

Systematic Literature Review: Penerapan Metode Algoritma C4.5 untuk Klasifikasi

Dewi Srifani¹, Bangkit Indarmawan Nugroho^{*2}, Nugroho Adhi Santoso³

¹Jurusan Studi Teknik Informatika, ^{2,3}Jurusan Program Studi Sistem Informasi, STMIK YMI Tegal
Jl.Pendidikan No 1 Margadana Kota Tegal, Jawa Tengah, Indonesia 52142

E-mailsrifanidewi@gmail.com¹, efbeterang@gmail.com², nugrohadhisantoso29@gmail.com³

Abstract—Data mining is extracting data in processing information with the aim of finding important patterns in piles of data. With data mining we can classify, predict, and make a decision. Classification is a way of grouping a data according to the characteristics of a data to be classified. In the process, the classification is divided into two, namely manually and with the help of technology. Manual classification is a classification carried out by humans without the help of technology, while classification with the help of technology has several algorithms, including C4.5, Naive Bayes, Fuzzy, and K-Nearest Neighbor. For classification, a systematic approach is used in the form of a systematic literature review (SLR). SLR is defined as a process in which the identification, assessment, and interpretation of all available research evidence is carried out with the aim of answering a number of research questions.

Abstrak—Data mining merupakan penggalian data dalam pengolahan informasi dengan tujuan untuk menemukan pola-pola penting dalam tumpukan data. Dengan data mining kita dapat melakukan pengklasifikasian, memprediksi, dan membuat suatu keputusan. Klasifikasi adalah cara pengelompokan suatu data menurut karakteristik dari suatu data yang akan di klasifikasikan. Dalam prosesnya, klasifikasi dibagi menjadi dua yaitu secara manual dan dengan bantuan teknologi. Klasifikasi secara manual yaitu klasifikasi yang dilakukan oleh manusia tanpa bantuan teknologi, sedangkan klasifikasi dengan bantuan teknologi memiliki beberapa Algoritma, antara lain C4.5, Naive bayes, Fuzzy, dan K-Nearest Neighbour. Dalam melakukan studi kepustakaan yang mengulas mengenai topik penerapan metode Algoritma C4.5 untuk klasifikasi, dipergunakan sebuah pendekatan secara sistematis berupa systematic literature review (SLR). SLR didefinisikan sebagai sebuah proses yang di dalamnya melakukan identifikasi, penilaian, dan interpretasi seluruh barang bukti penelitian yang tersedia dengan tujuan untuk menjawab sejumlah pertanyaan penelitian

Kata kunci : Algoritma C4.5, Data mining, Metode klasifikas, SLR.

I. PENDAHULUAN

Data mining merupakan penggalian data dalam pengolahan informasi dengan tujuan untuk menemukan pola-pola penting dalam tumpukan data[1]. Dengan data mining kita dapat melakukan pengklasifikasian, memprediksi, dan membuat suatu keputusan[2]. Klasifikasi adalah cara pengelompokan suatu data menurut karakteristik dari suatu data yang akan di klasifikasikan. Dalam prosesnya, klasifikasi dibagi menjadi dua yaitu secara manual dan dengan bantuan teknologi. Klasifikasi secara manual yaitu klasifikasi yang dilakukan oleh manusia tanpa bantuan teknologi, sedangkan klasifikasi dengan bantuan teknologi memiliki beberapa Algoritma, antara lain C4.5, Naive bayes, Fuzzy, dan K-Nearest Neighbour[3].

Naive bayes merupakan salah satu metode klasifikasi untuk memprediksi hasil dan perhitungan probabilitas, namun dalam pengambilan suatu keputusan sangat bergantung pada pengetahuan sebelumnya, serta akurasi yang kurang dalam asumsi variable yang Independent[4]. Fuzzy juga merupakan suatu metode klasifikasi untuk pengumpulan nilai linguistik dengan bobot yang berbeda, namun Fuzzy memiliki batasan nilai linguistik yang sangat berpengaruh pada akurasi metode tersebut[5]. K-Nearest Neighbour adalah metode klasifikasi untuk melakukan evaluasi nilai terdekat dalam suatu data serta menemukan 2

data dengan jarak terdekat, namun dalam metode ini memerlukan nilai parameter berdasarkan jarak yang belum jelas mengenai jarak apa yang perlu digunakan[4].

