumlah banyak dalam waktu yang bersamaan [2].

Salah satu pelayanan jasa yang ada di PT Pelabuhan Indonesia adalah terminal curah kering. Terminal curah kering merupakan fasilitas pelabuhan yang terdiri atas dermaga yang digunakan sebagai tempat kapal bersandar, lokasi penumpukan atau pembongkaran curah

kering yang dilengkapi dengan peralatan yang menunjang untuk kegiatan bongkar muat curah kering seperti Harbour Mobile Crane (HMC), Conveyor, dan gudang.

Seiring meningkatnya pembongkaran jenis kargo curah khususnya kedelai, gandum dan bungkil kedelai setiap tahunnya di pelabuhan Belawan serta untuk terciptanya pelayanan yang sesuai dengan harapan, maka perlu dilakukan pengembangan layanan dalam mendukung kesiapan peningkatan logistik Indonesia terutama dalam hal pembongkaran. Pemanfaatan *machine learning* menggunakan algoritma C4.5 dapat mempermudah dalam melakukan seleksi dan klasifikasi kelayakan kapal yang mendapat izin kegiatan pembongkaran. Penggunaan algoritma C4.5 akan menciptakan suatu pohon keputusan yang dapat menggambarkan hasil dari *data mining*, sehingga informasi yang didapatkan dari data tersebut akan lebih mudah dalam identifikasi [3].

Diharapkan dengan penelitian ini dapat membantu perusahaan sehingga tidak butuh waktu yang lama dalam penentuan izin kegiatan pembongkaran muatan kapal karena dalam penerapan algoritma C4.5 akan dapat membentuk suatu pola keputusan apakah kapal tersebut dinyatakan "Izinkan Bongkar" atau "Pending Bongkar".

BAHAN DAN METODE

Data mining merupakan metode yang dipergunakan untuk mengekstrak masalah

prediktif yang tersembunyi pada suatu database. Secara umum, *data mining* dapat dibagi menjadi dua kategori utama, yaitu:

1. *Deskriptive mining*, yaitu proses menemukan fitur-fitur kritis yang berasal dari data dalam database. Teknik penambangan data meliputi penambangan deskriptif, yaitu asosiasi, pengelompokan, dan penambangan sekuensial.
2. *Predictive*, yaitu proses penemuan pola yang berasal dari suatu data menggunakan variabel yang berbeda di masa yang akan datang [4].

Algoritma C4.5 merupakan sebuah algoritma yang terkenal digunakan oleh banyak peneliti di seluruh dunia. Algoritma C4.5 merupakan pengembangan lanjutan dari algoritma ID3 yang dikembangkan oleh J. Rose Quinlan [5]. Adapun cara algoritma C4.5 membuat pohon keputusan sebagai berikut [6]:

- a. Pilihlah atribut yang akan digunakan sebagai akar dan buat cabang untuk setiap nilai.
- b. Buatlah cabang untuk setiap nilai.
- c. Bagilah masalah pada sebuah cabang.
- d. Ulangi proses untuk setiap cabang sampai semua kasus di cabang memiliki kelas yang sama.

Saat memilih atribut yang akan dipilih sebagai *root*, didasarkan pada hasil nilai *gain* tertinggi dari atribut yang ada. Adapun cara untuk menghitung *gain* menggunakan rumus:

$$Gain(S, A) = Entropy(S) - \sum_{i=1}^n \frac{Si}{S} \times Entropy(Si)$$

Keterangan :

- S : Himpunan Kasus
- A : Atribut
- N : Jumlah Partisi Atribut A
- Si : Jumlah Kasus pada Partisi ke-i
- S : Jumlah Kasus dalam S

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

Keterangan :

- S : Himpunan Kasus
- A : Fitur
- N : Jumlah partisi S
- pi : Proporsi dari Si terhadap S

Decision tree merupakan representasi sederhana dari teknik klasifikasi untuk sejumlah kelas yang terbatas di mana simpul interior dan akar ditandai dengan nama atribut, lereng ditandai dengan nilai atribut, dan simpul tubuh dan daun ditandai dengan kelas yang tidak sejajar [7].

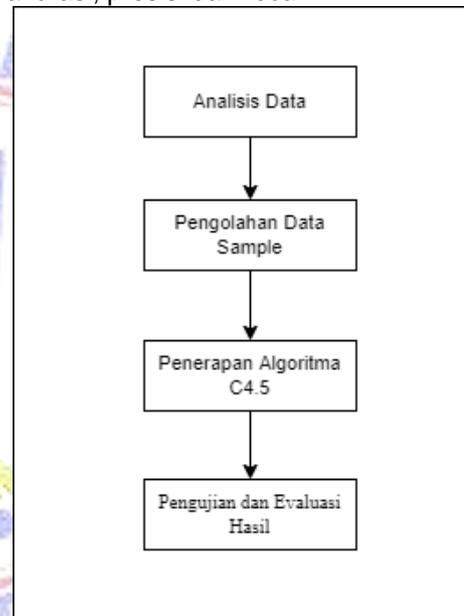
METODE PENELITIAN

Adapun tahapan kerja dalam proses penelitian yaitu :

1. Analisis Data
 Pada tahap ini, langkah pertama adalah menentukan rumusan masalah dari

penelitian, kemudian semua masalah yang ada dianalisis untuk menemukan cara pemecahan masalah yang ada.

