

ANALISIS DATA MINING PADA PRODUKSI IKAN AIR TAWAR DI KECAMATAN PANJALU MENGGUNAKAN METODE C4.5

Apip Novareza, Ruli Herdiana, Iin

Teknik Informatika, STMIK IKMI Cirebon

Jl. Perjuangan No.10B, Karyamulya, Kec. Kesambi, Kota Cirebon, Jawa Barat 45135

Kapip0909@gmail.com

ABSTRAK

Produksi perikanan merupakan sektor penting dalam industri pangan dan ekonomi global juga ketahanan pakan. Untuk meningkatkan *efisiensi* dan pengelolaan produksi perikanan, analisis data mining telah menjadi alat yang berharga dalam mengidentifikasi pola-pola yang tersembunyi dalam data produksi perikanan. Salah satu metode yang umum digunakan dalam analisis data mining adalah metode C4.5, yang merupakan algoritma pohon keputusan untuk tugas klasifikasi. Penelitian ini bertujuan untuk menerapkan metode C4.5 dalam analisis data produksi perikanan dengan fokus pada klasifikasi hasil produksi perikanan. Metode C4.5 digunakan untuk mengidentifikasi hubungan antara atribut-atribut yang terkait dengan produksi perikanan dan hasil produksi yang terjadi. Langkah-langkah dalam penerapan metode C4.5 melibatkan pemilihan atribut yang paling informatif, pembentukan pohon keputusan, dan proses pruning untuk menghindari overfitting. Hasil eksperimen menunjukkan bahwa metode C4.5 mampu menghasilkan model klasifikasi yang cukup akurat dalam memprediksi hasil produksi perikanan. Selain itu, pohon keputusan yang dibuat memberikan wawasan yang bermanfaat tentang komponen yang mempengaruhi produksi perikanan. Hasil penelitian ini diharapkan dapat memberikan kontribusi penting dalam upaya peningkatan produksi perikanan di Kecamatan Cihaurbeuti dan dapat menjadi dasar untuk pengambilan keputusan yang lebih efektif dalam mengelola sektor perikanan di wilayah tersebut. Penelitian ini juga dapat menjadi contoh bagi penelitian serupa di daerah lain yang menghadapi tantangan serupa dalam mengoptimalkan hasil produksi perikanan. Seluruh rangkaian tahapan pada penelitian ini dapat diukur efektif pada penerapan algoritma C4.5 dalam klasifikasi potensi produksi perikanan di Kecamatan Panjalu. Hasil pengukuran akurasi data yang diperoleh terhadap data produksi perikanan nilai akurasinya mencapai 94.44%. Dan factor yang paling berpengaruh terhadap tingkat produksi perikanan adalah luas kolam.

Kata Kunci : C4.5, Produksi, Perikanan.

1. PENDAHULUAN

Kecamatan Kecamatan Panjalu adalah salah satu wilayah yang memiliki potensi besar dalam sektor perikanan di wilayah tersebut. Kecamatan ini terletak di daerah Ciamis Jawa Barat. Daerah ini kaya akan hasil produksi, salah satunya adalah di bidang perikanan. Masyarakat di sana banyak yang mempunyai kolam ikan baik untuk produksi maupun budidaya. Di tengah perkembangan industri perikanan yang semakin pesat, pengetahuan terhadap factor-faktor yang menentukan terhadap tingkat produksi menjadi krusial untuk meningkatkan produktivitas dan ekonomi masyarakat setempat.

Kondisi Masyarakat yang minim pengetahuan menjadi salah satu hambatan dalam produksi perikanan. Masyarakat atau pembudidaya ikan air tawar perlu adanya pengetahuan tentang factor apa saja yang mempengaruhi terhadap tingkat produksi perikanan. Menurut [1] dalam jurnalnya Hasil penelitian menunjukkan bahwa; (1) Uji keseluruhan (Uji-F) menghasilkan variabel-variabel bebas yang terdiri dari luas kolam (X1), jumlah benih (X2), jumlah pakan (X3), pengalaman usaha (X4), dan tenaga kerja (X5) secara serentak berpengaruh signifikan terhadap produksi usaha budidaya ikan. Oleh karena itu, diperlukan analisis yang mendalam untuk menentukan factor-faktor mana yang berpengaruh terhadap tingkat produksi di Kecamatan Panjalu. Salah satu metode

yang dapat digunakan untuk melakukan analisis ini adalah metode C4.5, yang merupakan salah satu metode pengambilan keputusan berbasis pohon keputusan (decision tree). Yang mana akan mengklasifikasikan hasil produksi perikanan di kecamatan panjalu. Metode C4.5 adalah salah satu teknik data mining yang efektif untuk mengekstraksi pola-pola penting dari data. Dalam konteks ini, metode C4.5 dapat digunakan untuk memahami faktor-faktor yang paling berpengaruh terhadap tingkat produksi ikan di Kecamatan Panjalu, sehingga dapat membantu mengklasifikasi tingkat produksi.