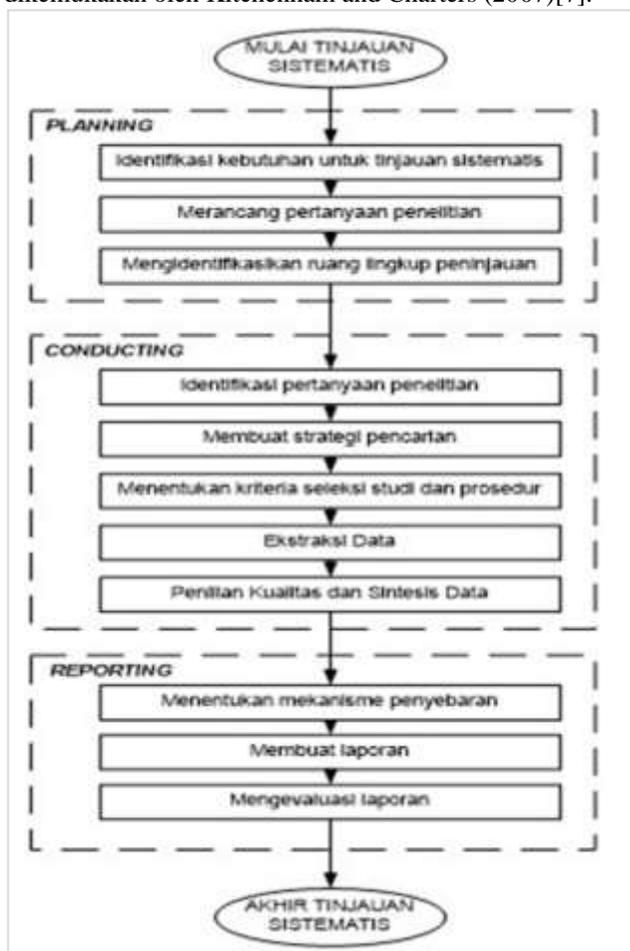
Algoritma C4.5 merupakan metode klasifikasi yang paling banyak digunakan serta memiliki keunggulan dibandingkan Algoritma yang lainnya. Kelebihan Algoritma C4.5 mudah diimplementasikan, dengan akurasi yang dapat diterima, dan efisien dalam penanganan atribut tipe diskrit serta dapat menangani atribut tipe numerik[6]

Dari banyaknya metode yang diusulkan untuk Klasifikasi, maka pada makalah ini akan dilakukan tinjauan pustaka systematic literature review penerapan metode Algoritma C4.5 untuk Klasifikasi.

II. METODE PENELITIAN

Dalam melakukan studi kepustakaan yang mengulas mengenai topik penerapan metode Algoritma C4.5 untuk klasifikasi, dipergunakan sebuah pendekatan secara sistematis berupa systematic literature review (SLR). SLR sekarang ini banyak dipergunakan sebagai metode pengulasan dalam penelitian di bidang rekayasa perangkat lunak. SLR didefinisikan sebagai sebuah proses yang di dalamnya melakukan identifikasi, penilaian, dan interpretasi seluruh barang bukti penelitian yang tersedia dengan tujuan untuk menjawab sejumlah pertanyaan penelitian. Studi

kepastakaan yang dilakukan untuk penelitian ini mengadopsi pada sumber atau panduan SLR yang dikemukakan oleh Kitchenham and Charters (2007)[7].



