

ANALISIS PERBANDINGAN DECISION TREE C4.5 DAN KNN DALAM PERIZINAN BONGKAR MUATAN KAPAL

Naurah Nazifah, Cahyo Prianto, Rolly Maulana Awangga

Program Studi Sarjana Terapan Teknik Informatika,
Universitas Logistik & Bisnis International, Jl Sariasih No.54, Kota Bandung, Indonesia
nnaurah998@gmail.com

ABSTRAK

Classification *Decision Tree* merupakan salah satu metode populer dalam analisis data dan pembelajaran mesin. Algoritma C4.5 adalah salah satu algoritma decision tree yang banyak digunakan karena kemampuannya dalam menghasilkan aturan keputusan yang dapat dipahami dengan mudah. Perizinan bongkar muatan kapal adalah proses krusial dalam operasi pelabuhan yang memastikan kapal dapat secara efisien dan aman melakukan bongkar muatan dalam upaya untuk meningkatkan efisiensi dan mengoptimalkan pengambilan keputusan perizinan. Penelitian ini bertujuan untuk melakukan analisis perbandingan metode machine learning antara algoritma *decision tree* C4.5 dengan algoritma *K-Nearest Neighbors* (KNN). Penulis sudah membandingkan kinerja algoritma-algoritma ini berdasarkan kriteria yang termasuk akurasi prediksi, dengan Classification *Decision Tree* menghasilkan peramalan unggul sebesar 98,33% dan 97,60% untuk algoritma KNN dalam investigasi ini. Hasil analisis bahwa pemilihan algoritma decision tree harus didasarkan pada tujuan spesifik analisis dan karakteristik data yang digunakan. Jika interpretabilitas aturan keputusan menjadi faktor utama, algoritma C4.5 tetap menjadi pilihan yang baik. Namun, jika akurasi prediksi dan penanganan data yang tidak seimbang menjadi prioritas, algoritma KNN dapat menjadi pilihan yang lebih baik.

Kata Kunci: Data Mining, Prediksi, Akurasi, C4.5, KNN

1. PENDAHULUAN

Analisis perbandingan algoritma Decision Tree merupakan topik yang penting dalam studi ilmu data dan kecerdasan buatan. Algoritma Decision Tree telah digunakan secara luas untuk memodelkan dan memprediksi data dalam berbagai bidang, termasuk pengambilan keputusan, klasifikasi, dan prediksi. Salah satu algoritma Decision Tree yang populer adalah algoritma C4.5. Algoritma ini merupakan pengembangan dari algoritma C4.5 adalah algoritma yang dikembangkan dari algoritma ID3 [1]. C4.5 memiliki beberapa perubahan dan peningkatan yang membuatnya lebih efektif dalam menangani masalah yang lebih kompleks. Dalam penelitian ini, peneliti bertujuan untuk melakukan analisis perbandingan antara algoritma Decision Tree C4.5 dengan algoritma K-Nearest Neighbors (KNN). Dengan *Decision Tree* dibentuk sebuah pohon keputusan dengan atribut data yang akan menghasilkan aturan-aturan dalam mengklasifikasi [2]. Dengan melakukan analisis perbandingan ini, diharapkan dapat menyediakan wawasan yang berharga bagi peneliti dan praktisi di bidang ilmu data dan kecerdasan buatan untuk memilih algoritma Decision Tree yang paling sesuai dengan kebutuhan mereka. Penelitian ini menggunakan beberapa pengujian seperti akurasi, dimana hal tersebut dapat menjadi nilai tambah [3].

Selanjutnya, peneliti akan menganalisis dan membandingkan hasil-hasil yang ditemukan dalam studi-studi tersebut. Mengeksplorasi kelebihan dan kelemahan algoritma Decision Tree C4.5 dibandingkan dengan algoritma KNN dalam hal akurasi prediksi dalam kemampuan mengatasi

masalah dengan atribut yang hilang atau kontinu. Algoritma machine learning bermanfaat dalam memberikan analisis yang benar dan membuat prediksi yang akurat [4].