Dalam studi kasus sebelumnya yang di lakukan oleh [2] dengan judul Analisis Data mining dengan Metode C. 45 pada Klasifikasi Kenaikan Rata-Rata Volume Perikanan Tangkap. Hasil penelitian ini bertujuan untuk melihat kenaikan rata-rata hasil ikan tangkapan yang Dapat disebabkan oleh faktor Jumlah nelayan, Kondisi Cuaca, Kondisi Air, Lebar Jaring dan jumlah trip berlayar. Dalam penelitian yang saat ini di kerjakan akan lebih spesifik tentang implementasi Algoritma C4.5 dan melihat mengenai factor apa yang sangat berkontribusi terhadap produksi perikanan di Kecamatan Cihaurbeuti.

Dalam study sebelumnya yang di lakukan oleh [3] yang berjudul analisis kualitas air pada lokasi budidaya ikan air tawar di kecamatan suwawa Tengah, Dimana hasil penelitian ini bertujuan untuk melihat

kelayakan air untuk budidaya ikan air tawar. Dalam penelitian yang saat ini di kerjakan akan menghasilkan klasifikasi produksi perikanan dimana salah satu atribut yang mempengaruhi terhadap potensi perikanan itu sendiri adalah sumber air.

Dalam study sebelumnya [4] Menjelaskan tentang Ukuran luas perkebunan yang bervariasi mempengaruhi terhadap produksi yang dihasilkan. Penelitian ini menggunakan metode C4.5 Dimana C4.5 merupakan metode yang digunakan untuk konstruksi decision tree atau pohon keputusan dalam pengambilan keputusan. Dalam penelitian yang saat ini di kerjakan menggunakan metode yang serupa yaitu C4.5.

Penelitian ini bertujuan untuk menerapkan metode C4.5 pada klasifikasi tingkat produksi perikanan di Kecamatan Panjalu. Di harapkan dapat memudahkan Masyarakat mengetahui factor-faktor yang mempengaruhi terhadap tingkat produksi perikanan. Dengan memanfaatkan data mengenai komoditi ikan jumlah pakan, jumlah benih, luas kolam, sumber air dan produksi. Penelitian ini juga bertujuan untuk mengidentifikasi hubungan antara atribut dengan hasil produksi. Dengan demikian penelitian ini berpotensi memberikan wawasan terhadap Masyarakat ataupun pemerintah daerah setempat. Penelitian ini juga bertujuan untuk mengoptimalkan produksi perikanan di kecamatan panjalu.

Penelitian ini menggunakan metode C4.5 yang merupakan salah satu metode data mining untuk mengetahui akurasi prediksi tingkat produksi perikanan di kecamatan panajalu. C4.5 juga memberikan tingkat akurasi yang ideal. Dimana menurut [5] dalam jurnalnya yang berjudul Penerapan Algoritma C4. 5 Untuk Prediksi Minat Penjurusan Siswa Di Smkn 1 Kademangan menjelaskan bahwa Algoritma C4.5 ialah algoritma yang diterapkan untuk menyusun pohon keputusan (Decision Tree). Pohon keputusan yaitu metode klasifikasi dan prediksi yang familiar.

Dengan melakukan analisis tingkat produksi perikanan menggunakan metode C4.5, diharapkan dapat melihat factor-fakto yang mempengaruhi terhadap tingkat produksi perikanan di Kecamatan Panajalu. Pada gilirannya dapat mengarah pada peningkatan produksi perikanan dan kesejahteraan masyarakat setempat. Analisis ini juga dapat menjadi dasar bagi pembuat kebijakan dalam mengembangkan strategi pengelolaan perikanan yang berkelanjutan di wilayah tersebut. Melalui pendekatan analisis data mining, diharapkan penelitian ini dapat memberikan kontribusi yang signifikan dalam menjaga keseimbangan antara eksploitasi keberlanjutan ekosistem perikanan di Kecamatan Panjalu. Dengan memahami faktor-faktor yang memengaruhi tingkat produksi perikanan, diharapkan langkah-langkah strategis yang lebih efektif dan tepat dapat diambil untuk menjaga produksi perikanan yang berkelanjutan. Di harapkan memberikan manfaat jangka panjang bagi Masyarakat

2. TINJAUAN PUSTAKA

Hasil Hasil literature review yang telah dilakukan pada jurnal-jurnal penelitian terkait topik Analisis data menggunakan metode C4.5 dapat dijabarkan sebagai berikut:

Paper [3] Bapak Suswanto menghadapi beberapa tantangan dalam usahanya, di antaranya adalah variasi hasil penjualan yang tidak stabil dan persaingan ketat dengan pesaing di pasar tempe. Belum adanya sistem pengolahan data menyebabkan kesulitan dalam mengelola hasil produksi, sedangkan kurangnya pengetahuan dalam menangani masalah operasional dapat membahayakan daya saing. Selain itu, potensi kerugian yang tidak diinginkan juga menjadi ancaman. Kelebihan dari algoritma C4.5 mencakup kemampuannya untuk menghasilkan pohon keputusan yang mudah diinterpretasikan, memiliki tingkat akurasi yang dapat diterima, efisiensi dalam mengelola atribut bertipe diskret, serta kemampuannya untuk menangani baik atribut bertipe diskret maupun numerik.