Gambar 1. Langkah-langkah SLR

Metodologi yang digunakan terdiri dari tiga fase utama yaitu: perencanaan (planning), pengerjaan (conducting), dan pelaporan (reporting)[8]. Dimana pada setiap fasenya memiliki langkah-langkah yang lebih spesifik, yang akan dijelaskan lebih mendalam pada bagian ini. Gambar 1 merangkum tahapan pelaksanaan tinjauan pustaka sistematis yang digunakan pada penelitian ini:

Fase awal berupa perencanaan (planning) yang dibagi ke dalam tiga tahap yaitu: (a) identifikasi kebutuhan untuk kajian sistematis, (b) merancang pertanyaan penelitian, dan (c) mendefinisikan ruang lingkup peninjauan. Pada fase kedua dilakukan perancangan (conducting) yang terdiri dari: (a) identifikasi pertanyaan penelitian, (b) membuat strategi pencarian, (c) menentukan kriteria seleksi studi dan prosedur, (d) ekstraksi data, (e) penilaian kualitas dan sintesis data. Sedangkan pada fase pelaporan (reporting) terdiri dari: (a) menentukan mekanisme penyebaran, (b) membuat laporan sesuai format, dan (c) mengevaluasi laporan. Fase yang paling krusial dan penting adalah pada fase conducting, karena merupakan tahapan pelaksanaan. Pada tahap ke-tiga dilakukan evaluasi secara iteratif terhadap protokol pengulasan literatur yang dipergunakan selama fase pengerjaan (conducting) dijalankan[7].

A. Identifikasi Pertanyaan Penelitian

Pertanyaan penelitian atau biasa disebut dengan *research question* (RQ) merupakan peran penting dalam strategi pencarian, ekstraksi data, serta tahap analisis. Identifikasi pertanyaan ditujukan untuk menemukan informasi sebanyak mungkin menggunakan strategi pencarian yang sesuai serta dapat menghindari bias publikasi.

Tabel 1. *Research question*

ID	Pertanyaan penelitian (RQ)	Tujuan
RQ1	Objek apa yang akan dipilih untuk Klasifikasi dalam penelitian	Identifikasi jenis objek yang akan dibahas pada penelitian
RQ2	Teknik metode apa yang digunakan atau diusulkan dalam penelitian	Untuk mendapatkan pengetahuan yang mendalam mengenai objek atau metode yang diusulkan pada penelitian
RQ3	Berapa tingkat akurasi yang didapat pada metode yang diusulkan dalam penelitian	Untuk mengetahui besarnya keakurasian pada metode atau teknik yang diusulkan
RQ4	Kelebihan dan kekurangan dari metode klasifikasi yang diusulkan dalam penelitian	Untuk mengetahui hasil akhir dari perbandingan beberapa metode pada klasifikasi di dalam penelitian

B. Strategi Pencarian

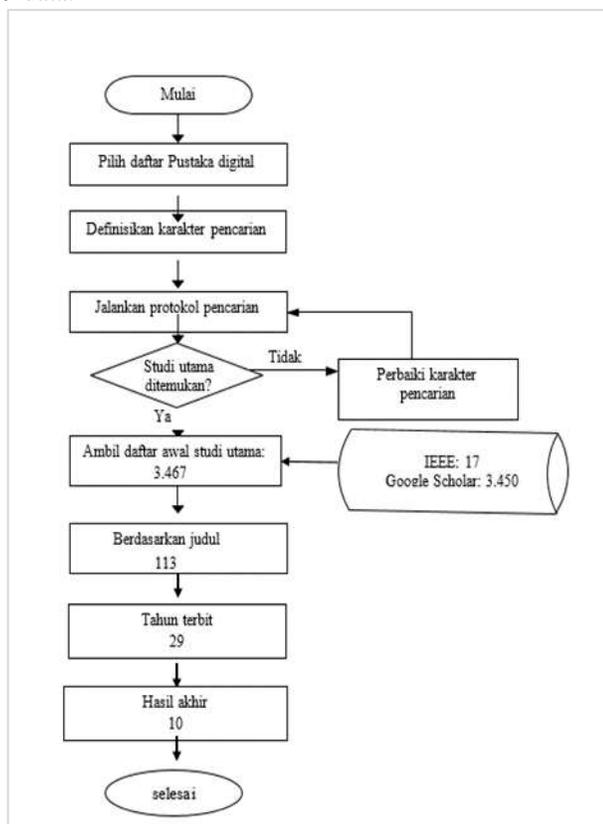
Menentukan istilah pencarian dengan menggunakan sumber data pencarian serta mempertimbangkan kriteria inklusi dan pengecualian, Pada penelitian ini, sumber data pencarian yang digunakan menggunakan perpustakaan elektronik *IEEE Xplore* (<https://ieeexplore.ieee.org/>) dan *Google Scholar* (<https://scholar.google.com/>) yang diakses mulai April 2022.

