
KLASIFIKASI KABUPATEN KOTA PROVINSI JAWA BARAT BERDASARKAN PENDAPATAN DARI SEKTOR PERTANIAN DENGAN ALGORITMA DECISION TREE

AMRIL MUTOI SIREGAR
AHMAD FAUZI

Jurusan Teknik informatika,
Universitas Buana Perjuangan karawang
Jalan Ronggo Waluyo Sirnabaya, Puseurjaya, Kec. Telukjambe Timur,
Kabupaten Karawang, Jawa Barat 41361
Email: amrilmutoi@ubpkarawang.ac.id, afauzi@ubpkarawang.ac.id

Abstract. *Indonesia as an agrarian country where the population is mostly farmers. It is expected that the government's attention to improve land management and production. Productive land is converted into housing and industry. Government need to maps land and productive land in order to make the right decision. Especially in West Java province which has vast agricultural land, every year there is a decrease in agricultural land. The agricultural sector includes many sectors, many farmers still live below the poverty line. Because facilities and financing are still minimal from the local government. Considering that almost all villages in the city district don't have true and accurate data on the condition of the potential of the village owned, especially the classification of income in agriculture. The purpose of this study is to classify regions that have a level of income from the agricultural sector, and analyze the income so that it can be better in decision making. The research method is how the Decision Tree algorithm for the classification of regencies and cities in West Java Province. Besides being able to classify regions, research also measures the success or accuracy of the algorithm used. The result of the research is to classify data automatically. After doing this research, with obtained Accuracy 90%.*

Keywords: data Mining, classification, Decision Tree, agriculture.

Abstrak. Indonesia sebagai negara agraris dimana penduduk sebagian besar petani. Diharapkan perhatian pemerintah untuk meningkatkan pengelolaan lahan dan produksinya. Lahan produktif beralih fungsi menjadi perumahan dan industri. Pemerintah memetakan lahan dan lahan produktif agar dapat mengambil keputusan yang tepat. Terutama propinsi Jawa Barat memiliki lahan pertanian yang luas, tetapi tiap tahun mengalami penurunan lahan pertanian. Sektor pertanian meliputi banyak bidang namun masih banyak petani kehidupannya dibawah garis kemiskinan. Karena fasilitas dan pembiayaan masih minim dari pemerintah setempat. Mengingat hampir semua desa yang ada di kabupaten kota tidak mempunyai data yang benar dan akurat tentang kondisi permasalahan dan potensi desa yang dimiliki, terutama klasifikasi pendapatan dibidang pertanian. Tujuan penelitian ini adalah klasifikasi daerah yang memiliki tingkat pendapatan dari sektor pertanian dan dengan menganalisa agar lebih baik dalam pengambilan keputusan. Metode penelitian adalah algoritma *Decision tree* untuk klasifikasikasi daerah kabupaten dan kota di propinsi Jawa Barat. Hasil penelitian adalah untuk mengklasifikasikan data secara otomatis dengan hasil akhir keberhasilan algoritmenya di peroleh tingkat Akurasi 90 %.

Kata Kunci: penambangan data, klasifikasi, *Decision Tree*, pertanian

PENDAHULUAN

Potensi pertanian di Jawa Barat tersebar secara merata di seluruh daerah, yang meliputi komoditas padi, palawija, dan hortikultural. Selain itu, jenis sayuran dan buah-buahan di daerah Jawa Barat memiliki potensi yang sangat menjanjikan. hasil pertanian utama di daerah

Jawa Barat 1997 adalah sebagai berikut: hasil produksi padi mencapai 10.352.650 ton dengan luas panen sawah dan ladang 2.040.680 ha; hasil produksi jagung 336.014 ton dengan luas panen 127.994 ha, dan hasil produksi ubi kayu 1.816.487 ton dengan luas panen 141.637 ha.

Lahan pertanian di Jawa Barat terus mengalami penyusutan seiring dengan kebutuhan sarana dan prasarana bagi penduduk di Jabar yang begitu banyak. Apalagi, di Jabar terdapat banyak proyek infrastruktur yang praktis turut menggerus lahan pertanian. Kepala Dinas Tanaman Pangan dan Hortikultura Provinsi Jabar Hendi Jatnika mengakui, lahan pertanian di Jabar terus berkurang. Meski tak hafal datanya secara pasti, dia menyatakan bahwa pihaknya berupaya mempertahankan produktivitas pertanian di Jabar. Kalau lahan, tidak akan bertambah, pasti itu. Jabar dengan penduduk yang begitu banyak, itu membutuhkan lahan buat pemukiman, sarana umum, usaha atau sektor lainnya. Nah, tentunya dengan lahan tersisa yang ada, kami harus pertahankan," kata Hendi di Lembang, Kabupaten Bandung Barat.

