

IMPLEMENTASI ALGORITMA C4.5 PADA PRODUKSI PERIKANAN DI KECAMATAN CIHAURBEUTI

Dira Nurrajiari¹, Bambang Irawan², Agus Bahtiar³, Edi Tohidi⁴

^{1,2,4} Teknik Informatika, STMIK IKMI Cirebon

³ Sistem Informasi, STMIK IKMI Cirebon

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

nengdir05@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 C.45, yang merupakan algoritma pohon keputusan untuk tugas klasifikasi. Penelitian ini bertujuan untuk menerapkan metode C.45 dalam analisis data produksi perikanan dengan fokus pada klasifikasi hasil produksi perikanan. Metode C.45 digunakan untuk mengidentifikasi hubungan antara atribut-atribut yang terkait dengan produksi perikanan dan hasil produksi yang terjadi. Langkah-langkah dalam penerapan metode C.45 melibatkan pemilihan atribut yang paling informatif, pembentukan pohon keputusan, dan proses pruning untuk menghindari overfitting. Hasil eksperimen menunjukkan bahwa metode C.45 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. Hasil pengukuran menyatakan Parameter *gain_ratio* adalah parameter yang paling mempengaruhi terhadap tingkat akurasi yang tinggi. Dalam penelitian ini tingkat akurasi hasil pengukuran akurasi data yang diperoleh terhadap data produksi perikanan nilai akurasinya mencapai 92,56%. Dan factor yang paling berpengaruh terhadap tingkat produksi perikanan adalah jumlah benih. Dengan mengetahui factor pengaruh utama yaitu jumlah benih ini memberikan wawasan berharga bagi pengambil keputusan, terutama bagi masyarakat pembudidaya dan pemerintah setempat.

Kata Kunci : C4.5, Produksi, Perikanan,

1. PENDAHULUAN

Kecamatan Cihaurbeuti terkenal dengan daerah yang hampir di setiap rumah mempunyai kolam ikan. Sebagian di jadikan sebagai budidaya dan sebagian di jadikan untuk konsumsi keluarga. Industri perikanan memiliki peran penting dalam mendukung pasokan pangan dan ekonomi suatu negara. Kecamatan Cihaurbeuti, sebagai salah satu wilayah dengan perikanan yang signifikan, berkontribusi terhadap kesejahteraan masyarakat setempat. Namun, dengan populasi yang terus meningkat dan masalah lingkungan yang kompleks, pendekatan yang cermat dalam pengelolaan produksi perikanan diperlukan untuk memastikan keberlanjutan hasil produksi yang optimal.

Dalam rangka mengoptimalkan pengelolaan produksi perikanan di Kecamatan Cihaurbeuti, diperlukan pendekatan analisis yang lebih mendalam untuk memahami faktor-faktor yang memengaruhi terhadap tingkat produksi perikanan. Dalam penelitian ini faktor” yang mempengaruhi terhadap tingkat produksi perikanan salah satunya yang akan di analisis adalah luas kolam, jumlah benih, jumlah pakan, komoditi ikan, sumber air, dan hasil produksi. Berdasarkan pengalaman empirik tidak hanya

komoditi ikan Luas kolam, jumlah benih, jumlah pakan, sumber air, dan hasil produksi saja , sangat memungkinkan ada factor lain yang mempengaruhi terhadap tingkat produksi. Salah satu pendekatan yang muncul dalam dekade terakhir adalah analisis data mining. Yang mana memanfaatkan teknik-teknik komputasional untuk mengidentifikasi pola dan informasi tersembunyi dalam data besar. Salah satu metode yang dapat digunakan dalam analisis data mining adalah metode C.45. Yang merupakan algoritma pohon keputusan yang telah terbukti efektif dalam tugas klasifikasi. Dengan menganalisis tingkat produksi perikanan menggunakan metode data mining, petani dan pengusaha perikanan dapat mengidentifikasi faktor-faktor yang berkontribusi pada peningkatan produksi. Mereka dapat menggunakan informasi ini untuk meningkatkan hasil produksi. Pada gilirannya dapat meningkatkan pendapatan dan kesejahteraan ekonomi mereka.

Dalam study kasus sebelumnya yang di lakukan oleh (Matondang et al., 2021a) 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 factor apa yang sangat berkontribusi terhadap tingkat produksi perikanan di Kecamatan Cihaurbeuti.