C. Kriteria seleksi studi

Dalam kriteria seleksi studi adalah dengan menentukan kriteria utama dalam penelitian yaitu penerapan metode Algoritma C4.5 untuk Klasifikasi. Adapun untuk menuju kriteria utama tersebut, perlu dilakukan langkah pembuatan filter dalam pemilihan dan penolakan sesuai dengan kriteria yang didefinisikan.

Pencarian diawali dengan pemilihan daftar pustaka

digital, kemudian mendefinisikan karakter pencarian dengan menjalankan protocol pencarian. Apabila studi utama telah ditemukan maka proses akan berjalan untuk mengambil daftar awal dari pencarian dengan perolehan data untuk IEEE 17 data, sedangkan Google Scholar 3.450 data, sedangkan apabila studi utama belum ditemukan maka perbaiki karakter pencarian tersebut. Jika proses berjalan dengan lancar atau studi utama telah ditemukan maka, proses filterasi atau penyaringan data akan berjalan. Jenis filter yang pertama adalah dengan mengambil studi utama sesuai judul dengan perolehan data 113 data. Kemudian di saring atau filter akhir dengan perolehan data terakhir yaitu 29 data.



Gambar 2. Seleksi data

Keterangan:

1. Study awal: “Penerapan Algoritma C4.5 untuk klasifikasi”.
2. Judul: “Systematic literature review penerapan Algoritma C4.5 untuk klasifikasi”.
3. Tahun terbit: Spesifikasi tahun 2022.
4. Hasil akhir pencarian: penambahan untuk spesifikasi di bidang Pendidikan dengan tahun terbit 2022.

Berikut literature yang di dapat dari seleksi studi pencarian diatas:

Tabel 2.
Literature review

No	Judul	Penulis	Tahun	Citasi
1	<i>Application of the C4.5 for classification of</i>	Nurul Abdilah ¹	2022	[9]

	<i>medical record data AT M.Djamil hospital based on the international disease code</i>	Muhamad Ikhsan ²		
2	Penerapan Algoritma C4.5 dalam mengklasifikasi status gizi balita pada posyandu desa demes kabupaten Lombok timur	Mahfuz Amri Muliawan Nur ¹ L M Samsul ²	2022	[10]
3	Penerapan Algoritma C4.5 untuk klasifikasi mahasiswa penerima bantuan social Covid-19	Eugenius Edsel Barito ¹ Jab Tji Beng ² Desi Arisandi ³	2022	[11]
4	Model aturan dalam menentukan prestasi nilai siswa di SMK GKPS 1 raya menggunakan Algoritma C4.5	Rofiana Simanullang ¹ Dedy Hartama ² Poningsih ³ Iin Parlina ⁴ Muhamad R. lubis ⁵	2022	[12]
5	Penerapan Algoritma decision Tree C4.5 untuk klasifikasi tingkat kesejahteraan keluarga pada desa tiga Dolok	Susi Fitriyah Demanik ¹ Anjar Wanto ² Indra Gunawan ³	2022	[13]
6	Penerapan Algoritma C4.5 dalam klasifikasi penerimaan peserta didik baru	Harnah Syuryanny Nurindah ¹ St.Syahdan ²	2022	[14]
7	Analisis dan prediksi	Muhamad Isra ¹	2022	[15]

	perilaku aparat sipil Negara dalam menempati jabatan fungsional menggunakan Algoritma C4.5			
8	Analisis dan penerapan Algoritma C4.5 untuk klasifikasi penerima Vaksinasi Covid-19 di kota Bengkulu	Pindo Putra Pratama ¹ Harry Witriyon o ² Rozli Toyib ³ Anisya Sonita ⁴	2022	[16]
9	Penerapan Data mining untuk rekomendasi beasiswa pada SD Maria Mediatrix menggunakan Algoritma C4.5	Eko Budiarto ¹ Rino ² Susanto Hariyant o ³ Dera Susilawa ti ⁴	2022	[17]
10	Implementasi Algoritma C4.5 untuk klasifikasi produk laris sepeda motor Honda pada CV Cendana motor cepiring	Tabitha Salsabila ¹ Sulastris ²	2022	[18]

D. Ekstraksi Data

Langkah selanjutnya dalam menentukan studi utama penelitian adalah ekstraksi data dengan iterasi hingga diperoleh data untuk menjawab pertanyaan penelitian (RQ) yang diajukan.