Kesimpulan dari analisis perbandingan ini akan memberikan gambaran yang jelas tentang keunggulan dan batasan dari algoritma *Decision Tree* C4.5 dibandingkan dengan algoritma *K-Nearest Neighbors* (KNN). Hasil penelitian ini diharapkan dapat memberikan panduan yang bermanfaat bagi peneliti dan praktisi dalam memilih algoritma Decision Tree dan K-Nearest Neighbors (KNN) yang sesuai dengan kasus yang sedang mereka hadapi.

2. TINJAUAN PUSTAKA

2.1. Decision Tree (C4.5)

Sebagai pemisahan rekursif dari ruang sampel, Decision Tree adalah algoritma klasifikasi. Pohon keputusan terdiri dari simpul-simpul yang membentuk pohon berakar, yang menyiratkan bahwa simpul akar berfungsi sebagai arah pohon. Node internal atau uji adalah node yang memiliki tepi keluar. Daun merujuk ke semua node lainnya. Setiap node internal dalam pohon keputusan memisahkan ruang instance menjadi dua atau lebih subruang sesuai dengan fungsi diskrit tertentu dari nilai atribut [5].

Sebuah pohon keputusan terdiri dari node inti, yang memberikan pengujian pada faktor atau kualitas input tertentu yang memisahkan data menjadi kelompok yang lebih kecil, dan node daun, yang mengkategorikan setiap observasi dalam segmen yang dihasilkan. Dalam karya ini, C4.5 menggunakan

gagasan informasi entropi untuk membangun sebuah pohon keputusan. Entropi sampel S dari pengamatan yang diklasifikasikan diberikan oleh

$$\text{Entropy}(S) = -p_1 \log_2(p_1) - p_0 \log_2(p_0),$$

Dimana p_1 adalah proporsi dari kelas yang nilainya 1 dan p_0 proporsi dari kelas yang nilainya 0 dalam sampel S . C4.5 memeriksa informasi Gain yang dinormalisasi (perbedaan entropi) sebagai hasil pemilihan properti untuk membagi data. Atribut yang dipilih untuk pengambilan keputusan adalah yang memiliki informasi Gain ternormalisasi terbesar. Kemudian, algoritme dipecah menjadi bagian yang lebih mudah dikelola [5].

2.2. K-Nearest Neighbor(KNN)

Machine learning menggunakan teknik K-Nearest Neighbors (KNN) untuk melakukan klasifikasi atau regresi berdasarkan data terdekat.

Menurut Han, Kamber, dan Pei (2011): "KNN adalah algoritma pembelajaran mesin yang sederhana namun efektif untuk klasifikasi dan regresi. Algoritma ini mengklasifikasikan objek baru berdasarkan mayoritas kategori dari K tetangga terdekat dalam dataset pelatihan, atau memprediksi nilai berdasarkan rata-rata nilai K tetangga terdekat.

"KNN adalah teknik pembelajaran mesin non-parametrik yang digunakan untuk klasifikasi dan regresi", menurut Mitchell (1997). Objek baru dikategorikan dalam kategorisasi menurut kategori mayoritas dari K tetangga terdekatnya. Regresi mengandalkan rata-rata nilai K untuk meramalkan nilai item baru yang merupakan tetangga terdekat.

Menurut Hastie, Tibshirani, dan Friedman (2009): "KNN adalah metode non-parametrik yang digunakan untuk klasifikasi dan regresi. Algoritma ini mengklasifikasikan objek baru dengan cara mencari K tetangga terdekatnya dalam ruang fitur dan memilih kategori mayoritas dari tetangga-tetangga tersebut. Dalam regresi, algoritma ini memprediksi nilai objek baru berdasarkan rata-rata nilai K tetangga terdekatnya.

Secara umum, para ahli mendefinisikan algoritma KNN sebagai metode non-parametrik yang mengklasifikasikan objek baru atau memprediksi nilai berdasarkan mayoritas kategori atau rata-rata nilai K tetangga terdekat dalam dataset pelatihan. KNN adalah algoritma yang sederhana namun efektif dalam banyak kasus, dan dapat digunakan untuk berbagai tugas dalam pembelajaran mesin.