Dalam study sebelumnya 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 di gunakan itu sendiri adalah sumber air. (Koniyo, 2020)

Dalam penelitian sebelumnya yang di lakukan oleh (Asmanah et al., 2009) menjelaskan hipotesis yang menyatakan bahwa variable pupuk, benih, luas lahan, dan produktivitas rtp berpengaruh terhadap produksi perikanan budidaya di terima. Maka dari itu dalam penelitian yang di lakukan saat ini attribute yang akan di teliti meliputi komoditi ikan, luas kolam, jumlah benih, jumlah pakan, sumber air, dan hasil produksi.

Penelitian ini bertujuan untuk menerapkan metode C.45 dalam konteks untuk menganalisis Faktor yang sangat berkontribusi terhadap tingkat produksi perikanan di Kecamatan Cihaurbeuti. Dengan memanfaatkan data historis mengenai luas kolam, jumlah pakan, jumlah benih, komoditi ikan, hasil produksi dan sumber air. penelitian ini bertujuan untuk mengidentifikasi hubungan antara variabel-variabel ini dan hasil produksi perikanan. Dengan demikian, penelitian ini ber memberikan wawasan yang lebih mendalam bagi para pemangku kepentingan di Kecamatan Cihaurbeuti untuk mengambil keputusan yang lebih baik dalam pengelolaan perikanan. Yang di harapkan dapat meningkatkan pendapatan dan kesejahteraan ekonomi mereka.

Penelitian ini menggunakan metode C.45 yang merupakan salah satu metode data mining untuk mengetahui tingkat akurasi prediksi produksi perikanan di kecamatan Cihaurbeuti. C.45 juga memberikan tingkat akurasi yang ideal. dimana menurut (Yulia et al., 2022) dalam jurnalnya menjelaskan bahwa Algoritma C4.5 ialah algoritma yang diterapkan untuk menyusun pohon keputusan (Decision Tree). Pohon keputusan yaitu metode klasifikasi dan prediksi yang familiar.

Keterbatasan warga dalam pengetahuan tentang budidaya ikan menjadi salah satu factor tidak terkondisikannya produksi disana. Melalui pendekatan analisis data mining, Penelitian ini diharapkan dapat memberikan kontribusi yang signifikan dalam menjaga keseimbangan produksi perikanan di Kecamatan Cihaurbeuti. 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 setempat serta lingkungan. Pada gilirannya dapat mengarah pada peningkatan produksi sektor perikanan dan kesejahteraan masyarakat setempat. Analisis ini juga dapat menjadi dasar bagi pembuat kebijakan dalam mengembangkan strategi pengelolaan perikanan yang berkelanjutan di wilayah tersebut.

2. TINJAUAN PUSTAKA

Hasil literature review yang telah dilakukan pada jurnal-jurnal penelitian terkait topik Analisis data dapat dijabarkan sebagai berikut :

Menurut [1] 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, 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.

Menurut [2] 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 C.45 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%.

Menurut [3] Menjelaskan tentang Ukuran luas perkebunan yang bervariasi mempengaruhi terhadap produksi yang dihasilkan. Penelitian ini menggunakan metode C.45 Dimana C4.5 merupakan metode yang digunakan untuk konstruksi decision tree atau pohon keputusan dalam pengambilan keputusan.

Menurut [4] 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 pertambahan tetapi hasil produksi karet menurun. Prediksi dilakukan menggunakan teknik Data Mining dengan menerapkan algoritma Decision Tree C4.5.

Menurut [5] Menjelaskan tentang Implementasi algoritma c. 45 untuk klasifikasi deteksi serangan pada protokol jaringan Dimana penulis [5] 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 C.45 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.

Menurut [6] Menjelaskan tentang Implementasi Algoritma Decision Tree C.45 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.

Menurut [7] 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.

Menurut [8] menejelaskan tentang Klasifikasi Pohon Keputusan Dengan Algoritma C45 Untuk Kasus Pemilihan Produk. Tujuan dari jurnal penelitian ini adalah untuk menggunakan algoritma C45 dalam membangun pohon keputusan untuk merekomendasikan produk kecantikan berdasarkan kriteria seperti jenis kulit, masalah kulit, dan usia konsumen. Hasilnya menunjukkan bahwa algoritma C45 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.