Tabel 3. Properti ekstraksi

Property	Pertanyaan penelitian
P1	Topik penelitian RQ1
P2	Metode Klasifikasi RQ2
P3	Indicator keakuratan RQ3
P4	Karakteristik metode yang diusulkan RQ4

1. Topik penelitian (P1)

Dari hasil pencarian, dalam studi utama Klasifikasi mempunyai beberapa metode selain Algoritma C4.5. terdapat beberapa metode dalam Klasifikasi seperti, Naïve Bayes, fuzzy, K-Nearest Neighbour dan masih banyak lagi.

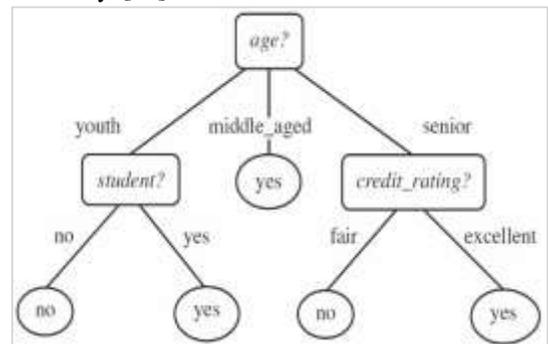
Dalam penelitian ini metode yang digunakan sebagai objek penelitian adalah Algoritma C4.5.

2. Metode Klasifikasi (P2)

• Algoritma C4.5

Metode C4.5 dapat digunakan untuk klasifikasi dengan cara membentuk pohon keputusan (decision tree) dari suatu data[19]. Langkah-langkah metode C4.5:

- Lakukan pengecekan untuk base case.
- Untuk setiap atribut yang ada di data pelatihan, hitung nilai information gain yang didapatkan jika membuat cabang pada atribut.
- Setelah didapatkan atribut yang memberikan nilai information gain tertinggi, lakukan percabangan pada pohon keputusan tersebut.
- Secara rekursif, lakukanlah langkah-langkah b dan c pada sub-atribut yang didapatkan, kemudian tambahkan kedalam percabangan-percabangan berikutnya[19].



Gambar 3. langkah Algoritma C4.5

• Naïve Bayes

Metode ini menggunakan pendekatan teorema Bayes, dengan asumsi setiap atribut bersifat saling bebas, tidak memiliki hubungan atau ketergantungan satu sama lain. Teorema Bayes yaitu menghitung probabilitas suatu kejadian berdasarkan suatu kondisi tertentu[19]

$$P(A/B) = \frac{P(A)P(B/A)}{P(B)} \quad (1)$$

• K-Nearest Neighbour

Merupakan salah satu metode klasifikasi yang menggunakan pendekatan supervised learning sehingga membutuhkan data pelatihan yang sudah dilabeli.

KNN membandingkan data pelatihan dan data pengujian, dimana data dideskripsikan menggunakan variable n. Klasifikasi data dilakukan dengan cara

menghitung tingkat kemiripan atau kedekatan dari data baru tersebut terhadap seluruh data pelatihan yang ada.

$$\text{dist}(X,Y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (2)$$

3. Indikator keakuratan (P3)

- Algoritma C4.5

Tabel 4. *Performance Algoritma C4.5*

Accuracy: 73,33%			
	True pj	True bpj	Class precision
Pred.pj	18	6	75.00%
Pred.bpj	2	4	66.67%
Class recal	90.00%	40.00%	

Sumber : jurnal penelitian[20]