Aturan *K-Nearest Neighbor* dapat diposisi sebagai sampel pelatihan dan kelas mereka ketika memutuskan mengenai data masuk baru serta Tujuan algoritma ini adalah untuk mengklasifikasikan objek baru berdasarkan nilai atribut dan data latih (Okfalisa et al, 2017).

$$d = \sqrt{\sum_{i=1}^n (a_i - b_i)^2}$$

Keterangan :

d = jarak

a = data ui/testing

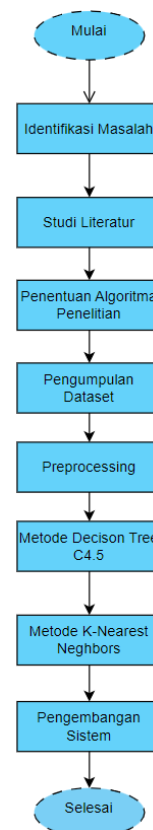
b = sample data

i = variable data

n = dimensi data

3. METODE PENELITIAN

3.1. Diagram Alur Metodologi Penelitian



Gambar 1. Alur Metodologi Penelitian

3.2. Identifikasi masalah

Pada tahapan awal dalam penelitian adalah identifikasi masalah, penulis melakukan analisa permasalahan yang terjadi ketika melakukan penelitian pada kegiatan perizinan bongkar muatan pada kapal. Permasalahan yang terjadi akan membuat suatu dasar pertimbangan dalam mempercepat rangkaian proses pengumpulan berkas persyaratan untuk melakukan kegiatan bongkar muatan pada kapal. Melihat dari permasalahan yang biasanya terjadi maka disarankan untuk adanya suatu sistem rekomendasi metode agar dapat memudahkan agent yang akan melakukan kegiatan bongkar muatan kapal di pelabuhan.

3.3. Studi Literature

Pada tahapan yang selanjutnya dilakukan persiapan data mentah, peneliti melakukan studi literature. Studi literature yang memuat sistematis mengenai kajian literature serta hasil penelitian yang sebelumnya dilakukan oleh peneliti lain yang

memiliki hubungan dengan penelitian yang akan peneliti lakukan. Studi literature sebelum melakukan penelitian memiliki tujuan yaitu untuk menentukan sebuah masalah agar dapat diteliti, melakukan pencarian yang relevan dengan masalah yang nantinya akan diteliti, melakukan pengkajian mengenai teori dasar yang relevan dengan masalah yang nantinya diteliti, melakukan pencarian mengenai landasan teori yang menjadi pedoman dalam pendekatan solusi masalah. Melakukan pengkajian terhadap hasil riset yang mempunyai kaitan dengan riset yang nantinya akan dilakukan, serta dapat melakukan penentuan atribut yang akan digunakan dalam menganalisa masalah.

3.4. Penentuan Algoritma Penelitian

Tahapan selanjutnya penentuan algoritma penelitian, pada penelitian mengenai rekomendasi perbandingan antara algoritma yang digunakan yaitu metode *Decision Tree C4.5* atau metode *K-Nearest Neighbor*.

3.5. Pengumpulan Dataset

Tahapan selanjutnya yaitu pengumpulan dataset, data yang dikumpulkan diambil dari situs kaggle.com dengan jumlah data yaitu 1500 data kaggle dan 300 data kegiatan bongkar muatan kapal pada pelabuhan. Data yang telah dikumpulkan berisi Cargo, weight(berat), BT (berthing time), ET (effective time), NOT (Not operation time), BWT (Berthing work time), LOA Vessel (Length of all), LOA USED.

3.6. Preprocessing

Tahapan preprocessing adalah suatu proses pengubahan bentuk data yang belum terstruktur menjadi data yang terstruktur yang dapat menyesuaikan kebutuhan dalam proses selanjutnya. Preprocessing dalam penelitian mempunyai tujuan yaitu data yang digunakan bisa diproses pada metode selanjutnya. Adapun tahapan preprocessing yaitu proses cleaning, proses cleaning adalah suatu proses yang dapat menghilangkan kata-kata yang missing. Proses preprocessing dilakukan supaya data dapat diproses oleh algoritma.