Berdasarkan penelitian yang dilakukan oleh [9] 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.

Berdasarkan penelitian yang dilakukan oleh [9] 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.

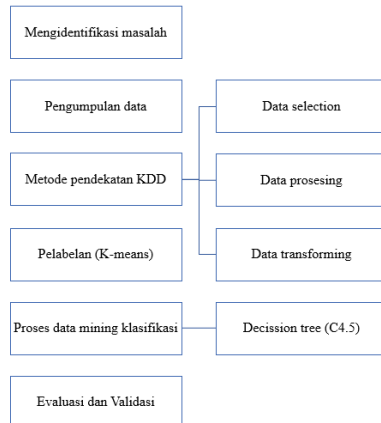
Berdasarkan penelitian yang dilakukan oleh [10] pada tahun 2021 yang berjudul “Implementasi Data Mining Untuk Prediksi Penyakit Diabetes Dengan Algoritma C4.5” penelitian ini bertujuan untuk membuat model prediksi menggunakan Data Mining Algoritma C4.5 yang menghasilkan sebuah pohon keputusan serta pengujian 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.

Berdasarkan penelitian yang dilakukan oleh [11] 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

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.

Bagan Tahapan Metode Penelitian



Gambar 1. 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 Mininng, C4.5, Decission tree, serta bahan bacaan lain yang mendukung
- Metode Pendekatan KDD : Melakukan Data Selection, Data Prossesing, 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

4.1. Proses klasifikasi

Import Data - Format your columns.

☐ Replace errors with missing values ⓘ

	Komoditi ik...	Luas kolam	Jumlah pa...	Jumlah Be...	Sumber air	produksi
	polynomial	integer	integer	integer	polynomial	integer
1	Nila	56	152	1120	Ingasi	112
2	Lele	140	320	2800	sungai	280
3	Mujair	140	320	2240	Sungai	280
4	Nila	140	320	2240	Sungai	280
5	Nila	140	320	2240	Sungai	280
6	Nila	56	152	896	Sumber mata air	112
7	Nila	140	320	2240	Sungai	280
8	Nila	70	180	1120	Sungai	140
9	Nila	308	656	4928	Sungai	616
10	Nila	140	320	2240	Sungai	280
11	Nila	154	348	2464	Sungai	308
12	Nila	140	320	2240	Sungai	280

no problems.

Previous Finish Cancel

Gambar 2. Statistik data

Langkah awal pembuatan pohon keputusan (*decision tree*) di mulai dengan memasukan data yang akan di ke dalam rapid miner menggunakan operator *read excel* yang mana operator ini dapat di gunakan untuk memuat data dari Microsoft excel.

Dalam memilih label maka di masukan operator *set role* yang bertujuan merubah peran atribut. Bisa di lihat pada gambar di bawah ini :



Gambar 3. Operator *set role*

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

Tabel 1. Parameter yang di atur pada *set role*

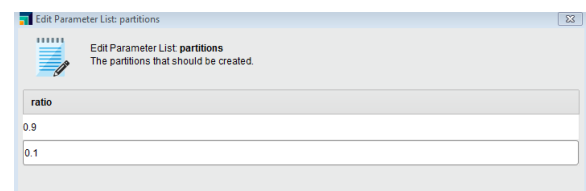
NO.	Parameter	Isi
1	Attribute nama	Produksi
2	Target role	Label

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

Row No.	Tingkat prod...	Komoditi ikan	Luas kolam	Jumlah pakan	Jumlah Benih	Sumber air	produksi
1	rendah	Nila	56	152	1120	Ingasi	112
2	sedang	Lele	140	320	2800	sungai	280
3	sedang	Mujair	140	320	2240	sungai	280
4	sedang	Nila	140	320	2240	sungai	280
5	sedang	Nila	140	320	2240	sungai	280
6	rendah	Nila	56	152	896	Sumber mata...	112
7	sedang	Nila	140	320	2240	sungai	280
8	rendah	Nila	70	180	1120	sungai	140
9	tinggi	Nila	308	656	4928	sungai	616
10	sedang	Nila	140	320	2240	sungai	280
11	sedang	Nila	154	348	2464	sungai	308
12	sedang	Nila	140	320	2240	sungai	280
13	sedang	Nila	196	432	3136	sungai	392
14	rendah	Nila	70	180	1120	sungai	140
15	sedang	Nila	154	348	3080	sungai	308
16	rendah	Nila	84	208	1344	sungai	168

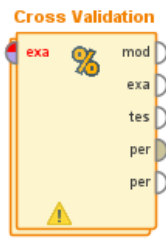
Gambar 4. Statistik data

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.