Paper [4] menjelaskan tentang Analisis Data mining dengan Metode C. 45 pada Klasifikasi Kenaikan Rata-Rata Volume Perikanan Tangkap dimana penulis (Matondang et al., 2021b) dalam pengerjaannya menggunakan Data mining, Algoritma C4.5 dan Rapidminer yang mana dalam jurnalnya dapat di simpulkan bahwa Data yang digunakan berasal dari situs resmi BPS Simlaungun dan mencakup informasi mengenai hasil tangkapan ikan beserta faktornya. Setelah melakukan proses pembersihan data, data tersebut dianalisis dengan menggunakan algoritma C4.5 dan dilakukan pengujian dengan menggunakan perangkat lunak RapidMiner. Hasil dari RapidMiner menunjukkan hasil yang konsisten, yaitu atribut Jumlah nelayan memiliki tingkat akurasi sebesar 100%.

Paper [5] Menjelaskan tentang Ukuran luas perkebunan yang bervariasi mempengaruhi terhadap produksi yang dihasilkan. Penelitian ini menggunakan metode C4.5 Dimana C4.5 merupakan metode yang digunakan untuk konstruksi decision tree atau pohon keputusan dalam pengambilan keputusan.

Paper [6] Menjelaskan tentang Prediksi produksi karet dilakukan pada PT. Perkebunan Nusantara VII Trikora Lampung Selatan. Mengenai jumlah hasil produksi beserta luas areal perkebunan yang disetiap tahunnya mengalami fluktuatif. Kondisi fluktuatif tersebut disebabkan oleh luas areal perkebunan mengalami penambahan tetapi hasil produksi karet menurun. Prediksi dilakukan menggunakan teknik Data Mining dengan menerapkan algoritma Decision Tree C4.5.

Paper [7] Menjelaskan tentang Implementasi algoritma c. 45 untuk klasifikasi deteksi serangan pada protokol jaringan Dimana penulis [8] menjelaskan bahwa Berdasarkan hasil penelitian, data yang diambil pada tanggal 28 Agustus 2020 mencakup 125543 data, di mana terdapat 28 paket data yang terdeteksi sebagai worm trojan. Hasil analisis klasifikasi data jaringan

menggunakan algoritma C4.5 menunjukkan tingkat akurasi sebesar 93,62%. Penelitian ini fokus pada klasifikasi berdasarkan protokol jaringan yang digunakan, namun selanjutnya dapat melibatkan pengukuran parameter lain dalam jaringan dan penerapan algoritma data mining yang berbeda. Dimana Metode penelitian yang digunakan yaitu terdiri dari beberapa tahapan yaitu Pengambilan data, processing data, Analisis dan monitoring.

Paper [9] Menjelaskan tentang Implementasi Algoritma Decision Tree C4.5 Pada Klasifikasi Penyakit Tuberkulosis. Tujuan dari jurnal penelitian ini adalah untuk menggunakan algoritma Decision Tree C4.5 dalam membantu diagnosis penyakit tuberkulosis, dengan harapan dapat memberikan informasi yang cepat dan tepat kepada tenaga medis. Hasilnya menunjukkan bahwa implementasi algoritma C4.5 dalam klasifikasi penyakit tuberkulosis menggunakan data mining dapat memudahkan proses klasifikasi berdasarkan subtype yang ada dan meminimalisir kesalahan dalam proses diagnosa. Algoritma ini mampu mengklasifikasikan dan menganalisis data penyakit tuberkulosis dengan akurat, serta disajikan dalam beberapa tampilan aplikasi data mining untuk klasifikasi penyakit tuberkulosis.

Paper [10] menjelaskan tentang Analisis Klasifikasi C4.5 Pada Pola Pembayaran Sepeda Motor Adira Cabang Pematangsiantar. Tujuan dari jurnal penelitian ini adalah untuk menganalisis pola pembayaran sepeda motor di Cabang Adira Pematangsiantar dan mengidentifikasi masalah yang mungkin terjadi dalam pola pembayaran tersebut. Hasilnya menunjukkan bahwa algoritma C4.5 mampu memberikan solusi dengan tingkat akurasi sebesar 86,67% dalam menganalisis pola pembayaran sepeda motor.

Paper [11] menjelaskan tentang Klasifikasi Pohon Keputusan Dengan Algoritma C4.5 Untuk Kasus Pemilihan Produk. Tujuan dari jurnal penelitian ini adalah untuk menggunakan algoritma C4.5 dalam membangun pohon keputusan untuk merekomendasikan produk kecantikan berdasarkan kriteria seperti jenis kulit, masalah kulit, dan usia konsumen. Hasilnya menunjukkan bahwa algoritma C4.5 efektif dalam menghasilkan 4 produk yang direkomendasikan berdasarkan data latih sebanyak 206 baris. Selain itu, pengujian menunjukkan akurasi sebesar 96% dalam merekomendasikan produk perawatan kulit berdasarkan kondisi kulit pelanggan.