Pertanian adalah kegiatan pemanfaatan sumber daya hayati yang dilakukan manusia untuk menghasilkan bahan pangan, bahan baku industri, atau sumber energi, serta untuk mengelola lingkungan hidupnya. Kegiatan pemanfaatan sumber daya hayati yang termasuk dalam pertanian biasa dipahami orang sebagai budidaya tanaman atau bercocok tanam (bahasa Inggris: *crop cultivation*) serta pembesaran hewan ternak (*raising*), meskipun cakupannya dapat pula berupa pemanfaatan mikroorganisme dan bioenzim dalam pengolahan produk lanjutan, seperti pembuatan keju dan tempe, atau sekadar ekstraksi semata, seperti penangkapan ikan atau eksploitasi hutan.

Ekonomi pertanian adalah aktivitas ekonomi yang terkait dengan produksi, distribusi, dan konsumsi produk dan jasa pertanian mengkombinasikan produksi pertanian dengan teori umum mengenai pemasaran dan bisnis adalah sebuah disiplin ilmu yang dimulai sejak akhir abad ke 19, dan terus bertumbuh sepanjang abad ke-20.

Pohon keputusan (*Decision tree*)

Pohon keputusan (*Decision tree*) adalah struktur pohon, yang mana setiap node leaf atau daun mempresentasikan kelompok data kelas tertentu. Level node paling atas disebut dengan root, yang memiliki pengaruh terbesar pada suatu kelas, dan dapat sebagai rule yang paling pertama yang mempengaruhi.

Decision tree memiliki beberapa algoritme yang termasuk seperti ID3, C,45 C,50, algoritme tersebut sangat populer didalam klasifikasi dan ekstraksi rule, ini termasuk dari datamining. Klasifikasi itu proses mencari pola dari sekumpulan data atau fungsi yang dapat memisahkan data kelas yang satu dengan yang lain (Quinlan, 1987).

Tanaman Pangan adalah segala jenis tanaman yang di dalamnya terdapat karbohidrat dan protein sebagai sumber energi manusia. Pembentukan model prediksi menggunakan metode C4.5. Pada algoritma C4.5 dilakukan perhitungan entropy dan information gain. Sebagai atribut sumber untuk memperoleh node akar dan node lainnya. Berdasarkan hasil klasifikasi menggunakan algoritma C4.5 menunjukkan bahwa algoritma C4.5 cocok digunakan untuk memprediksi hasil panen padi(Sri Maya Sari Sianturi, n.d.).

Tanaman padi merupakan tanaman pangan yang rentan terserang hama. Pengenalan terhadap jenis hama yang menyerang merupakan langkah awal yang sangat penting untuk menunjang keberhasilan dalam usaha pengendaliannya. Hama tanaman padi tersebut dapat menjadi kendala bagi petani untuk bisa meningkatkan produksi. Dengan diadakannya klasifikasi pada hama padi menggunakan algoritma C4.5 diharapkan para petani dapat segera mengetahui jenis hama padi dan tingkat keanasannya. Sehingga diharapkan mampu menangani hawa secara tepat, agar tidak terjadi kerusakan dan gagal panen(Sularno & Anggraini, 2017).

Algoritma ID3 membentuk pohon keputusan dengan metode *divide and conquer* data secara rekursif dari atas ke bawah. Strategi pembentukan *Decision Tree* dengan algoritma ID3 adalah:

1. Pohon sebagai node tunggal baik akar dan root untuk representasikan semua data.
2. Setelah node root dibentuk, data pada akar akan diukur dengan information gain untuk dipilih atribut mana yang akan dijadikan atribut pembaginya.
3. Sebuah cabang dari atribut yang pilih menjadi pembagi dan kata akan distribusikan ke dalam cabang masing masing
4. Algoritma Decision Tree secara rekurdif, ketika sebuah atribut telah dipilih menjadi node pembagi atau cabang, maka atribut tersebut tidak diikutkan lagi dalam perhitungan nilai information gain.
5. Proses rekursif akan berhenti jika salah satu sesuai kondisi terpenuhi.
 - a. Semua data dari anak cabang telah termasuk dalam kelas yang sama
 - b. Semua atribut yang dipakai, tetapi masih tersisa data dalam kelas berbeda.
 - c. Tidak terdapat data pada anak cabang baru

Dalam Analisa keputusan, pohon keputusan dan diagram pengaruh yang terkait erat digunakan sebagai alat pendukung keputusan visual dan analitik, dimana nilai yang diterapkan atau utilities dari alternative yang bersaing dihitung (Olson, D. L., & Delen, 2008).