- *Naïve Bayes*

Tabel 5. *Performance Naïve Bayes*

Accuracy: 80.00%			
	True pj	True bpj	Class precision
Pred.pj	17	3	85.00%
Pred.bpj	3	7	70.00%
Class recal	85.00%	70.00%	

Sumber : jurnal penelitian[20]

- *K-Nearest Neighbour*

Tabel 6. *Pperformance KNN*

Accuracy: 70.00%			
	True pj	True bpj	Class precision
Pred.pj	16	5	76.19%
Pred.bpj	4	5	55.56%
Class recal	80.00%	50.00%	

Sumber: jurnal penelitian[20]

Dari data akurasi diatas diperoleh tingkat akurasi:

- *Algoritma C4.5* : 73.33%
- *Naïve Bayes* : 80.00%
- *KNN* : 70.00%

Dari hasil akhir data diatas metode *Naïve Bayes* lebih akurat. Tetapi ada satu studi kasus yang menunjukkan

keakurasian Algoritma C4.5 lebih tinggi dari *Naïve Bayes*. Berikut penjelasannya.

Tabel 7. *Perbandingan Kinerja*

Model	Accuracy	Execution time
C4.5	96.4%	0 s
Naïve Bayes	95.11%	0 s

Dari data studi kasus diatas, membuktikan kelayakan akurasi kinerja Algoritma C4.5 lebih tinggi dari *Naïve Bayes*. Table diatas diambil dari studi kasus untuk jurnal penelitian yang berjudul “Perbandingan Kinerja Algoritma C4.5 dan *Naïve Bayes* untuk Klasifikasi Penerima Beasiswa” tahun 2018[21].

Jadi dari penjelasan tersebut diperoleh metode yang diusulkan untuk Klasifikasi adalah Algoritma C4.5

4. Karakteristik metode yang diusulkan (P4)

Kelebihan Algoritma C4.5 adalah dapat membuat pohon keputusan, sehingga lebih efisien dalam menangani atribut diskrit maupun numeric pada Klasifikasi[22].

Sedangkan untuk kelemahan Algoritma C4.5 adalah skalabilitas yang hanya dapat digunakan dan disimpan secara keseluruhan pada waktu yang bersamaan[22].

E. *Penilaian kualitas studi dan sintesis data*

Pada tahap ini merupakan tahapan yang memberikan bantuan dalam interpretasi data yang telah diekstraksi. Penilaian kualitas studi ini akan memberikan bukti untuk jawaban RQ dari data yang telah dikumpulkan sebelumnya.

Sedangkan untuk sintesis data adalah sebagai interpretasi visualisasi dari data yang semula merupakan data yang disajikan mentah. Tahap ini dilakukan guna memperkuat analisa dan presentasi peneliti dalam mengolah dan menyajikan data mengenai Klasifikasi.[7]

III. HASIL DAN PEMBAHASAN

A. *RQ1 - Objek apa yang akan dipilih untuk Klasifikasi dalam penelitian.*

Dari hasil pencarian, dalam studi utama Klasifikasi mempunyai beberapa metode selain Algoritma C4.5. terdapat beberapa metode dalam Klasifikasi seperti, *Naïve Bayes*, fuzzy, *K-Nearest Neighbour* dan masih banyak lagi.

Dalam penelitian ini metode yang digunakan sebagai objek penelitian adalah Algoritma C4.5.

B. *RQ2 - Teknik metode klasifikasi apa yang digunakan atau diusulkan dalam penelitian.*

Pada bagian ini disajikan hasil analisa yang menunjukkan beberapa metode yang ada di Klasifikasi. Dari beberapa metode tersebut disimpulkan untuk penelitian ini menggunakan metode Algoritma C4.5 yang digunakan sebagai objek penelitian untuk Penerapan Metode Algoritma C4.5 untuk Klasifikasi.

C. *RQ3 - Berapa tingkat akurasi yang didapat pada metode yang diusulkan dalam penelitian*

Dari hasil akurasi pada penjelasan di bagian Metodologi, hasil akhir akurasi menunjukkan bahwa Algoritma C4.5 memiliki akurasi lebih tinggi dari beberapa metode Klasifikasi lainnya. Walaupun di beberapa studi kasus menunjukkan bahwa Naïve bayes lebih tinggi tingkat akurasinya, tetapi untuk tingkat akurasi kinerja Algoritma C4.5 terbukti lebih tinggi dari Naïve Bayes, seperti pada Tabel 6 diatas.