Paper [12] Berdasarkan penelitian yang dilakukan oleh Tukino pada tahun 2023 yang berjudul “Analisis Prediksi Penjualan Sayuran Hidroponik Menggunakan Algoritma C4.5 pada PT Batam Indo Agri Perkasa” PT Batam Indo Agri Perkasa mengalami kesulitan dalam memprediksi stok Sayuran Hidroponik yang akan habis karena banyaknya pembeli dan variasi selera konsumen. Oleh karena itu, perusahaan perlu mengembangkan sistem prediksi penjualan Sayuran Hidroponik yang dapat menjadi

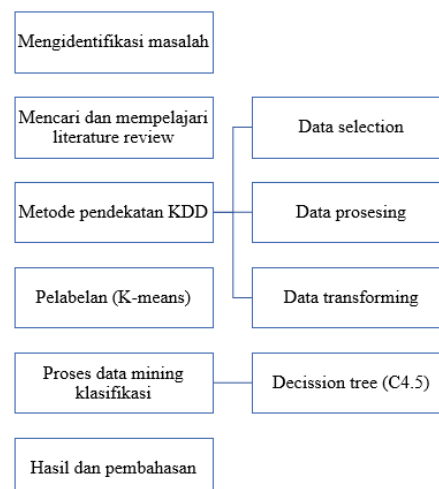
strategi yang sangat membantu dalam mencapai tujuan tersebut dengan meningkatkan stabilitas usaha dan keuntungan.

Paper [13] Berdasarkan penelitian yang dilakukan oleh Sanni Ucha Putri, Eka Irawan, Fitri Rizky pada tahun 2021 yang berjudul “Implementasi Data Mining Untuk Prediksi Penyakit Diabetes Dengan Algoritma C4.5” penelitian ini bertujuan untuk membuat model prediksi menggunakan Data Mining Algoritma C4.5 yang menghasilkan sebuah pohon keputusan serta pengujian yang dilakukan dengan menggunakan Rapidminer agar pencegahan terhadap penyakit diabetes. Menggunakan beberapa atribut klasifikasi yaitu berat badan, usia, tekanan darah, denyut nadi dan kadar gula darah. Hasil dari penelitian ini akan dijadikan sebagai acuan untuk dapat melihat apakah seseorang beresiko terkena penyakit diabetes atau tidak berdasarkan atribut yang telah ditetapkan.

Paper [14] Berdasarkan penelitian yang dilakukan oleh Chella Aprianti, Muhammad Faishal, Yuyun Umaidah pada tahun 2021 yang berjudul “Penerapan Data Mining Untuk Klasifikasi Penjualan Baju Muslim Dimasa Pandemi Covid-19 Menggunakan Metode Algoritma C4.5 (Studi Kasus: Garaya Collection)” Kebutuhan strategi khusus untuk meningkatkan penjualan, oleh karena itu penerapan klasifikasi menggunakan Metode data mining. Penerapan algoritma C4.5 di Garaya Collection untuk menentukan penjualan pakaian yang sedang laku Sangat baik, laris manis, dan tidak laris manis. Kemudian menggunakan C4.5 Metode algoritma pada rapidminer, dilakukan dengan memasukkan data pada Kategori, Merek Barang, Harga, Waktu Jual, dan Penjualan. dan akan diproses dan masuk bentuk pohon keputusan. Setelah itu, rapidminer akan menghasilkan mana yang terjual, terjual, dan tidak laku.

3. METODE PENELITIAN

3.1. Metode pendekatan KDD



Gambar 1. Tahapan Metode Penelitian

Dalam penelitian ini menggunakan Metode penelitian KDD (Knowledge Discovery in Databases) yang mana KDD merupakan suatu pendekatan sistematis untuk mengekstraksi pengetahuan yang berharga dari data. KDD merupakan suatu proses yang melibatkan beberapa langkah yang saling terkait.

Tahapan Metode Penelitian:

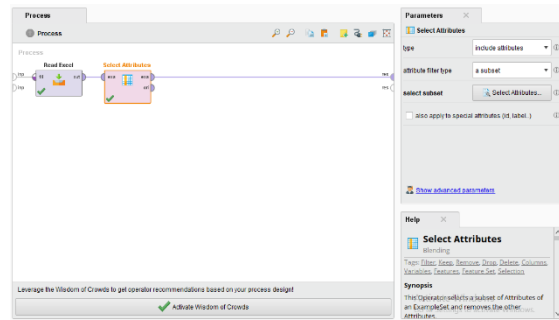
- Mengidentifikasi masalah: Mengidentifikasi Masalah yang ada di lingkungan. masalah yang di ambil mengenai potensi produksi di Kecamatan Panjalu
- Study Literatur: Literatur diambil dari berbagai sumber yaitu berupa artikel, jurnal ilmiah tentang metode KDD, Data Mining, C4.5, Decision tree, serta bahan bacaan lain yang mendukung
- Metode Pendekatan KDD: Melakukan Data Selection, Data Processing, Data Transformasi
- Pelabelan (K-means): Menentukan label menggunakan algoritma K-means
- Proses data mining klasifikasi: Mengklasifikasikan data menggunakan metode C4.5 dengan tools rapid miner
- Hasil dan Pembahasan: Pada tahapan ini akan dilakukan analisa data yang dihasilkan dari metode klasifikasi.