Pohon keputusan adalah sebagai pendukung keputusan menggunakan model keputusan seperti pohon / tree. Ini salah satu cara algoritma yang hanya berisi pernyataan kontrol berisi (Kamiński, Jakubczyk, & Szufel, 2018). *Decision tree* terdiri dari tiga jenis simpul:

- a. Node keputusan diwakilkan oleh kuadrat
- b. Probabilitas node diwakilkan lingkaran
- c. Node akhir diwakilkan oleh segitiga

Tree adalah suatu metode klasifikasi. Klasifikasi adalah proses menemukan kumpulan pola atau atau fungsi yang mendeskripsikan dan memisahkan kelas data dengan yang lain, agar dapat digunakan dalam memprediksi data yang belum memiliki kelas tertentu.

Menurut (Fergie Joanda Kaunang, Reymon Rotikan, 2018) pertanian merupakan salah satu sektor industri yang menjadi bagian pekerjaan dalam pemenuhan kebutuhan pokok seperti tanaman pangan pertanian, mengolah dengan metode algoritme Decision Tree J48, dengan tujuan memberikan sumbangsih dibidang pertanian terlebih khusus pada petani. Menurut (Herman Santoso Pakpahan, Fenny Indar, 2018) untuk usaha menurunkan persentasi angka kemiskinan Kabupaten Kutai Kartanegara membantu sebuah kebijakan dalam bentuk dan jaminan sosial melalui pemberian santunan kepada warga tidak mampu (SWTM) pada masyarakat petani. Menurut (Siti Maesaroh, 2017) melakukan Analisa terhadap peningkatan padi dengan melihat beberapa factor penting untuk mendukung produktivitas pertanian, menggunakan metode klasifikasi dengan mengukur akurasi dengan metode uji K Fold Cross Validation.

Untuk melakukan pemisahan split objek dilakukan tes terhadap atribut dengan mengukur tingkat ketidakmurnian pada suatu simpul atau node. Pada algoritme C 4.5 menggunakan rasio perolehan / gain ratio. Sebelum rasio diperoleh, perlu menghitung dulu nilai informasi dalam satuan bits dari suatu kumpulan objek. Cara menghitung dilakukan rumus entropy.

$$\text{Entropy (S)} = \sum_{j=1}^k -p_j \log_2 p_j \quad 1$$

Dimana:

S = Hinpunan Dataset

K = Banyaknya partisi S

Pj = Probabilitas jumlah Ya dan No dibagi total data

Kemudian menghitung perolehan informasi dari output data atau variabel dependent y yang dikelompokkan berdasarkan atribut A, dinotasikan dengan gain (y, A). Perolehan informasi, gain (y, A), dari atribut A relative terhadap output data y adalah dengan rumus information gain:

$$\text{Gain (A)} = \text{Entropy (S)} - \sum_{i=1}^k \frac{S_i}{S} \times \text{Entropy (S}_i) \quad 2$$

Dimana:

S = Ruang data sample yang digunakan untuk training

A = Atribut

I Si I = Jumlah saple Nilai V

I S I = Jumlah seluruh Sample

Entropy (Si) = Entropy sample Sample yang memiliki i

Klasifikasi

Klasifikasi adalah merupakan salah satu topik utama dalam *data mining* atau *machine learning*. Klasifikasi adalah suatu pengelompokan data dimana data yang digunakan tersebut mempunyai kelas label atau target. Sehingga algoritme untuk menyelesaikan masalah klasifikasi dikategorisasikan ke dalam *supervised learning* atau pembelajaran yang terawasi. Maksud dari pembelajarannya adalah data label atau target ikut berperan sebagai supervisor mengawasi proses pembelajaran dalam mencapai akurasi atau presisi tertentu (Olson, D. L., & Delen, 2008).