D. RQ4 - Kelebihan dan kekurangan dari metode klasifikasi yang diusulkan dalam penelitian.

Kelebihan Algoritma C4.5 adalah dapat membuat pohon keputusan, sehingga lebih efisien dalam menangani atribut diskrit maupun numeric pada Klasifikasi[22].

Sedangkan untuk kelemahan Algoritma C4.5 adalah skalabilitas yang hanya dapat digunakan dan disimpan secara keseluruhan pada waktu yang bersamaan[22].

IV. KESIMPULAN

Systematic literature review untuk penerapan metode klasifikasi dengan refrensi beberapa jurnal yang diambil dari tahun 2018-2022 ini mendapatkan total seleksi pencarian sebesar 29 jurnal. Dari ke 29 jurnal tersebut, kemudian melewati tahap ekstraksi data dengan penambahan spesifikasi penggunaan dibidang pendidikan, sehingga mendapatkan total 10 jurnal sebagai parameter dari Penerapan Metode Algoritma C4.5 untuk Klasifikasi.

Klasifikasi adalah cara pengelompokan suatu data menurut karakteristik dari suatu data yang akan di klasifikasikan. Dalam prosesnya, klasifikasi dibagi menjadi dua yaitu secara manual dan dengan bantuan teknologi. Klasifikasi secara manual yaitu klasifikasi yang dilakukan oleh manusia tanpa bantuan teknologi, sedangkan klasifikasi dengan bantuan teknologi memiliki beberapa Algoritma, antara lain C4.5, Naive bayes, Fuzzy, dan K-Nearest Neighbour.

Dari beberapa metode klasifikasi diatas, disimpulkan bahwa Algoritma C4.5 adalah metode dengan nilai akurasi tertinggi di bidang kinerjanya untuk pengklasifikasian suatu data.

Sedangkan Algoritma C4.5 itu sendiri merupakan metode klasifikasi yang digunakan untuk klasifikasi suatu data dengan cara membuat suatu pohon keputusan dari suatu data tersebut