4. HASIL DAN PEMBAHASAN

Pada penelitian ini dataset tentang produksi perikanan di Kecamatan Panjalu terdapat 1000 data. Dalam data ini terdapat 19 atribut data yaitu nama, nik, Tanggal lahir, jenis kelamin, agama, Pendidikan terakhir, jumlah tanggungan, alamat, desa, kecamatan, kabupaten, provinsi, komoditi ikan, jenis ikan, luas kolam, jumlah pakan, jumlah benih, sumber air, dan produksi. Bisa di lihat pada gambar berikut.

Gambar 2. Dataset

Pada tahap ini di lakukan data selection. Tahapan ini bertujuan untuk memilih data yang akan di gunakan dalam proses klasifikasi melalui tools Rapidminer. Tahapan seleksi di lakukan di Rapid Miner menggunakan operator select attribute. Tahapan data selection adalah di mulai dengan memasukan operator read excel yang bertujuan untuk memuat data yang ada di excel, yang mana data awal terdapat 1000 data dan 19 atribut. Lanjut dengan memasukan operator select attribute seperti pada gambar di bawah ini.



Gambar 3. select attribute

Parameter set role yang di gunakan tampak pada tabel di bawah ini.

Tabel 1. Parameter dan Attribute yang di pilih pada Operator Select Attribute

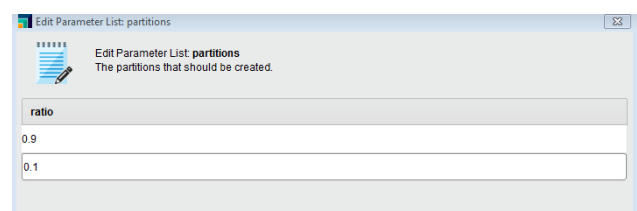
NO.	Parameter	Isi
1	Attribute nama	Produksi
2	Target role	Label

Hasil dari perubahan menjadi label dapat di lihat pada gambar berikut :

Row No.	TINGKAT PR...	KOMODITI IK...	SUMBER AIR	LUAS KOLAM	JUMLAH PA...	JUMLAH BE...	PRODUKSI
1	tinggi	Nila	Sumber mata...	140	320	1800	280
2	tinggi	Nila	Sungai	280	600	1800	560
3	rendah	Lele	Sungai	60	160	5000	120
4	tinggi	Nila	Sungai	140	320	6000	280
5	rendah	Mujair	Sungai	40	120	5600	80
6	rendah	Mujair	Sungai	50	140	1000	100
7	rendah	Mujair	Sungai	40	120	8000	80
8	rendah	Mujair	Sungai	40	120	960	80
9	tinggi	Nila	Sungai	140	320	1280	280
10	rendah	Nila	Sungai	70	180	5040	140
11	tinggi	Mujair	Sungai	140	320	5600	280
12	tinggi	Nila	Sungai	300	640	5600	600
13	rendah	Mujair	Sungai	40	120	8400	80
14	rendah	Mujair	Sungai	40	120	6000	80
15	rendah	Mujair	Sungai	70	180	3000	140
16	rendah	Mujair	Sungai	50	140	1400	100

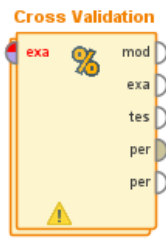
Gambar 4. Hasil Pemodelan

Dalam menentukan data training dan data testing di gunakan operator split data, dalam operator split data dibagi menjadi 90% untuk data training dan 10% di gunakan untuk data testing.

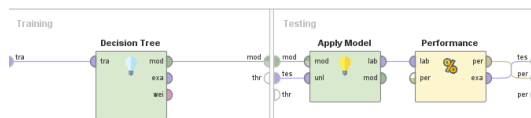


Gambar 5. Edit Parameter Split data

Untuk memperoleh hasil akurasi yang maksimal maka di gunakan operator cross validation seperti pada gambar berikut :



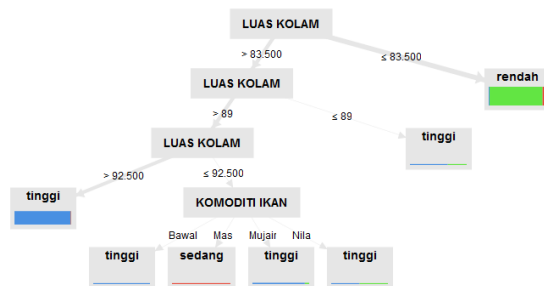
Gambar 6. Cross Validation



Gambar 7. Desain Model

Gambar di atas adalah langkah pemodelan decision tree. Dalam proses ini yang di gunakan adalah operator decision tree untuk menghasilkan pohon keputusan yang dapat di gunakan untuk klasifikasi. Dilanjutkan dengan memasukan operator apply model yang berguna untuk menerapkan model pada data set. Langkah terakhir adalah dengan memasukan operator performance (classification) yang berguna untuk evaluasi kinerja klasifikasi.