Confusion matrix

Confusion matrix adalah suatu pengukuran terhadap kinerja suatu system klasifikasi yang merupakan hal penting. Kinerja klasifikasi menggambarkan seberapa baik sistem klasifikasi data. *Confusion matrix* merupakan salah satu metode yang dapat digunakan untuk mengukur kinerja dalam metode klasifikasi. Dasarnya *confusion matrix* mengandung informasi yang membandingkan klasifikasi yang dilakukan oleh system dengan hasil klasifikasi yang sebenarnya. Pada pengukuran kinerja menggunakan *confusion matrix*, terdapat 4 istilah sebagai representasi hasil proses klasifikasi. Keempat istilah tersebut adalah *True Positive (TP)*, *True Negative (TN)*, *False Positive (FP)* dan *False Negative (FN)*. Pada jenis klasifikasi binary yang hanya memiliki 2 keluaran kelas, *confusion matrix* dapat disajikan seperti pada Tabel berikut.

Tabel 1. *Confusion matrix*

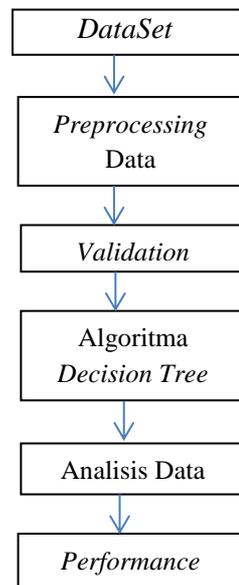
Kelas	Positif	Negatif
Positif	TP (<i>True Positive</i>)	FN (<i>False Negative</i>)
Negatif	FP (<i>False Positive</i>)	TN (<i>True Negative</i>)

Keterangan:

- TP adalah *True Positive*, yaitu jumlah data positif yang terklasifikasi dengan benar
- TN adalah *True Negative*, yaitu jumlah data negatif yang terklasifikasi dengan benar.
- FN adalah *False Negative*, yaitu jumlah data negatif namun terklasifikasi salah.
- FP adalah *False Positive*, yaitu jumlah data positif namun terklasifikasi salah

METODE

Metode penelitian adalah langkah-langkah yang digunakan agar penelitian ini terstruktur dengan baik, menyebabkan proses penelitian dapat dipahami dan diikuti oleh pihak lain. Penelitian yang akan dilakukan adalah bagaimana klasifikasi daerah menggunakan algoritme *decision tree*. Objek penelitian ini adalah data penghasilan sektor pertanian dipropinsi Jawa Barat, pada tahun 2013. Tersedia di laman dinas pertanian: <https://data.go.id>. Berikut langkah langkah penelitian yang digunakan dibawah ini.



Gambar 1. Diagram Alir yang digunakan

Penjelasan sebagai berikut:

Data set yang digunakan adalah database dari pemerintah bagian statistic yang memiliki banyak atribut dibidang pertanian yaitu Tanaman Padi dan Palawija, Tanaman hortikultura, Tanaman Perkebunan, Peternakan, Budidaya ikan di tambak/air payau, Budidaya ikan di kolam air tawar, Budidaya ikan di sawah, Budidaya ikan di perairan umum, Budidaya ikan hias, Penangkapan ikan di laut, Penangkapan ikan di perairan umum, Tanaman Kehutanan, Pemungutan hasil hutan/Penangkapan satwa liar, Jasa Pertanian dan pembibitan tanaman. Total atribut yang digunakan adalah 16 atribut, dengan jumlah kabupaten kota yang ada di Jawa Barat.

Preprocessing adalah proses setelah didapatkan dataset mentah dari situs pemerintah.

Preprocessing yang digunakan adalah *data selection, data integration, data transformastion, data reduction*. Proses ini yang dilakukan untuk *preprocessing* data sehingga data siap digunakan menjadi *task relevant data*.

Validation adalah memastikan data yang digunakan sudah valid dan kesesuaian atribut yang digunakan untuk mengambil informasi yang dibutuhkan seperti transformasi data sehingga didapatkan data yang sudah dikurangi tingkat kompleksitasnya, kemudian akan memudahkan proses pengolahan karena ukuran data yang diproses sudah menjadi lebih kecil.

Algoritme Decision Tree adalah pengolahan data dimana yang pertama melakukan pengolahan algoritma decision tree secara otomatis untuk decision tree dengan tools *software Rapid Miner*.