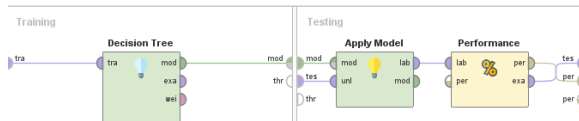


Gambar 5. Ratio yang di atur pada *split data*

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



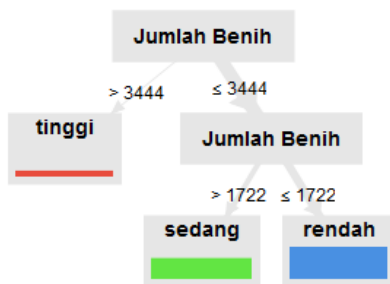
Gambar 6. Cross validation



Gambar 7. Pemodelan decision tree

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.

4.2. Decision tree



Gambar 8. Pohon keputusan

Setelah uji coba pada penelitian data mining prediksi produksi perikanan di Kecamatan Cihaurbeuti menggunakan algoritma C4.5 maka tampilan pohon keputusan dapat di lihat pada gambar 8 pohon keputusan. Berikut ini rules yang di hasilkan dari pohon keputusan

NO	RULES	KEPUTUSAN
1	Jika jumlah benih lebih dari ($>$) 3444 maka	Tinggi
2	jika jumlah benih kurang dari sama dengan ($\leq$) 3444 yaitu jumlah benih lebih besar dari ($>$) 1722 maka	Sedang
3	jika jumlah benih lebih kecil sama dengan ($\leq$) dari 1722 maka	Rendah

Gambar 9. Rules pohon keputusan

4.3. Confusion matrix

Pada parameter decision tree ada beberapa kinerja yang dapat mempengaruhi hasil atau nilai akurasi dari prediksi perikanan di antaranya gain ratio, information gain, gini index, dan accuracy. Berikut

Hasil akurasi dari criteria pada parameter decision tree:

4.4. Accuracy menggunakan parameter Gain_ratio

accuracy: 92.56% +/- 2.67% (micro average: 92.56%)

	true rendah	true sedang	true tinggi	class precision
pred. rendah	505	0	0	100.00%
pred. sedang	0	321	67	82.73%
pred. tinggi	0	0	7	100.00%
class recall	100.00%	100.00%	9.46%	

Gambar 10. Confusion matrix Gain_Ratio

Hasil pengukuran akurasi data menggunakan parameter gain_ratio yang di peroleh dari data training nilainya 92,56%. Dari gambar tersebut diketahui prediksi rendah dengan true rendah mencapai 505 dengan hasil pencapaian nilai presisi sebesar 100.00%. Diketahui prediksi sedang dengan true sedang mencapai 321 dan true tinggi mencapai 67 dengan hasil pencapaian nilai presisi sebesar 82,73%. Sedangkan prediksi tinggi dengan true tinggi mencapai 7 dengan hasil pencapaian nilai presisi sebesar 100.00%. Untuk recall data yang sangat tinggi terdapat pada class recall true rendah dan true sedang mencapai 100.00% sedangkan untuk class recall tinggi mencapai 9.46%.

