

Penerapan Data Mining Menggunakan Algoritma C4.5 Dalam Mengukur Tingkat Kepuasan Para Wali Siswa Pada SMK Letris Indonesia 2

Rengga Herdiansyah^{1*}, Muhammad Ramdani¹

¹Fakultas Ilmu Komputer, Teknik Informatika, Universitas Pamulang, Jl. Raya Puspittek No. 46, Kel. Buaran, Kec. Serpong, Kota Tangerang Selatan. Banten 15310, Indonesia

Email: 1*dosen01101@unpam.ac.id, 2muhammadramdani121.mr60@gmail.com

(* : coresponding author)

Abstrak— Pendidikan adalah proses yang bertujuan menciptakan lingkungan belajar yang kondusif agar siswa dapat mencapai potensi penuh mereka. Algoritma C4.5 adalah metode pohon keputusan (*decision tree*) yang populer karena menghasilkan model klasifikasi yang mudah dipahami dan efektif dalam membandingkan data dengan berbagai jenis atribut. Pendidikan adalah upaya mempersiapkan peserta didik melalui kegiatan orientasi, kegiatan belajar mengajar, penugasan, dan kegiatan aktivitas pendidikan lainnya yang dapat memengaruhi sikap dan perilaku untuk meningkatkan mutu belajar siswa. Pendidikan melibatkan kegiatan yang kompleks, memiliki aspek yang komprehensif, dan dipengaruhi oleh banyak variabel. Mutu pendidikan suatu bangsa ditentukan oleh mutu pendidikannya. Pendidik harus memiliki kualifikasi dan tingkat kemahiran yang memenuhi standar nasional pendidikan. Sekolah yang dimiliki merupakan sekolah menengah kejuruan, juga dikenal sebagai SMK Letris Indonesia, yang awalnya memiliki bangunan kecil di Raya Jombang dan mendapat kepercayaan masyarakat berkat siswa yang bersekolah di SMK Letris Indonesia. Pada tahap pertama, SMK Letris berkomitmen untuk membangun gedung sekolah berlantai empat dengan total enam belas ruang belajar dan administrasi, sehingga SMK Letris Indonesia 2 memiliki 1434 siswa dengan program inti sebagai berikut: multimedia, teknologi jaringan dan jaringan, akuntansi, bisnis daring dan pemasaran, otomatisasi dan tata kelola perkantoran, dan perbankan. Algoritma C4.5 adalah metode pohon keputusan populer yang membantu menurunkan ambang batas kepuasan wali siswa berdasarkan indikator yang relevan.

Kata Kunci: Data Mining, Algoritma C4.5, Pendidikan, Sekolah Menengah Kejuruan

Abstract— *Education is a process aimed at creating a conducive learning environment so that students can reach their full potential. The C4.5 algorithm is a popular decision tree method because it produces classification models that are easy to understand and effective at comparing data with different types of attributes. Education is the effort to prepare students thru orientation activities, teaching and learning activities, assignments, and other educational activities that can influence attitudes and behavior to improve student learning quality. Education involves complex activities, has comprehensive aspects, and is influenced by many variables. The quality of a nation's education is determined by the quality of its educators. Educators must possess qualifications and a level of proficiency that meet national education standards. The school owned is a vocational high school, also known as SMK Letris Indonesia, which initially had a small building on Raya Jombang and gained public trust thanks to the students attending SMK Letris Indonesia. In the first phase, SMK Letris is committed to building a four-story school building with a total of sixteen classrooms and administrative spaces, so that SMK Letris Indonesia 2 has 1434 students with the following core programs: multimedia, network and network technology, accounting, online business and marketing, office automation and governance, and banking. The C4.5 algorithm is a popular decision tree method that helps determine the satisfaction threshold of student guardians based on relevant indicators.*

Keywords: Data Mining, C4.5 Algorithm, Education, Vocational High School

1. PENDAHULUAN

Perkembangan teknologi informasi yang begitu maju saat ini, menyebabkan tingkat akurasi suatu data sangat dibutuhkan dalam kehidupan sehari-hari. Setiap informasi yang ada menjadi suatu hal penting untuk menentukan setiap keputusan dalam situasi tertentu. Hal ini menyebabkan penyediaan informasi menjadi sarana untuk dianalisa dan diringkas menjadi suatu pengetahuan dari data yang bermanfaat ketika pengambilan suatu keputusan dilakukan (Aini & Nasution, 2024). Selain itu juga telah mempengaruhi kehidupan manusia disektor pendidikan yang menunjukkan pertumbuhan signifikan dalam dekade terakhir. Pendidikan sendiri merupakan suatu proses terencana yang bertujuan menciptakan lingkungan belajar kondusif agar peserta didik dapat mengembangkan potensi secara optimal (Tilaar, 2012). Dalam konteks Indonesia, Sekolah

Menengah Kejuruan (SMK) menjadi salah satu jalur strategis bagi siswa untuk memperoleh keterampilan yang relevan dengan kebutuhan dunia kerja maupun wirausaha (Kementerian Pendidikan dan Kebudayaan, 2020).