Analisis data adalah melihat hasil dari algoritma decision tree, serta di analisa data dengan informasi yang berharga. Dan permasalahan yang ditemui pada saat penelitian ini

Hasil / performance adalah hasil dari pengolahan data yang sudah dilakukan akan ditentukan *rule* yang terbentuk untuk memberikan suatu keputusan dari algoritma dan menampilkan hasil akurasi dari algoritma decision tree.

HASIL DAN PEMBAHASAN

Hasil penelitian yang didapatkan merupakan hasil klasifikasi menggunakan *RapidMiner* sebagai tools perangkat lunak untuk pengolahan data. Dengan menggunakan prinsip dalam algoritma *data mining*, *RapidMiner* dapat mengekstrak pola-pola dari data set yang besar dengan mengkombinasikan kecerdasan buatan, metode statistika, dan *database*. Untuk memudahkan

penggunanya dalam melakukan pengujian data yang sangat banyak dengan menggunakan banyak operator.

Hasil ditampilkan secara visual berupa grafik, penelitian berupa klasifikasi dengan dataset semua kab/kota yang ada di Jawa Barat dengan menggunakan 16 atribut. Menampilkan hasil *performance* dari algoritma *Decision Tree*. Hasil penelitian seperti dibawah ini:

Decision Tree akan mengidentifikasi atribut mana yang paling besar atau dominan dari hasil pertanian di propinsi Jawa Barat, lalu hasilnya di ilustrasikan dalam bentuk pohon keputusan agar lebih mudah memahaminya.



Gambar 2. Model Decision Tree

Berikut *rule* yang didapatkan dari dataset yang digunakan:

1. Penangkapan satwa liar > 12.070 dikategorikan tinggi
2. Penangkapan satwa liar ≤ 12.070 , penangkapan ikan hasil perairan umum > 27.405 dikategorikan tinggi
3. Penangkapan satwa liar ≤ 12.070 , penangkapan ikan hasil perairan umum ≤ 27.405 , tanaman kebun > 4.685 dikategorikan sedang
4. Penangkapan satwa liar ≤ 12.070 , penangkapan ikan hasil perairan umum ≤ 27.405 , tanaman kebun ≤ 4.685 dikategorikan rendah

Langkah selanjutnya adalah menampilkan *confusion matrix* untuk mengevaluasi algoritma *Decision Tree* untuk klasifikasi. Dari tabel *confusion matrix* tersebut untuk mendapatkan nilai akurasi yang merupakan perbandingan antara data yang terklasifikasi benar dengan keseluruhan data.

Tabel 2. *performance accuracy*

Accuracy : 90 %

	True Tinggi	True Sedang	True rendah
Pred tinggi	3	0	0
Pred sedang	0	6	0
Pred rendah	0	3	18
Class recall	100 %	66.67%	100%

Berikut dibawah ini hasil *performance* penelitian ini, dengan menampilkan hasil setiap atribut untuk dapat dijadikan dasar ambil keputusan. Berikut atribut yang digunakan dalam ini adalah: Nama nama kabupaten/kota adalah Bogor, Sukabumi, Cianjur, Bandung, Garut, Tasikmalaya, Ciamis, Kuningan, Cirebon, Majalengka, Sumedang, Indramayu, Subang, Purwakarta, Karawang, Bekasi, Bandung Barat, Kota Bogor, Kota Sukabumi, Kota Bandung, Kota Cirebon, Kota Bekasi, Kota Depok, Kota Cimahi, Kota Tasikmalaya, Kota Banjar.