2. Pengolahan Data Sample
 Tahap selanjutnya adalah pengolahan data, yang melibatkan beberapa langkah untuk membuat pohon keputusan. Pengolahan data sampel dilakukan setelah pembersihan data, integrasi dan transformasi data, dan teknik reduksi data.
3. Penerapan Algoritma C4.5
 Setelah data terkumpul, dilakukan analisis data untuk mengadaptasi pengolahan data dengan algoritma C4.5 Implementasi algoritma C4.5 dilakukan dengan aplikasi Orange Data Mining.
4. Pengujian dan Evaluasi Hasil
 Suatu pengujian dilakukan untuk menentukan apakah pemeriksaan yang dilakukan telah mencapai tujuan yang diharapkan. Setelah hasil klasifikasi diperoleh hasilnya dinilai dengan validasi silang (*confusion matrix*) untuk menentukan akurasi, presisi dan *recall*.



Gambar 1. Tahapan Kerja Penelitian

HASIL DAN PEMBAHASAN

Perhitungan Algoritma C4.5

Akan dilakukan proses perhitungan menggunakan algoritma C4.5 dengan atribut yang digunakan untuk penentuan izin kegiatan pembongkaran muatan kapal yaitu sandar kapal, dokumen SPB, inpeksi karantina, HMC1, HMC2, Conveyor, gudang, dimana pada atribut tersebut akan mempengaruhi hasil dari perhitungan *entropy* dan *gain*. Terdapat 32 data latih yang mana data tersebut akan menjadi data awal untuk perhitungan dengan rumus algoritma C4.5.

Berikut merupakan data latih yang akan nantinya akan digunakan untuk menentukan kriteria izin kegiatan pembongkaran muatan kapal dengan algoritma C4.5 :

Tabel 1. Data Latih

No.	Sandar Kapal	Dokumen SPB	Inpeksi Karantina	HMC 1	HMC 2	Conveyor	Gudang	Status
1	SUDAH	SUDAH	SUDAH	BAIK	BAIK	BAIK	KOSONG	Izinkan Bongkar
2	SUDAH	BELUM	BELUM	PROBLEM	PROBLEM	PROBLEM	TERISI	Pending Bongkar
3	SUDAH	SUDAH	BELUM	PROBLEM	PROBLEM	PROBLEM	TERISI	Pending Bongkar
30	SUDAH	BELUM	SUDAH	PROBLEM	BAIK	BAIK	TERISI	Pending Bongkar
31	SUDAH	BELUM	BELUM	PROBLEM	BAIK	BAIK	TERISI	Pending Bongkar
32	BELUM	BELUM	BELUM	PROBLEM	BAIK	BAIK	TERISI	Pending Bongkar

Data yang ada lalu akan diproses menggunakan rumus untuk mencari nilai dari *entropy* dan nilai *gain* tertinggi pada masing-masing atribut. Berikut merupakan perhitungan untuk mencari simpul akar dari pohon keputusan :

Tabel 2. Perhitungan Node 1 Algoritma C4.5

Node 1			Jumlah	Izinkan Bongkar	Pending Bongkar	Entropy	Gain
	Total			32		3	29
SANDAR KAPAL							
		SUDAH	18	3	15	0,6500	0,0832
		BELUM	14	0	14	0,0000	
DOKUMEN SPB							
		SUDAH	19	3	16	0,6292	0,0752
		BELUM	13	0	13	0,0000	
INPEKSI KARANTINA							
		SUDAH	22	3	19	0,5746	0,0538
		BELUM	10	0	10	0,0000	
HMC 1							
		BAIK	16	2	14	0,5436	0,0084
		PROBLEM	16	1	15	0,3373	
HMC 2							
		BAIK	17	2	15	0,5226	0,0056
		PROBLEM	15	1	14	0,3534	
CONVEYOR							
		BAIK	24	3	21	0,5436	0,0412
		PROBLEM	8	0	8	0,0000	
GUDANG							
		TERISI	15	0	15	0,0000	0,0917
		KOSONG	17	3	14	0,6723	

Berdasarkan perhitungan *entropy* dan *gain* node 1 pada table 2, nilai *gain* tertinggi yaitu atribut Gudang dengan dua nilai atribut yaitu "Terisi" dan "Kosong". Nilai dari atribut Gudang "Terisi" telah mengklasifikasikan kasus menjadi sebuah keputusan yaitu "Pending Bongkar". Langkah berikutnya adalah menghitung cabang pada node 1.1 yang terdiri dari dua simpul.