3.7. Metode Decision Tree

Metode yang digunakan pada penelitian ini adalah *Decision Tree C4.5* yang dipakai untuk mempermudah klasifikasi kapal apakah diizinkan melakukan kegiatan bongkar muatan kapal atau

pending bongkar muatan kapal dengan berdasarkan menurut atribut terkait. *Decision Tree* membangun pohon awal dengan menggunakan seluruh dataset pelatihan sebagai input kemudian setiap atribut dalam dataset akan diuji sebagai kandidat pemisah (splitter) untuk membuat node dalam pohon. Memilih atribut terbaik menggunakan metode Information Gain Ratio untuk memilih atribut terbaik sebagai pemisah (splitter), Information Gain Ratio mengukur seberapa baik atribut tersebut membagi dataset berdasarkan informasi yang diberikan, atribut dengan Information Gain Ratio tertinggi akan dipilih sebagai atribut terbaik untuk membuat node dalam pohon. Setelah atribut terbaik dipilih, sebuah node baru akan dibuat dalam pohon berdasarkan atribut tersebut, Nilai atribut tersebut akan digunakan sebagai aturan pemisah (split rule) untuk membagi dataset menjadi subset yang lebih kecil. Kemudian Setelah pohon selesai dibangun, nilai prediksi atau kelas target akan ditentukan untuk setiap daun dalam pohon, Nilai prediksi dapat berupa mayoritas kelas pada subset yang terkait dengan daun atau nilai regresi jika menggunakan regresi.

3.8. Metode K-Nearest Neighbor

Metode K-Nearest Neighbor dapat melakukan klasifikasi mengenai obyek dari data training yang mempunyai jarak terdekat dengan objek tersebut serta nilai K merupakan banyaknya nilai yang terdekat. KNN, jarak Euclidean sering digunakan untuk mengukur jarak antara dua titik dalam ruang fitur. Adapun cara kerja dari KNN yaitu melakukan pencarian kecocokan atau kemiripan dari kelompok objek data training dengan objek data baru atau data testing yang terdapat pada database.

3.9. Pengembangan Sistem

Tahap pengembangan sistem adalah ketika sistem yang dikembangkan dirancang dengan tujuan menghasilkan artikel ilmiah ini sehingga benar-benar menghasilkan sistem rekomendasi. Perancangan dan analisis sistem merupakan langkah-langkah dalam proses pengembangan sistem.

4. HASIL DAN PEMBAHASAN

Sebagai bahan referensi untuk penelitian ini, peneliti telah mengumpulkan beberapa penelitian yang memiliki relevansi seperti berikut :

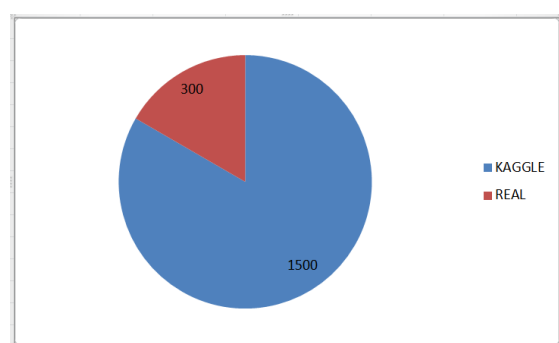
Tabel 1. Penelitian Terdahulu

No	Nama Peneliti	Tahun	Judul dan Kesimpulan Penelitian
1.	[6].	2018	Judul : “perbandingan minat siswa smu pada metode klasifikasi menggunakan 5 algoritma” Kesimpulan : “Teknik Decision Tree (C4.5) memiliki nilai presisi sebesar 89,09% dan nilai akurasi sebesar 84,00% lebih besar dibandingkan dengan metode lainnya”.
2.	[7].	2021	Judul : “Perbandingan metode algoritma k-nn & metode algoritma c4.5 pada analisa kredit macet (studi kasus pt tungtung textile bintang)” Kesimpulan : “Berdasarkan temuan penelitian “Perbandingan Metode Algoritma KNN dan Metode Algoritma C4.5 pada Analisis Kredit Macet (Studi Kasus Koperasi PT. Akurasi