Cianjur	10.1	1.8	39.4	4.8	9.2	15.7	9.5	15.9	TINGGI	1.0	0.0	0.0	TINGGI
Garut	4.1	6.9	2.1	7.0	17.7	10.7	8.2	10.7	SEDANG	0.0	1.0	0.0	SEDANG
Tasikmalaya	4.8	0.0	0.0	5.5	3.7	16.3	7.5	12.7	SEDANG	0.0	1.0	0.0	SEDANG
Cirebon	0.5	43.0	0.5	2.4	3.3	0.2	3.0	0.5	SEDANG	0.0	0.2	0.8	RENDAH
Majalengka	1.9	0.0	0.0	6.6	8.1	1.7	5.4	1.4	RENDAH	0.0	0.2	0.8	RENDAH
Indramayu	0.0	12.9	15.5	3.5	1.8	0.0	7.6	0.0	SEDANG	0.0	0.2	0.8	RENDAH
Purwakarta	0.0	0.0	12.7	3.3	0.9	3.8	2.0	4.8	SEDANG	0.0	1.0	0.0	SEDANG
Karawang	0.0	8.4	8.6	4.1	1.5	0.2	4.9	0.7	SEDANG	0.0	0.2	0.8	RENDAH
Bandung Barat	0.0	0.0	2.1	9.3	11.7	1.1	4.6	4.8	SEDANG	0.0	1.0	0.0	SEDANG
Kota Bandung	0.0	0.0	0.0	0.4	0.1	0.0	0.2	0.0	RENDAH	0.0	0.2	0.8	RENDAH
Kota Cirebon	1.9	0.8	0.0	0.2	0.1	0.0	0.0	0.0	RENDAH	0.0	0.2	0.8	RENDAH
Kota Depok	0.0	0.0	0.0	0.2	0.7	0.0	0.1	0.0	RENDAH	0.0	0.2	0.8	RENDAH
Kota Cimahi	0.0	0.0	0.0	0.4	0.3	0.0	0.1	0.0	RENDAH	0.0	0.2	0.8	RENDAH
Kota Tasikmalaya	0.0	0.0	0.0	0.3	0.1	0.1	0.5	0.3	RENDAH	0.0	0.2	0.8	RENDAH
Kota Banjar	4.8	0.0	0.0	0.5	0.2	0.3	0.5	0.7	RENDAH	0.0	0.2	0.8	RENDAH

Gambar 3. Tabel hasil penelitian perbandingan hasil klasifikasi manual dengan algoritma *Decision Tree*

Hasil klasifikasi pada penelitian ini adalah memiliki ketidaksamaan klasifikasi yaitu kabupaten Cirebon, karawang dan Bandung Barat. Permasalahan adalah kegagalan algoritma *Decision Tree* dalam kasus ini adalah ke tidak berhasil pada klasifikasi sedang, sedangkan klasifikasi tinggi dan rendah hasil sama dengan hasil klasifikasi algoritma *Decision Tree*.

Kab/Kota	Klasifikasi	prediction(Klasifikasi)
Bogor	TINGGI	TINGGI
Cianjur	TINGGI	TINGGI
Bandung	RENDAH	RENDAH
Tasikmalaya	SEDANG	SEDANG
Kota Bogor	RENDAH	RENDAH
Kota Sukabumi	RENDAH	RENDAH
Kota Bekasi	RENDAH	RENDAH
Bandung	RENDAH	RENDAH
Tasikmalaya	SEDANG	SEDANG
Ciamis	TINGGI	TINGGI
Sumedang	RENDAH	RENDAH
Subang	RENDAH	RENDAH
Kota Bogor	RENDAH	RENDAH
Kota Sukabumi	RENDAH	RENDAH
Kota Depok	RENDAH	RENDAH
Tasikmalaya	SEDANG	SEDANG
Cirebon	SEDANG	RENDAH
Majalengka	RENDAH	RENDAH
Purwakarta	SEDANG	SEDANG
Karawang	SEDANG	RENDAH
Bekasi	RENDAH	RENDAH
Kota Sukabumi	RENDAH	RENDAH
Kota Bandung	RENDAH	RENDAH
Kota Bekasi	RENDAH	RENDAH
Kota Cimahi	RENDAH	RENDAH
Kota Banjar	RENDAH	RENDAH
Bogor	TINGGI	TINGGI
Cianjur	TINGGI	TINGGI
Garut	SEDANG	SEDANG
Tasikmalaya	SEDANG	SEDANG
Cirebon	SEDANG	RENDAH
Majalengka	RENDAH	RENDAH
Indramayu	SEDANG	RENDAH
Purwakarta	SEDANG	SEDANG
Karawang	SEDANG	RENDAH
Bandung Barat	SEDANG	SEDANG
Kota Bandung	RENDAH	RENDAH
Kota Cirebon	RENDAH	RENDAH
Kota Depok	RENDAH	RENDAH
Kota Cimahi	RENDAH	RENDAH
Kota Tasikmalaya	RENDAH	RENDAH
Kota Banjar	RENDAH	RENDAH

Gambar 4. Tabel ketidaksesuaian hasil klasifikasi

Untuk penelitian sebelumnya yang berhubungan dengan penelitian ini hanya membahas *rule* yang dihasilkan algoritma C.45 sedangkan penelitian ini mendapatkan *rule* dan akurasi algoritma yang digunakan.