4.5. Accuracy menggunakan parameter Information_gain

accuracy: 91.78% +/- 0.57% (micro average: 91.78%)

	true rendah	true sedang	true tinggi	class precision
pred. rendah	505	0	0	100.00%
pred. sedang	0	321	74	81.27%
pred. tinggi	0	0	0	0.00%
class recall	100.00%	100.00%	0.00%	

Gambar 11 confusion matrix Informstion_Gain

Hasil pengukuran accuracy data menggunakan parameter information_gain yang di peroleh dari data training nilainya 91,87%. Dari gambar tersebut diketahui prediksi rendah dengan true rendah mencapai 505, true sedang 0 dan true tinggi 0 dengan hasil pencapaian nilai presisi sebesar 100.00%. Diketahui prediksi sedang dengan true rendah 0 true sedang mencapai 321 dan true tinggi mencapai 74 dengan hasil pencapaian nilai presisi sebesar 81,27%. Sedangkan prediksi tinggi dengan true rendah 0, true sedang 0 dan true tinggi 0 dengan hasil pencapaian nilai presisi sebesar 0.00%. Untuk recall data yang sangat tinggi terdapat pada class recall true rendah dan true sedang mencapai 100.00% sedangkan untuk class recall tinggi 0,00%

4.6. Accuracy menggunakan parameter Gini_Index

accuracy: 91,78% +/- 0,57% (micro average: 91,78%)				
	true rendah	true sedang	true tinggi	class precision
pred rendah	505	0	0	100,00%
pred sedang	0	321	74	81,27%
pred tinggi	0	0	0	0,00%
class recall	100,00%	100,00%	0,00%	

Gambar 12 Confusion matrix Gini_Index

Hasil pengukuran accuracy data menggunakan parameter Gini_Index yang di peroleh dari data training nilainya 91,87%. Dari gambar tersebut diketahui prediksi rendah dengan true rendah mencapai 505, true sedang 0 dan true tinggi 0 dengan hasil pencapaian nilai presisi sebesar 100,00%. Diketahui prediksi sedang dengan true rendah 0 true sedang mencapai 321 dan true tinggi mencapai 74 dengan hasil pencapaian nilai presisi sebesar 81,27%. Sedangkan prediksi tinggi dengan true rendah 0, true sedang 0 dan true tinggi 0 dengan hasil pencapaian nilai presisi sebesar 0,00%. Untuk recall data yang sangat tinggi terdapat pada class recall true rendah dan true sedang mencapai 100,00% sedangkan untuk class recall tinggi 0,00%

4.7. Accuracy menggunakan parameter Accuracy

accuracy: 91,78% +/- 0,57% (micro average: 91,78%)				
	true rendah	true sedang	true tinggi	class precision
pred rendah	505	0	0	100,00%
pred sedang	0	321	74	81,27%
pred tinggi	0	0	0	0,00%
class recall	100,00%	100,00%	0,00%	

Gambar 13. Confusion matrix Accuracy

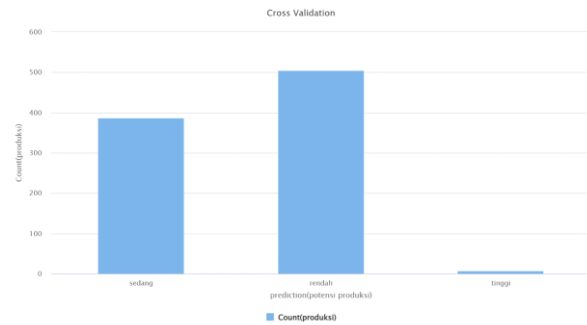
Hasil pengukuran accuracy data menggunakan parameter Accuracy yang di peroleh dari data training nilainya 91,87%. Dari gambar tersebut diketahui prediksi rendah dengan true rendah mencapai 505, true sedang 0 dan true tinggi 0 dengan hasil pencapaian nilai presisi sebesar 100,00%. Diketahui prediksi sedang dengan true rendah 0 true sedang mencapai 321 dan true tinggi mencapai 74 dengan hasil pencapaian nilai presisi sebesar 81,27%. Sedangkan prediksi tinggi dengan true rendah 0, true sedang 0 dan true tinggi 0 dengan hasil pencapaian nilai presisi sebesar 0,00%. Untuk recall data yang sangat tinggi terdapat pada class recall true rendah dan true sedang mencapai 100,00% sedangkan untuk class recall tinggi 0,00%.

Tabel 2. Akurasi berdasarkan criterion parameters

Criteria	Accuracy
Gain_Ratio	92,56 %
Information_Gain	91,78%
Gini_Index	91,78%
Accuracy	91,78%

Dari tabel di atas di ketahui bahwa parameter yang berpengaruh terhadap nilai akurasi terbaik adalah Gain_Ratio dengan akurasi sebesar 92,56%.