4.1. Decision tree



Gambar 8. Decission Tree

Berdasarkan gambar di atas dapat di deskripsikan:

- Jika luas kolam lebih besar ($>$) dari 83,5 yaitu luas kolam lebih besar dari ($>$) 89 dengan luas kolam lebih besar ($>$) dari 92,5 maka tingkat produksi masuk kategori tinggi.
- Jika luas kolam lebih besar ($>$) dari 83,5 dengan luas kolam lebih besar ($>$) dari 89 yaitu luas kolam lebih kecil sama dengan ($\leq$) 92,5 dengan komoditi ikan bawal maka tingkat produksi masuk kategori tinggi.
- Jika luas kolam lebih besar ($>$) dari 83,5 dengan luas kolam lebih besar ($>$) dari 89 yaitu luas kolam lebih kecil sama dengan ($\leq$) 92,5 dengan komoditi ikan mas maka potensi produksi masuk kategori sedang.
- Jika luas kolam lebih besar ($>$) dari 83,5 dengan luas kolam lebih besar ($>$) dari 89 yaitu luas kolam lebih kecil sama dengan ($\leq$) 92,5 dengan komoditi ikan mujair maka tingkat produksi masuk kategori tinggi.

- Jika luas kolam lebih besar ($>$) dari 83,5 dengan luas kolam lebih besar ($>$) dari 89 yaitu luas kolam lebih kecil sama dengan ($\leq$) 92,5 dengan komoditi ikan nila maka tingkat produksi masuk kategori tinggi.
- Jika luas kolam lebih besar ($>$) dari 85,5 yaitu luas kolam lebih kecil sama dengan ($\leq$) 89 maka tingkat produksi masuk ke dalam kategori tinggi
- Jika luas kolam lebih kecil sam dengan ($\leq$) 85,5 maka tingkat produksi rendah

4.2. Confusion matrix

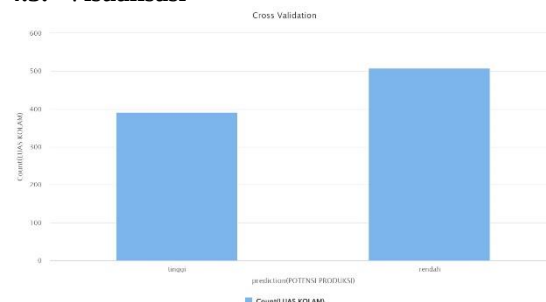
accuracy: 94.44% +/- 0.91% (micro average: 94.44%)

	true tinggi	true rendah	true sedang	class precision
pred tinggi	376	10	6	95.92%
pred rendah	8	474	26	93.31%
pred sedang	0	0	0	0.00%
class recall	97.92%	97.93%	0.00%	

Gambar 9. Accuracy Data Training

Hasil pengukuran akurasi data yang di peroleh dari data training nilainya 94,44%. Dari gambar tersebut diketahui prediksi tinggi dengan true tinggi mencapai 376, true rendah 10, dan true sedang 6 dengan hasil pencapaian nilai presisi sebesar 95.92%. Diketahui prediksi rendah dengan true tinggi mencapai 8, true rendah 474, dan true sedang mencapai 26 dengan hasil pencapaian nilai presisi sebesar 93.31%. Sedangkan prediksi sedang dengan true tinggi 0 true rendah 0 dan true sedang juga 0 dengan hasil pencapaian nilai presisi sebesar 0.00%. Untuk recall data yang sangat tinggi terdapat pada class recall true rendah mencapai 97.93%.

4.3. Visualisasi



Gambar 10. Visualisasi

Berdasarkan gambar visualisasi di atas tingkat produksi berdasarkan luas kolam dengan kategori tinggi memiliki rentang nilai mulai dari 0-392 sedangkan potensi produksi yang masuk kategori rendah memiliki rentang nilai mulai dari 0-509.