- [6] S. Liang, A. Ma, S. Yang, Y. Wang, and Q. Ma, "A Review of Matched-pairs Feature Selection Methods for Gene Expression Data Analysis," *Comput. Struct. Biotechnol. J.*, vol. 16, pp. 88–97, 2018, doi: 10.1016/j.csbj.2018.02.005.
- [7] D. J. Hidayat, R. Z. Fathiyana, and G. B. Dharmawan, "Systematic Literature Review : Pengembangan Sistem Intelijen untuk Prediksi Suhu," *J. Informatics Commun. Technol.*, vol. 3, no. 2, pp. 94–107, 2022, doi: 10.52661/j_ict.v3i2.93.
- [8] B. A. Kitchenham and S. Charters, "Guidelines for performing Systematic Literature Reviews in Software Engineering. EBSE Technical Report EBSE-2007-01. School of Computer Science and Mathematics, Keele University," no. January, pp. 1–57, 2007.
- J. Mantik, N. Abdillah, and M. Ihsan, "APPLICATION OF THE C4 . 5 ALGORITHM FOR CLASSIFICATION OF MEDICAL RECORD DATA AT M . DJAMIL HOSPITAL BASED ON THE INTERNATIONAL DISEASE CODE," vol. 6, no. 36, pp. 576–581, 2022.
- [9] J. Informatika, T. Vol, A. Gizi, P. Desa, D. Damai, and K. Kunci, "PENERAPAN ALGORITMA C4.5DALAM MENGLASIFIKASI STATUS GIZI BALITA PADA POSYANDU DESA DEMAS DAMAI KABUPATEN LOMBOK TIMUR," vol. 5, no. 1, 2022.
- [10] P. Algoritma et al., "MAHASISWA PENERIMA BANTUAN SOSIAL COVID-19".
- [11] [12] M. A. C., "Model Aturan dalam Menentukan Prestasi Nilai Siswa di SMK GKPS 1 Raya," vol. 2, no. 1, 2022.
- [12] S. F. Damanik, A. Wanto, and I. Gunawan, "Penerapan Algoritma Decision Tree C4 . 5 untuk Klasifikasi Tingkat Kesejahteraan Keluarga pada Desa Tiga Dolok," vol. 1, pp. 21–32, 2022.
- P. Algoritma and C. Dalam, "Penerapan algoritma c4.5 dalam klasifikasi penerimaan peserta didik baru," vol. 1, pp. 1–6, 2022.
- [13] J. Informasi, "Analisis dan Prediksi Perilaku Aparatur Sipil Negara dalam Menempati Jabatan Fungsional Menggunakan Algoritma C4 . 5," vol. 4, no. Eselon V, pp. 58–63, 2022, doi: 10.37034/jidt.v4i1.186.
- [14] A. Dan, P. Algoritma, C. Untuk, K. P. V. Covid-, and D. I. K. Bengkulu, "Analisis dan penerapan algoritma c4.5 untuk klasifikasi penerima vaksinasi," vol. 1, pp. 105–112, 2022.
- [15] C. Algoritma, "Penerapan Data Mining Untuk Rekomendasi Beasiswa Pada SD Maria Mediatrix Menggunakan," vol. 2, 2022.
- [16] T. Salsabilla, "Implementasi algoritma c4.5 untuk klasifikasi produk laris sepeda motor honda pada cv cendana motor cepiring 1) 1,2)," vol. 7, no. 2, pp. 164–171, 2022.
- [17] [19] Y. Lukito and A. R. Chrismanto, "Perbandingan Metode - Metode Klasifikasi Untuk Indoor Positioning System," vol. 1, pp. 123–131, 2015.
- [18] P. A. C, D. A. N. Naive, and B. Untuk, "Perbandingan algoritma c4.5, knn, dan naive bayes untuk penentuan model klasifikasi penanggung jawab bsi entrepreneur center," vol. 14, no. 2, pp. 169–174, 2018.
- [19] C. Anam and H. B. Santoso, "Perbandingan Kinerja Algoritma C4 . 5 dan Naive Bayes untuk Klasifikasi Penerima Beasiswa," vol. 8, no. 1, pp. 13–19, 2018.
- [20] S. W. Siahaan, ; Kristin, D. R. Sianipar, ; P P P A N W Fikrul, I. R. H. Zer, and ; Dedy Hartama, "Penerapan Algoritma C4.5 Dalam Meningkatkan Kemampuan Bahasa Inggris Pada Mahasiswa," *PETIR J. Pengkaj. dan Penerapan Tek. Inform.*, vol. 13, no. 2, pp. 229–239, 2020, [Online]. Available: <https://doi.org/10.33322/petir.v13i2.1029>

DAFTAR PUSTAKA

- [1] A. A. Nababan, M. Khairi, and B. S. Harahap, "Implementation of K-Nearest Neighbors (KNN) Algorithm in Classification of Data Water Quality," vol. 6, no. 36, pp. 30–35, 2022.
- [2] Y. Mardi, "Data Mining : Klasifikasi Menggunakan Algoritma C4.5," *Edik Inform.*, vol. 2, no. 2, pp. 213–219, 2017, doi: 10.22202/ei.2016.v2i2.1465.
- [3] F. A. D. Aji Prasetya Wibawa, Muhammad Guntur Aji Purnama, Muhammad Fathony Akbar, "Metode-metode Klasifikasi," *Pros. Semin. Ilmu Komput. dan Teknol. Inf.*, vol. 3, no. 1, p. 134, 2018.
- [4] N. Pada and K. Kanker, "Journal of Dinda," vol. 2, no. 1, pp. 11–20, 2022.
- [5] A. Mulyanto, "PENENTUAN RANGKING EVALUASI PRESTASI BELAJAR SISWA MENGGUNAKAN METODE EVALUASI FUZZY BAI-CHEN," vol. 6, no. 2, pp. 1–8, 2021.