Di era digital, data menjadi aset penting yang dapat dimanfaatkan untuk mendukung pengambilan keputusan di berbagai bidang, termasuk pendidikan. Sekolah tidak hanya berperan sebagai lembaga penyelenggara pembelajaran, tetapi juga sebagai pengelola data yang mencakup informasi akademik, administrasi, dan umpan balik dari para pemangku kepentingan, termasuk wali siswa (Kharis & Zili, 2022). Salah satu pendekatan yang dapat digunakan untuk menganalisis data tersebut adalah data mining, yaitu proses penggalian informasi bermakna dari kumpulan data besar dengan memanfaatkan metode statistika, pembelajaran mesin, dan algoritma komputas (Galih, 2019).

Algoritma C4.5 merupakan salah satu metode decision tree yang populer karena kemampuannya menghasilkan model klasifikasi yang mudah diinterpretasikan serta efektif dalam menangani data dengan berbagai tipe atribut (Sari, Giwangga, Fitriyanti, Putra, & Irawan, 2024) (Damanik, Wanto, & Gunawan, 2022). Dalam konteks SMK Letris Indonesia 2, penerapan algoritma ini dapat membantu mengukur tingkat kepuasan wali siswa berdasarkan indikator-indikator tertentu, seperti kualitas pembelajaran, fasilitas sekolah, dan komunikasi antara pihak sekolah dengan wali siswa. Analisis ini diharapkan dapat menjadi dasar evaluasi yang objektif bagi pihak sekolah dalam meningkatkan mutu layanan pendidikan.

Pendidikan merupakan upaya mempersiapkan peserta didik melalui orientasi, kegiatan belajar mengajar, penugasan dan kegiatan aktivitas pendidikan lainnya yang dapat mempengaruhi sikap dan perilaku untuk meningkatkan mutu belajar siswa. Pendidikan merupakan hal terpenting bagi setiap manusia. Pendidikan melibatkan kegiatan yang kompleks, mempunyai aspek yang luas, dan dipengaruhi oleh banyak variable (Maryani, 2020). Mutu pendidikan suatu negara ditentukan oleh mutu pendidiknya. Pendidik harus memiliki keterampilan dan kualifikasi yang memenuhi standar nasional Pendidikan.

Pada awal tahun pelajaran 2011-2012 kami membuka sebuah sekolah kejuruan yang pertama, dengan diawali kerjasama dengan salah satu pengusaha yang berkantor di gedung Pesona di wilayah Tanah Kusir Kebayoran Baru, Jakarta. Dengan bekerjasama dengan Yayasan yang berniat membuat salah satu sekolah kejuruan dengan biaya sekolah yang terjangkau masyarakat menengah, maka lahirlah sekolah dengan nama SMK LETRIS INDONESIA. Nama SMK LETRIS INDONESIA adalah perwujudan dedikasi yayasan dalam membantu negara dalam dunia pendidikan dengan mencantumkan kata Indonesia pada nama sekolah. Pada awal berdiri kami hanya menempati sebuah ruko di jalan Raya Jombang yang berukuran kecil dan mendapat kepercayaan masyarakat dengan peserta didik yang bergabung di SMK LETRIS INDONESIA hanya 20 siswa saja.

SMK Letris berkomitmen menjalankan lembaga pendidikan tersebut maka pihak yayasan membangun gedung sekolah dengan 4 lantai pada tahap pertama dengan jumlah ruang belajar dan ruang administrasi sebanyak 16 ruangan. Pada tahun kedua tepatnya tahun pelajaran 2012-2013 sudah menempati gedung baru sampai sekarang dan alhamdulilah perkembangan siswa dan sarana prasarana sekolah terus meningkat dari tahun ke tahunnya. Pada akhirnya jumlah siswa kami berjumlah ribuan siswa sampai sekarang. SMK Letris Indonesia 2 merupakan salah satu Yayasan Leo Sutrisno yang beralamat di Jl. Siliwangi No. 55 Pondok Benda, Pamulang, Tangerang Selatan.. Peserta didik SMK Letris Indonesia 2 berjumlah 1434 anak dengan tujuh program keahlian, yaitu: Multimedia, Teknik Jaringan dan Jaringan, Rekayasa Perangkat Lunak, Akutansi, Bisnis Daring dan Pemasaran, Otomatisasi dan Tata Kelola Perkantoran, dan Perbankan Syariah.