PENUTUP

Simpulan

Tahapan terakhir penelitian ini adalah tahap pengambilan kesimpulan dengan dataset sebanyak 16 atribut pada tahun 2013. Hasil klasifikasi kabupaten kota berdasarkan hasil bidang pertanian dengan algoritma *Decision Tree* dengan akurasi 90%. Dan *rule* yang didapatkan dari algoritma *Decision Tree* adalah ada 4 *rule* yang akan dijadikan dasar keputusan dalam klasifikasi. Rule 1. Penangkapan satwa liar > 12.070 dikategorikan tinggi; rule 2 Penangkapan satwa liar ≤ 12.070 , penangkapan ikan hasil perairan umum > 27.405 dikategorikan tinggi; rule 3 Penangkapan satwa liar ≤ 12.070 , penangkapan ikan hasil perairan umum ≤ 27.405 , tanaman kebun > 4.685 dikategorikan sedang; rule 4 Penangkapan satwa liar ≤ 12.070 , penangkapan ikan hasil perairan umum ≤ 27.405 , tanaman kebun ≤ 4.685 dikategorikan rendah.

Saran

Dengan hasil dan kesimpulan dari penelitian ini penulis memberikan beberapa saran adalah:

1. Pengujian dengan data sama perlu menggunakan algoritma yang lain.
2. Pengujian bisa digunakan *tools* yang lain.
3. Bisa jadi referensi bagi peneliti ke depannya dalam menganalisa data sejenis atau permasalahan yang mirip.

DAFTAR PUSTAKA

- Fergie Joanda Kaunang, Reymon Rotikan, G. S. T. (2018). Tree, Pemodelan Sistem Prediksi Tanaman Pangan Menggunakan Algoritma Decision. *Cogito Smart Journal*, 4(1), 213-218. Retrieved from https://www.researchgate.net/publication/326054689_Pemodelan_Sistem_Prediksi_Tanaman_Pangan_Menggunakan_Algoritma_Ddecision_Tree
- Herman Santoso Pakpahan, Fenny Indar, M. W. (2018). Penerapan Algoritma CART Decision Tree pada Penentuan Penerima Program Bantuan Pemerintah Daerah Kabupaten Kutai Kartanegara. *JURTI*, 2(1), 27–36. Retrieved from e-journals.unmul.ac.id/index.php/INF/article/download/1372/1058
- Kamiński, B., Jakubczyk, M., & Szufel, P. (2018). A framework for sensitivity analysis of decision trees. *Central European Journal of Operations Research*, 26(1), 135–159. <https://doi.org/10.1007/s10100-017-0479-6>
- Olson, D. L., & Delen, D. (2008). *Advanced Data Mining Techniques*. Springer.
- Quinlan, J. R. (1987). Simplifying decision trees. *International Journal of Man-Machine Studies*, 27(3), 221–234. [https://doi.org/10.1016/S0020-7373\(87\)80053-6](https://doi.org/10.1016/S0020-7373(87)80053-6)
- Siti Maesaroh, K. (2017). Sistem Prediksi Produktifitas Pertanian Padi Menggunakan Data Mining. *Unmul*, 7(2), 25–30. Retrieved from eprints.dinus.ac.id/16925/1/jurnal_16115.pdf
- Sri Maya Sari Sianturi, N. A. H. (n.d.). ANALISA DATA PERTANIAN TANAMAN PANGAN UNTUK MEMREDIKSI HASIL PANEN DENGAN DATA MINING ALGORITMA C.45 (STUDI KASUS: DINAS TANAMAN PANGAN dan HOLTIKUTURA PROVINSI SUMUT). *Jurnal Pelita Informatika*, 18(2), 212–219.
- Sularno, S., & Anggraini, P. (2017). PENERAPAN ALGORITMA C4.5 UNTUK KLASIFIKASI TINGKAT KEGANASAN HAMA PADA TANAMAN PADI (Studi Kasus : Dinas Pertanian Kabupaten Kerinci). *Jurnal Sains Dan Informatika*, 3(2), 161. <https://doi.org/10.22216/jsi.v3i2.2779>