No	Nama Peneliti	Tahun	Judul dan Kesimpulan Penelitian
			metode KNN adalah 45,21%, sedangkan akurasi algoritma C4.5 adalah 61,64”.
3.	[8].	2023	Judul : “Perbandingan Algoritma Naive Bayes Dan Decision Tree(C4.5) dalam Klasifikasi Dosen Berprestasi” Kesimpulan : “hasil dengan nilai akurasi sebesar 94,80%, nilai standar deviasi sebesar 3,79%, nilai presisi sebesar 95,25% untuk kelas prestasi, dan nilai presisi sebesar 94,15% untuk kelas underachievement”.
4.	[9].	2022	Judul : “analisis algoritma c4.5 dan fuzzy sugeno untuk optimasi rule base fuzzy” Kesimpulan : “Pendekatan fuzzy Sugeno yang digabungkan dengan algoritma C4.5 menghasilkan temuan perbandingan dengan akurasi sebesar 88,57%”.
5.	[10].	2019	Judul : “Perbandingan Algoritma Random Forest, Decision Stump, Naive Bayes, Bayesian Network dan Algoritma C4.5 Untuk Prediksi Pola Kartu Poker” Kesimpulan : “Pendekatan C4.5 berada di urutan kedua, diikuti oleh Random Forest, Decision Stump, Naive Bayes, dan Bayesian Network. Jika dibandingkan dengan pendekatan lain yang hanya membutuhkan waktu 0 hingga 0,08 detik, metode ini memakan waktu lama yaitu 2,26 detik”.
6.	[11].	2021	Judul : “perbandingan algoritma k-nearest neighbor, decision tree, dan naive bayes untuk menentukan kelayakan pemberian kredit” Kesimpulan : “Pada pengujian kelima, algoritma K-Nearest Neighbor (K-NN) mencapai skor akurasi tertinggi sebesar 93,33%”.
7.	[12].	2019	Judul : “Penerapan Algoritma C4.5 untuk Prediksi Kepuasan Penumpang Bus Rapid Transit (BRT) Trans Semarang” Kesimpulan : “perbandingan data pelatihan dan pengujian 60% : 40% nilai akurasi 92,5%, presisi 92,42%, dan recall 98,39%, pengujian 2 dari perbandingan 70% : 30% adalah nilai akurasi 93,33%”.
8.	[13].	2022	Judul : “Perbandingan algoritma c 4.5 dan algoritma naive bayes untuk klasifikasi pekerja migran Indonesia” Kesimpulan : “Metode Nave Bayes memiliki nilai akurasi sebesar 58,29% dan tingkat kesalahan sebesar 41,71%, sedangkan metode C 4.5 memiliki tingkat akurasi sebesar 84,84% dan tingkat kesalahan penyajian sebesar 15,16%”.
9.	[14].	2021	Judul : “Perbandingan Algoritma C4.5 Dan Id3 Untuk Prediksi Ketepatan Waktu Lulus Mahasiswa” Kesimpulan : “Dataset diuji menggunakan tiga komposisi berbeda, dan temuan menunjukkan bahwa metode ID3 dan algoritma C4.5 sebesar 30% dari data uji. Secara keseluruhan, berdasarkan temuan implementasi, Algoritma C4.5 mengungguli algoritma ID3 dalam hal akurasi”.
10.	[15].	2020	Judul : “Implementasi algoritma c4.5 untuk memprediksi keterlambatan pembayaran sumbangan pembangunan pendidikan sekolah menggunakan python” Kesimpulan : “Angka akurasi, presisi, dan daya ingat masing-masing adalah 73%, 71%, dan 71%, berdasarkan temuan dari deskripsi hasil dan hasil prediksi dengan Algoritma C4.5”.

4.1. Deskripsi Data

Perizinan bongkar muatan kapal merujuk pada informasi yang terkait dengan izin atau persetujuan yang diberikan untuk membongkar kapal yang mengandung bahan mutan. Bongkar mutan kapal mengacu pada proses pembongkaran kapal yang telah terpapar atau terkontaminasi oleh bahan atau zat yang berpotensi membahayakan manusia atau lingkungan. Deskripsi data tentang data perizinan bongkar mutan kapal dapat mencakup beberapa informasi penting, antara lain: Identitas Kapal (Informasi tentang kapal yang akan dibongkar mutan, seperti nama kapal, nomor registrasi, jenis kapal, ukuran, dan tahun pembuatan) dan juga dokumen perizinan (Informasi tentang dokumen-dokumen perizinan yang diperlukan untuk melaksanakan proses bongkar mutan kapal, seperti izin dari otoritas maritim atau lingkungan setempat, sertifikat keselamatan, dan persyaratan hukum yang harus dipenuhi).