Tabel 3. Perhitungan Node 1.1 Algoritma C4.5

Node 1.1			Jumlah	Izinkan Bongkar	Pending Bongkar	Entropy	Gain
	Total			17		3	14
SANDAR KAPAL							
		SUDAH	7	3	4	0,9852	0,2666
		BELUM	10	0	10	0,0000	
DOKUMEN SPB							
		SUDAH	10	3	7	0,8813	0,1539
		BELUM	7	0	7	0,0000	
INPEKSI KARANTINA							
		SUDAH	12	3	9	0,8113	0,0996
		BELUM	5	0	5	0,0000	
HMC 1							
		BAIK	10	2	8	0,7219	0,0040
		PROBLEM	7	1	6	0,5917	
HMC 2							
		BAIK	10	2	8	0,7219	0,0040
		PROBLEM	7	1	6	0,5917	
CONVEYOR							
		BAIK	15	3	12	0,7219	0,0353
		PROBLEM	2	0	2	0,0000	

Dari tabel 3 diperoleh nilai *gain* tertinggi adalah atribut Sandar Kapal dengan nilai 0.2666, jadi pada node 1.1 adalah atribut Sandar Kapal. Berikut adalah gambar pohon keputusan sampai dengan node terakhir :



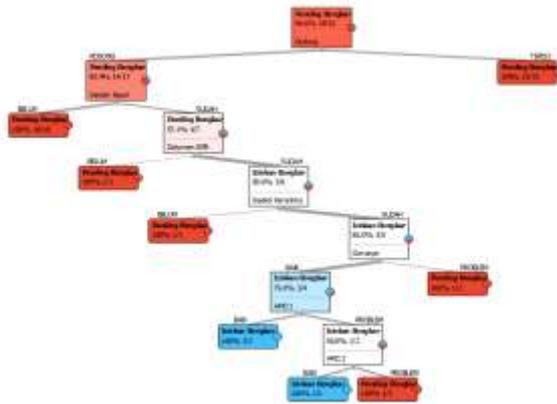
Gambar 2. Pohon Keputusan Penentuan Izin Pembongkaran Muatan Kapal

Dari pohon keputusan yang dihasilkan maka dapat diperoleh kesimpulan sebagai berikut :

- IF "Gudang" bernilai "Terisi" THEN "Pending Bongkar".
- IF "Gudang" bernilai "Kosong" AND "Sandar Kapal" bernilai "Belum" THEN "Pending Bongkar".
- IF "Gudang" bernilai "Kosong" AND "Sandar Kapal" bernilai "Sudah" AND "Dokumen SPB" bernilai "Belum" THEN "Pending Bongkar".
- IF "Gudang" bernilai "Kosong" AND "Sandar Kapal" bernilai "Sudah" AND "Dokumen SPB" bernilai "Sudah" AND "Inpeksi Karantina" bernilai "Belum" THEN "Pending Bongkar".
- IF "Gudang" bernilai "Kosong" AND "Sandar Kapal" bernilai "Sudah" AND "Dokumen SPB" bernilai "Sudah" AND "Inpeksi Karantina" bernilai "Sudah" AND "Conveyor" bernilai "Problem" THEN "Pending Bongkar".
- IF "Gudang" bernilai "Kosong" AND "Sandar Kapal" bernilai "Sudah" AND "Dokumen SPB" bernilai "Sudah" AND "Inpeksi Karantina" bernilai "Sudah" AND "Conveyor" bernilai "Baik" AND "HMC 1" bernilai "Baik" THEN "Izinkan Bongkar".
- IF "Gudang" bernilai "Kosong" AND "Sandar Kapal" bernilai "Sudah" AND "Dokumen SPB" bernilai "Sudah" AND "Inpeksi Karantina" bernilai "Sudah" AND "Conveyor" bernilai "Baik" AND "HMC 1" bernilai "Baik" THEN "Izinkan Bongkar".

- "Problem" AND "HMC 2" bernilai "Problem" THEN "Pending Bongkar".
- h. IF "Gudang" bernilai "Kosong" AND "Sandar Kapal" bernilai "Sudah" AND "Dokumen SPB" bernilai "Sudah" AND "Inpeksi Karantina" bernilai "Sudah" AND "Conveyor" bernilai "Baik" AND "HMC 1" bernilai "Problem" AND "HMC 2" bernilai "Baik" THEN "Izinkan Bongkar".