Pada penelitian ini dilakukan penerapan algoritma C4.5 untuk menganalisis faktor yang berpengaruh terhadap tingkat produksi perikanan di Kecamatan Panjalu. Metode analisis data yang digunakan yaitu KDD (Knowledge Discovery in Database) dan tools yang digunakan pada penelitian ini adalah RapidMiner. Tahapan analisis data yang dilakukan yaitu sebagai berikut:

Data selection dilakukan dengan pemilihan atribut yang sesuai untuk digunakan dalam penelitian dan semua atribut yang terdapat dalam data diambil sebanyak 6 atribut dari 19 atribut yang terdapat dalam data karena dirasa relevan dengan penelitian ini. Kemudian, tahapan preprocessing dilakukan dalam penelitian ini untuk mengecek data duplikat, kosong atau hilang. Namun pada tahap ini keseluruhan data tidak terdapat missing value ataupun duplikat. Selanjutnya tahap transformasi data, dalam penelitian ini dilakukan menggunakan operator set role untuk mengubah attribute nama menjadi id, dan menggunakan operator nominal to numerical untuk mengubah data nominal ke bentuk numeric. Sebelum masuk ke dalam tahap data mining di karenakan data belum mempunyai label di lakukanlah pelabelan data dengan K-medoids. Pada tahapan data mining diterapkan algoritma C4.5 untuk menganalisis faktor yang berpengaruh terhadap tingkat produksi perikanan di Kecamatan Panjalu. Kemudian melakukan tahap evaluasi dengan mengetahui nilai akurasi yang dihasilkan dari data tersebut. Seluruh rangkaian tahapan pada penelitian ini dapat diukur efektif pada penerapan algoritma C4.5 dalam klasifikasi potensi produksi perikanan di Kecamatan Panjalu.

Berdasarkan hasil evaluasi dan uji coba pada penelitian data mining prediksi potensi produksi perikanan di Kecamatan Panjalu menggunakan algoritma C4.5 bahwa tingkat produksi dapat diprediksi dengan nilai akurasi. Hasil pengukuran akurasi data yang diperoleh terhadap data produksi perikanan nilai akurasinya mencapai 94.44%. Tingkat akurasi sebesar 94.44% mencerminkan tingkat keandalan algoritma C4.5 dalam melakukan klasifikasi tingkat produksi perikanan. Keandalan ini menjadi faktor penting dalam memberikan keyakinan kepada pemangku kepentingan, termasuk masyarakat, pemerintah setempat, dan pelaku usaha perikanan, untuk mengandalkan hasil prediksi dalam pengambilan keputusan terkait produksi perikanan di Kecamatan Panjalu.

Berdasarkan gambar 4.18 di atas pada penelitian ini terdapat factor yang paling berpengaruh terhadap tingkat produksi di Kecamatan Panjalu yaitu luas kolam. Pada penelitian yang telah di lakukan sebelumnya hipotesis yang menyatakan bahwa variable pupuk, benih, luas lahan dan produktivitas rtp berpengaruh terhadap produksi perikanan budidaya di terima dengan tingkat signifikansi mencapai 99%. (Asmanah et al., 2009). Maka dari hasil penelitian inipun relevan dengan penelitian terdahulu bawasannya luas lahan berpengaruh terhadap hasil produksi. Dengan di ketahuinya factor yang paling berpengaruh terhadap tingkat produksi perikanan ini diharapkan dapat digunakan sebagai dasar dalam pengelolaan perikanan yang lebih efektif dan berkelanjutan.

5. KESIMPULAN DAN SARAN

Berdasarkan evaluasi model dan uji coba, dapat disimpulkan bahwa penerapan algoritma C4.5 dalam klasifikasi tingkat produksi perikanan di Kecamatan Panjalu cukup efektif.

Algoritma C4.5 dalam prediksi tingkat produksi perikanan di Kecamatan Panjalu menghasilkan tingkat akurasi sebesar 94.44%. Hal ini menandakan bahwa model yang dikembangkan mampu memberikan prediksi yang sangat akurat terkait tingkat produksi perikanan.

Hasil evaluasi model menunjukkan bahwa luas kolam menjadi faktor paling berpengaruh terhadap tingkat produksi perikanan di Kecamatan Panjalu. Temuan ini relevan dengan penelitian terdahulu.

Berdasarkan hasil penelitian yang menggunakan algoritma C4.5 untuk menganalisis tingkat produksi perikanan di Kecamatan Panjalu, disarankan untuk mengembangkan model dengan mempertimbangkan penambahan atribut, seperti kondisi lingkungan dan suhu air, guna meningkatkan akurasi prediksi. Perluasan cakupan data juga perlu diperhatikan untuk memperoleh informasi lebih komprehensif. Implementasi sistem pemantauan periodik terhadap faktor-faktor yang berpengaruh, seperti luas kolam, diharapkan dapat memastikan keberlanjutan dan aktualitas model. Melibatkan pihak terkait dan menyelenggarakan program edukasi kepada petani perikanan dapat mendukung penerapan hasil penelitian. Penggunaan algoritma alternatif, uji coba terhadap variabilitas waktu, dan keterlibatan komunitas lokal juga perlu dipertimbangkan untuk memperkuat validitas dan adopsi model. Pengembangan sistem pendukung keputusan dan penelitian lanjutan dapat memberikan dukungan lebih lanjut dalam pengelolaan potensi produksi perikanan di wilayah tersebut.