Gambar 2. Diagram Deskripsi Data

Dalam penelitian ini penulis melakukan pengumpulan dataset, data yang dikumpulkan diambil dari situs kaggle.com dengan jumlah data yaitu 1500 data kaggle dan 300 data kegiatan bongkar muatan kapal pada pelabuhan. Data yang telah dikumpulkan berisi Cargo, weight (berat), BT (berthing time), ET (effective time), NOT (Not operation time),

BWT(Berthing work time), LOA Vessel (Lenght of all), LOA USED.

4.2. Analisis Pemodelan KNN

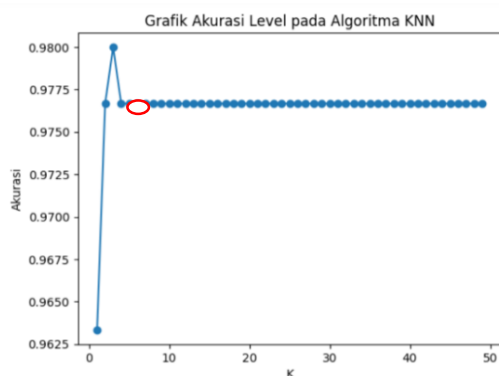
Unit data yang berbeda memiliki dampak yang signifikan terhadap temuan analisis saat menggunakan k-NN. Oleh karena itu, lebih baik untuk menormalkan atau mengubah data sebelum mengklasifikasikan menggunakan k-NN. Akurasi setiap k dihitung dengan menggunakan matriks konfusi untuk mendapatkan k yang ideal. Dimulai dengan k=1 sampai k=50 selanjutnya diukur akurasi untuk setiap k. Dari ke-50 k pilih akurasi optimum. K dengan akurasi optimum selanjutnya digunakan sebagai k optimum. Nilai-nilai akurasi untuk setiap k di jelaskan pada gambar berikut:

```
print('Akurasi Train set: ', accuracy_score(y_train, model_
print('Akurasi Test set: ', accuracy_score(y_test, y_pred))
```

Akurasi Test set: 0.9766666666666667

Tabel 2. Nilai Akurasi untuk K=1 s.d K=50

K1	96.33	K26	97.60
K2	97.60	K27	97.60
K3	98.00	K28	97.60
K4	97.60	K29	97.60
K5	97.60	K30	97.60
K6	97.60	K31	97.60
K7	97.60	K32	97.60
K8	97.60	K33	97.60
K9	97.60	K34	97.60
K10	97.60	K35	97.60
K11	97.60	K36	97.60
K12	97.60	K37	97.60
K13	97.60	K38	97.60
K14	97.60	K39	97.60
K15	97.60	K40	97.60
K16	97.60	K41	97.60
K17	97.60	K42	97.60
K18	97.60	K43	97.60
K19	97.60	K44	97.60
K20	97.60	K45	97.60
K21	97.60	K46	97.60
K22	97.60	K47	97.60
K23	97.60	K48	97.60
K24	97.60	K49	97.60
K25	97.60	K50	97.60



Gambar 3. Grafik Nilai Akurasi

Pada Gambar 3 dalam grafik nilai akurasi terlihat bahwa k yang optimum adalah k=4 dengan akurasi 97.60%.

4.3. Analisis Pemodelan Decision Tree

Pemodelan dengan data pelatihan 300 titik data adalah langkah pertama dalam teknik pemodelan pohon klasifikasi. Kinerja model pohon klasifikasi menggunakan matriks konfusi dan minsplit 300 pada data pelatihan ditunjukkan di bawah ini untuk memberikan hasil sebagai berikut.:

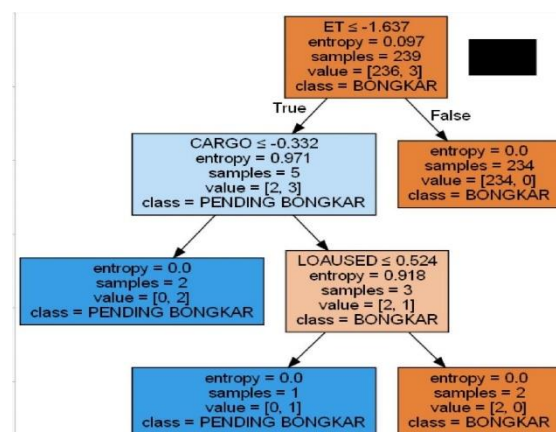
```
X_train, X_test, y_train, y_test = train_test_split(df,
print('Train set :', X_train.shape, y_train.shape)
print('Test set :', X_test.shape, y_test.shape)
```

```
Train set : (239, 8) (239,)
Test set : (60, 8) (60,)
```

```
print('Akurasi Decision Tree :', accuracy_score(y_test, y_pred))
```

Akurasi Decision Tree : 0.9833333333333333

Hasil entropi yang diperoleh dapat berkisar antara 0 hingga 1. Nilai entropi yang lebih rendah menunjukkan tingkat ketidakteraturan yang lebih rendah atau ketidakpastian yang lebih rendah dalam data, sementara nilai entropi yang lebih tinggi menunjukkan tingkat ketidakteraturan yang lebih tinggi atau ketidakpastian yang lebih tinggi dalam data. Berikut adalah hasil entropi dengan menggunakan metode pemisahan seperti Information Gain atau Gain Index :



Gambar 4. Hasil Pohon Keputusan

4.4. Pemilihan Metode Terbaik

Setelah penerapan teknik k-NN dan Decision Tree C4.5 pada data Izin Bongkar Muat Kapal, diberikan ringkasan akurasi dari kedua pendekatan tersebut.

Metode Klasifikasi	Akurasi
KNN	97.60%
Decision Tree	98.33%

Model yang menggunakan pendekatan *Decision Tree* C4.5 memiliki akurasi yang paling maksimal pada data training dan data testing, sesuai tabel. Oleh

karena itu, teknik Pohon Keputusan klasifikasi adalah metode yang paling efektif dalam penelitian ini untuk peramalan izin operasi bongkar muat kapal di Pelabuhan X yang mungkin termasuk bongkar atau pending bongkar.

5. KESIMPULAN

Berdasarkan hasil yang diperoleh, saat membandingkan nilai akurasi, pendekatan *Decision Tree* C4.5 atau dikenal dengan *Classification Decision Tree* memiliki nilai akurasi yang paling besar jika dibandingkan dengan metode K-Neighbor (KNN). Jadi teknik *Classification Decision Tree* menghasilkan peramalan unggul sebesar 98,33% untuk investigasi ini. menggunakan pendekatan Pohon Klasifikasi, Atribut-atribut yang digunakan dalam aturan di *Decision Tree* adalah Cargo, weight(berat muatan yang akan melakukan bongkar), BT(berthing time), ET (effective time), NOT (Not operation time), BWT(Berthing work time), LOA Vessel (Lenght of all), LOA USED (Lenght of All Used) selama 1 tahun.