Data tersebut kemudian akan dilakukan pengujian menggunakan Orange Data Mining. Untuk hasil pengujiannya dapat dilihat pada gambar 3 :



Gambar 3. Pengujian Pohon Keputusan Menggunakan Aplikasi Orange

Confusion matrix dapat digunakan untuk melakukan perhitungan akurasi, presisi dan recall yang dihasilkan oleh aplikasi Orange Data Mining. Di bawah ini adalah hasil perhitungan *confusion matrix*, yang dapat dilihat pada gambar 4.

		Predicted		
		Izinkan Bongkar	Pending Bongkar	
Actual	Izinkan Bongkar	2	1	3
	Pending Bongkar	4	25	29
		6	26	32

Gambar 4. Perhitungan Dari *Confusion Matrix*

Dari hasil *confusion matrix* diatas didapatkan informasi bahwa :

1. Dari 6 jumlah data kapal dengan klasifikasi Izinkan Bongkar terdapat 2 data kapal dengan prediksi yang sesuai dengan aktual dan 4 data kapal yang diprediksi izinkan bongkar tetapi ternyata secara aktual pending bongkar.
2. Dari 26 jumlah data kapal dengan klasifikasi Pending Bongkar terdapat 25 data kapal dengan prediksi yang sesuai dengan aktual dan 1 data kapal yang diprediksi Pending bongkar ternyata secara aktual izinkan bongkar.

Berikut hasil dari validasi akurasi, presisi dan recall yang hasilnya dapat terlihat pada gambar 5.

Model	CA	F1	Precision	Recall
Tree - C4.5	0.844	0.886	0.903	0.844

Gambar 5. Hasil Validasi Akurasi, Presisi Dan Recall

KESIMPULAN

Setelah menganalisa dan menguji Algoritma C4.5 untuk mengambil keputusan apakah kapal akan dinyatakan "Izinkan Bongkar" atau "Pending Bongkar", sehingga dapat ditarik kesimpulan bahwa:

1. Hasil dari analisa data kapal, terdapat 7 atribut predictor yang akan digunakan yaitu Gudang, Sandar Kapal, Dokumen SPB, Inpeksi Karantina, Conveyor, HMC 1, HMC 2.
2. Algoritma C4.5 dapat diimplementasikan untuk menentukan keputusan apakah kapal tersebut dinyatakan "Izinkan Bongkar" atau "Pending Bongkar" dengan mengolah data yang digunakan sebagai data latih dalam pembentukan suatu pohon keputusan yang mana akan diperoleh aturan untuk memprediksi keputusan.
3. Berdasarkan data latih sebanyak 32 data kapal hasil evaluasi penelitian menggunakan aplikasi Orange Data Mining bahwa algoritma C4.5 mendapatkan nilai akurasi yaitu 84%, presisi 90% dan recall 84%.

UCAPAN TERIMAKASIH

Terimakasih banyak kepada pihak-pihak yang sudah memberikan dukungan kepada penulis baik dari segi moril ataupun materil.

DAFTAR PUSTAKA

- [1] Sulistiowati (2018). *Tata Kelola Pelayanan Indonesia Kerangka Pengaturan Dan Kelembagaan Kegiatan Pelayanan Di Indonesia*. 1st ed. Depok: Rajawali Pers, p.252.
- [2] Aditya, K., 2016. *Analisis Kinerja Pelayanan Bongkar Muat Pada Terminal Jamrud Berdasarkan Model Sistem Antrian (Studi Kasus Pada Terminal Jamrud Pt. Pelabuhan Indonesia Iii (Persero) Cabang Tanjung Perak)* (Doctoral dissertation, Universitas Brawijaya).
- [3] Buulolo, E., Silalahi, N. and Rahim, R., 2017. C4. 5 Algorithm To Predict the Impact of the Earthquake. *Entropy (S)*, 1, p.1.
- [4] Noviyanto, N., 2020. Penerapan Data Mining dalam Mengelompokkan Jumlah Kematian Penderita COVID-19 Berdasarkan

- Negara di Benua Asia. *Paradigma-Jurnal Komputer dan Informatika*, 22(2), pp.183-188.
- [5] Kacung, S. and Santoso, B., 2018. Sistem Deteksi Dini Untuk Meningkatkan Performance Kelulusan Mahasiswa Dengan ID3.
- [6] Tanjung, D.Y.H., 2021. Optimalisasi Algoritma C4. 5 untuk Prediksi Kerusakan Mesin ATM. *INFOSYS (INFORMATION SYSTEM) JOURNAL*, 6(1), pp.12-21.
- [7] Nasrullah, A.H., 2021. Implementasi Algoritma Decision Tree Untuk Klasifikasi Produk Laris. *Jurnal Ilmiah Ilmu Komputer Fakultas Ilmu Komputer Universitas Al Asyariah Mandar*, 7(2), pp.45-51