4.8. Visualisasi



Gambar 14. Visualisasi

Berdasarkan gambar visualisasi di atas produksi berdasarkan Jumlah benih dengan kategori sedang memiliki rentang nilai 0-388, Sedangkan produksi yang masuk ke dalam kategori rendah memiliki rentang nilai mulai dari 0-505 dan produksi dengan kategori tinggi memiliki rentang nilai mulai dari 0-7.

Berdasarkan hasil evaluasi dan uji coba pada penelitian data mining prediksi produksi perikanan di Kecamatan Cihaurbeuti 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 92,56%.

Berdasarkan hasil uji coba pada penelitian ini terdapat factor yang paling berpengaruh terhadap produksi di Kecamatan Cihaurbeuti yaitu jumlah benih. 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. Maka dari hasil penelitian ini relevan dengan penelitian terdahulu bawasannya jumlah benih berpengaruh terhadap hasil produksi. Dengan di ketahuinya factor yang paling berpengaruh ini menjadi wawasan baru bagi pengambil keputusan yaitu masyarakat dan pemerintah terkait. Dimana masyarakat pembudidaya dapat memahami pentingnya pengelolaan jumlah benih dalam kegiatan perikanan. Selain itu, pemerintah setempat dapat fokus pada penyuluhan kepada masyarakat terkait peran jumlah benih dalam meningkatkan hasil produksi perikanan. Hal ini diharapkan dapat meningkatkan efektivitas dan keberlanjutan pengelolaan perikanan di Kecamatan Cihaurbeuti.

Berdasarkan hasil uji coba menggunakan 4 criterion parameter yang tersedia yaitu gain_ratio, information_gain, gini_index dan Accuracy diketahui bahwa parameter gain_ratio menjadi parameter yang mempengaruhi nilai akurasi terbaik mencapai 92,56%

5. KESIMPULAN DAN SARAN

Parameter gain_ratio memiliki pengaruh terbesar terhadap nilai akurasi model, mencapai 92,56%. Oleh karena itu, dalam mengoptimalkan model prediksi, perhatian khusus harus diberikan pada parameter ini. Dengan mengetahui factor pengaruh utama yaitu

jumlah benih ini memberikan wawasan berharga bagi pengambil keputusan, terutama bagi masyarakat pembudidaya dan pemerintah setempat. Untuk meningkatkan efektivitas produksi perikanan di Kecamatan Cihaurbeuti, disarankan agar masyarakat pembudidaya lebih memfokuskan pengelolaan pada faktor kunci, terutama jumlah benih, yang telah terbukti menjadi variabel paling berpengaruh. Program penyuluhan perlu ditingkatkan untuk meningkatkan pemahaman masyarakat tentang pentingnya jumlah benih dalam produksi perikanan. Pemerintah setempat dapat merancang kebijakan berbasis bukti dengan melibatkan kolaborasi antara lembaga penelitian, pemerintah, dan komunitas pembudidaya. Pengembangan model prediksi harus terus berlanjut dengan mempertimbangkan faktor tambahan dan optimalisasi parameter lainnya. Monitoring dan evaluasi terus-menerus perlu diimplementasikan untuk memastikan pembaruan model berdasarkan data produksi perikanan yang aktual. Selain itu, pelatihan terkait penggunaan model prediksi perlu diberikan kepada pengambil keputusan dan praktisi di bidang perikanan. Dengan demikian, upaya kolaboratif ini diharapkan dapat menciptakan pengelolaan perikanan yang lebih efektif dan berkelanjutan di Kecamatan Cihaurbeuti.

DAFTAR PUSTAKA

- [1] 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.
- [2] 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.
- [3] 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.
- [4] 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>
- [5] 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.
- [6] S. A. Rahmadhan, M. S. Asih, and N. Wulan, "Implementasi Algoritma Decision Tree C.45 Pada Klasifikasi Penyakit Tuberkulosis," *Semin. Nas. Teknol. Inf. Komun.*, vol. 4, no. 2, pp. 230–238, 2020.
- [7] 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>
- [8] "Klasifikasi pohon keputusan dengan algoritma C4.5 untuk kasus pemilihan produk," *J. Innov. Res. Knowl.*, vol. 3, no. 5, pp. 955–964, 2023.
- [9] 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.
- [10] 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.