DAFTAR PUSTAKA

- [1] N. H. Inda Arifin and W. J. Shudiq, "Algoritma Decision Tree Dengan Menggunakan Pruning dan Missing Value Untuk Prediksi Kredit Macet," *COREAI J. Kecerdasan Buatan, Komputasi dan Teknol. Inf.*, vol. 3, no. 1, pp. 38–45, 2022, doi: 10.33650/coreai.v3i1.4124.
- [2] P. Issn, A. Fatkhudin, M. Y. Febrianto, F. A. Artanto, M. W. N. Hadinata, and R. Fahlevi, "Mahasiswa Program Studi Manajemen Informatika Umpp," vol. 8, no. 2, pp. 83–86, 2022.
- [3] M. Ardiansyah, A. Sunyoto, and E. T. Luthfi, "Analisis Perbandingan Akurasi Algoritma Naïve Bayes Dan C4.5 untuk Klasifikasi Diabetes," *Edumatic J. Pendidik. Inform.*, vol. 5, no. 2, pp. 147–156, 2021, doi: 10.29408/edumatic.v5i2.3424.
- [4] K. L. Kohsasih and Z. Situmorang, "Analisis Perbandingan Algoritma C4.5 dan Naïve Bayes Dalam Memprediksi Penyakit Cerebrovascular," *J. Inform.*, vol. 9, no. 1, pp. 13–17, 2022, doi: 10.31294/inf.v9i1.11931.
- [5] R. Annisa, "Analisis Komparasi Algoritma Klasifikasi Data Mining Untuk Prediksi Penderita Penyakit Jantung," *J. Tek. Inform. Kaputama*, vol. 3, no. 1, pp. 22–28, 2019, [Online]. Available: <https://jurnal.kaputama.ac.id/index.php/JTIK/article/view/141/156>
- [6] A. Purwanto *et al.*, "Perbandingan Minat Siswa Smu Pada Metode Klasifikasi Menggunakan 5 Algoritma," vol. 2, no. 1, pp. 43–47, 2018.
- [7] A. Setianingrum, A. Hindayanti, D. M. Cahya, and D. S. Purnia, "Perbandingan Metode Algoritma K-NN & Metode Algoritma C45 Pada Analisa Kredit Macet (Studi Kasus PT Tugmung Textil Bintan)," *EVOLUSI J. Sains dan Manaj.*, vol. 9, no. 2, pp. 78–92, 2021, doi: 10.31294/evolusi.v9i2.11166.
- [8] A. Supriyadi, "Perbandingan Algoritma Naive Bayes dan Decision Tree (C4 . 5) dalam Klasifikasi Dosen Berprestasi," vol. 7, no. 1, pp. 39–49, 2023.
- [9] J. M. Jonson Manurung, B. S. Bosker Sinaga, P. M. H. Paska Marto Hasugian, L. Logaraj, and S. R. Sethu Ramen, "Analisis Algoritma C4.5 Dan Fuzzy Sugeno Untuk Optimasi Rule Base Fuzzy," *J. Sist. Inf. dan Ilmu Komput. Prima(JUSIKOM PRIMA)*, vol. 5, no. 2, pp. 166–171, 2022, doi: 10.34012/jurnalsisteminformasidanilmukomputer.v5i2.2488.
- [10] F. M. Januarsyah, E. Zuhairi, and R. M. Firsandaya, "Perbandingan Algoritma Random Forest, Decision Stump, Naïve Bayes, Bayesian Network dan Algoritma C4.5 Untuk Prediksi Pola Kartu Poker," *Pros. Annu. Res. Semin.*, vol. 5, no. 1, pp. 122–126, 2019, [Online]. Available: <https://archive.ics.uci.edu/ml/datasets/Poker+Hand>.
- [11] T. T. Muryono, A. Taufik, and I. Irwansyah, "Perbandingan Algoritma K-Nearest Neighbor, Decision Tree, Dan Naive Bayes Untuk Menentukan Kelayakan Pemberian Kredit," *Infotech J. Technol. Inf.*, vol. 7, no. 1, pp. 35–40, 2021, doi: 10.37365/jti.v7i1.104.
- [12] A. S. Febriarini and E. Z. Astuti, "Penerapan Algoritma C4.5 untuk Prediksi Kepuasan Penumpang Bus Rapid Transit (BRT) Trans Semarang," *Eksplora Inform.*, vol. 8, no. 2, pp. 95–103, 2019, doi: 10.30864/eksplora.v8i2.156.
- [13] S. Sukmawati, H. Februariyanti, A. Jananto, U. Stikubank Semarang Jl Tri Lomba Juang No, and S. Kode Pos, "I N F O R M a T I K a Perbandingan Algoritma C 4.5 Dan Algoritma Naive Bayes Untuk Klasifikasi Pekerja Migran Indonesia," *J. Inform. Manaj. dan Komput.*, vol. 14, no. 1, pp. 7–16, 2022.
- [14] T. Faizah and A. Jananto, "Perbandingan Algoritma C4 . 5 Dan Id3 Untuk Prediksi," *Jatisi*, vol. 8, no. 2, pp. 980–990, 2021.
- [15] V. S. Ginting, K. Kusriani, and E. Taufiq, "Implementasi Algoritma C4.5 untuk Memprediksi Keterlambatan Pembayaran Sumbangan Pembangunan Pendidikan Sekolah Menggunakan Python," *Inspir. J. Teknol. Inf. dan Komun.*, vol. 10, no. 1, pp. 36–44, 2020, doi: 10.35585/inspir.v10i1.2535.