DAFTAR PUSTAKA

- [1] K. A. Pamungkas, D. P. Sudjatmiko, and A. Zaini, "Analisis Faktor-Faktor Yang Mempengaruhi Produksi Ikan Nila di Kecamatan Lingsar Kabupaten Lombok Barat," no. 1, 2023.
- [2] D. Yulia, A. P. Kusuma, and D. F. H. Permadi, "Penerapan Algoritma C4. 5 Untuk Prediksi Minat Penjurusan Siswa Di Smkn 1 Kademangan," *JATI (Jurnal Mhs. ...)*, 2022.
- [3] Y. Maulana, R. Winanjaya, and F. Rizk, "Klasifikasi Penjualan Tempe Dengan Data Mining Menggunakan Algoritma C4.5," *Bull. Comput. Sci. Res.*, vol. 2, no. 2, pp. 53–58, 2023.
- [4] M. R. Matondang, M. R. Lubis, and H. Satria, "Analisis Data mining dengan Metode C. 45 pada Klasifikasi Kenaikan Rata-Rata Volume Perikanan Tangkap," *J. Penerapan Kecerdasan Buatan*, vol. 2, no. 2, pp. 74–81, 2021.
- [5] D. Damayanti, "Implementasi Algoritma C4.5 Prediksi Produksi Komoditas Tanaman Perkebunan Berdasarkan Luas Lahan," *TIN Terap. Inform. Nusantara*, vol. 2, no. 10, pp. 571–

- 579, 2022, doi: 10.47065/tin.v2i10.1026.
- [6] N. P. Setyadi, "Penerapan Data Mining Untuk Prediksi Hasil Produksi Karet Menggunakan Algoritma Decision Tree C4. 5," *J. Teknol. Pint.*, vol. 2, no. 7, pp. 1–11, 2022, [Online]. Available: <http://teknologipintar.org/index.php/teknologipintar/article/view/203>
- [7] I. Anggraeni and S. Andriani, "Implementasi algoritma c. 45 untuk klasifikasi deteksi serangan pada protokol jaringan," *Komputasi J. Ilm. Ilmu ...*, 2021, [Online]. Available: <https://journal.unpak.ac.id/index.php/komputasi/article/view/3562>
- [8] I. Anggraeni and S. Andriani, "Implementasi algoritma c. 45 untuk klasifikasi deteksi serangan pada protokol jaringan," *J. Ilm. Ilmu Komput. dan Mat.*, vol. 18, no. 2, pp. 62–68, 2021.
- [9] S. A. Rahmadhani, M. S. Asih, and ..., "Implementasi Algoritma Dececion Tree C. 45 Pada Klasifikasi Penyakit Tuberkolosis. In Seminar Nasional Teknologi Informasi," *SEMINAR ...*, download.garuda.kemdikbud.go.id, 2020. [Online]. Available: <http://download.garuda.kemdikbud.go.id/article.php?article=2573771&val=24090&title=IMPLEMENTASI ALGORITMA DECISION TREE C45 PADA KLASIFIKASI PENYAKIT TUBERKULOSIS>
- [10] F. A. Pambudi, A. P. Windarto, M. Fauzan, J. T. Hardinata, and R. Winanjaya, "Analisis Klasifikasi C4. 5 Pada Pola Pembayaran Sepeda Motor Adira Cabang Pematangsiantar," *Kaji. Ilm. Inform. dan Komput.*, vol. 1, no. 4, pp. 160–169, 2021, [Online]. Available: <http://djournals.com/klik/article/view/144>
- [11] D. Kurniawati, E. Wahyuningsih, and D. P. Welfianti, "KLASIFIKASI POHON KEPUTUSAN DENGAN ALGORTIMA C45 UNTUK KASUS PEMILIHAN PRODUK," *J. Innov. Res. Knowl.*, vol. 3, no. 5, pp. 950–958, 2023, [Online]. Available: <https://bajangjournal.com/index.php/JIRK/article/view/6676>
- [12] Tukino, "Analisis Prediksi Penjualan Sayuran Hidroponik Menggunakan Algoritma C4.5 pada PT Batam Indo Agri Perkasa," *J. Desain Dan Anal. Teknol.*, vol. 2, no. 2, pp. 129–137, 2023, doi: 10.58520/jddat.v2i2.30.
- [13] S. U. Putri, E. Irawan, and F. Rizky, "Implementasi Data Mining Untuk Prediksi Penyakit Diabetes Dengan Algoritma C4.5," *Penerapan Sist. Inf. (Komputer Manajemen)*, vol. 2, no. 1, pp. 39–46, 2021.
- [14] C. Aprianti, M. Faishal, and Y. Umaidah, "Penerapan Data Mining Untuk Klasifikasi Penjualan Baju Muslim Dimasa Pandemi Covid-19 Menggunakan Metode Algoritma C4.5 (Studi Kasus: Garaya Collection)," *J. Ilm. Wahana Pendidik.*, vol. 8, no. 1, pp. 104–112, 2021, doi: 10.5281/zenodo.5816231.