Teknologi informasi telah mengalami banyak perkembangan seiring berjalaninya waktu. Teknologi mulai digunakan dalam berbagai bidang kehidupan, termasuk pendidikan (Nento & Manto, 2023). Banyaknya siswa menjadi salah satu alasan penggunaan teknologi komputer. Mengelola data dalam jumlah besar tentunya membutuhkan sumber daya, biaya, dan waktu untuk mengolah data tersebut. Data mining merupakan suatu teknologi untuk mengolah dan mengekstraksi data dalam jumlah yang besar dan beragam. *Data mining* juga merupakan proses analisis untuk menemukan pola dan memprediksi sekumpulan data untuk memperoleh pengetahuan dengan menggunakan teknik pengenalan pola seperti teknik statistik dan matematika. Hasil dari proses data mining ini sangat akurat tergantung dari algoritma yang digunakan. Penggunaan algoritma juga harus disesuaikan dengan jenis data yang perlu (Marisa, Maukar, & Akhriza, 2021). Algoritma C4.5

digunakan dalam *Data Mining* sebagai *Decision Tree* yang dapat digunakan untuk menghasilkan keputusan, berdasarkan *sampel* data tertentu Algoritma C4.5 lebih efektif, hasil ketepatan prediksi dan nilai kesalahan (*error rate*) lebih baik dari algoritma lain Pengukuran kinerja yang dilakukan menggunakan sekelompok data uji untuk mengetahui *presentase precision, recall* dan *accuracy*, menunjukkan bahwa algoritma C4.5 memiliki tingkat akurasi yang lebih tinggi dari pada algoritma lain (Rahmayanti & Rusdiana, 2022).

Pengukuran tingkat kepuasan para wali siswa dapat dilakukan dengan berbagai cara, salah satunya adalah dengan menggunakan metode *data mining*. *Data mining* merupakan proses mengekstraksi informasi yang berguna dari data yang besar. Salah satu metode *data mining* yang dapat digunakan untuk mengukur tingkat kepuasan adalah algoritma C4.5 (Rahmayanti & Rusdiana, 2022).

Data mining adalah metode untuk menemukan informasi baru yang berguna dari kumpulan data yang besar dan dapat membantu dalam pengambilan keputusan Beberapa pandangan mengenai *data mining*, seperti penemuan pengetahuan atau pengenalan pola. Pengenalan pola dipergunakan sebagai pengetahuan yang diekstraksi dari dalam bongkahan data (Girsang, Ginting, & Hustasuhut, 2022).

Algoritma C4.5 adalah salah satu jenis pohon keputusan yang paling populer. Pohon keputusan berguna untuk eksplorasi data, mencari hubungan tersembunyi antara beberapa variabel input kandidat dan variabel target. Algoritma C4.5 dapat menangani data digital dan diskrit menggunakan *ratio gain*. Objek menggunakan konsep *entropy*. Dengan menggunakan pohon keputusan, peneliti tidak perlu memperkirakan distriusi ukuran tinggi badan atau parameter tertentu dari distriusi kelas. Karena metode ini menggunakan kriteria yang lebih sedikit untuk masingmasing dari simpul dalam tanpa mengurangi kualitas keputusan yang dihasilkan secara signifikan (Girsang, Ginting, & Hustasuhut, 2022).

2. METODOLOGI PENELITIAN

Salah satu elemen paling penting yang mepengaruhi seberapa baik suatu organisasi bidang pendidikan beroperasi adalah kepuasan orang tua wali siswa. Studi ini menggunakan pendekatan kuantitatif dengan metode survei untuk menilai sikap siswa terhadap layanan pendidikan yang disediakan oleh SMK Letris Indonesia 2. Pendekatan ini umum digunakan dalam studi pendidikan di Indonesia karena memungkinkan standarisasi hubungan antara kualitas layanan sekolah dan kepuasan individu (Efendi, Hapsari, Wulandari, Inayati, & Rahmi, 2025).

Metode penelitian adalah proses dalam pelaksanaan penelitian secara sistematis. Metode penelitian yang digunakan pada penelitian ini adalah sebagai berikut:

1. Mengidentifikasi Masalah

Langkah pertama melibatkan analisis konteks sekolah untuk mengidentifikasi masalah, indikator, dan variabel yang relevan. Fokus penelitian ini adalah pada tingkat kepuasan orang tua/wali siswa serta faktor-faktor yang mungkin mempengaruhinya (seperti kualitas pendidikan, fasilitas, dan komunikasi antara sekolah dan orang tua). Hal ini sejalan dengan keyakinan lokal bahwa layanan sekolah (empati, bukti fisik, jaminan, daya tanggap, dan keandalan) berkorelasi dengan kebutuhan siswa. (Efendi, Hapsari, Wulandari, Inayati, & Rahmi, 2025)

2. Pengumpulan data

Data primer diperoleh melalui kuesioner daring. Tautan kuesioner diberikan kepada kepala sekolah untuk diberikan kepada siswa di setiap tingkatan, setelah itu diberikan kepada orang tua atau wali siswa. Metode survei yang berani dipilih karena efektif untuk populasi besar, memfasilitasi pengumpulan respons, dan selaras dengan kebutuhan praktis layanan pendidikan di Indonesia (Kurnia & Ningtias, 2025).

3. Sumber Data dan Variabel Penelitian

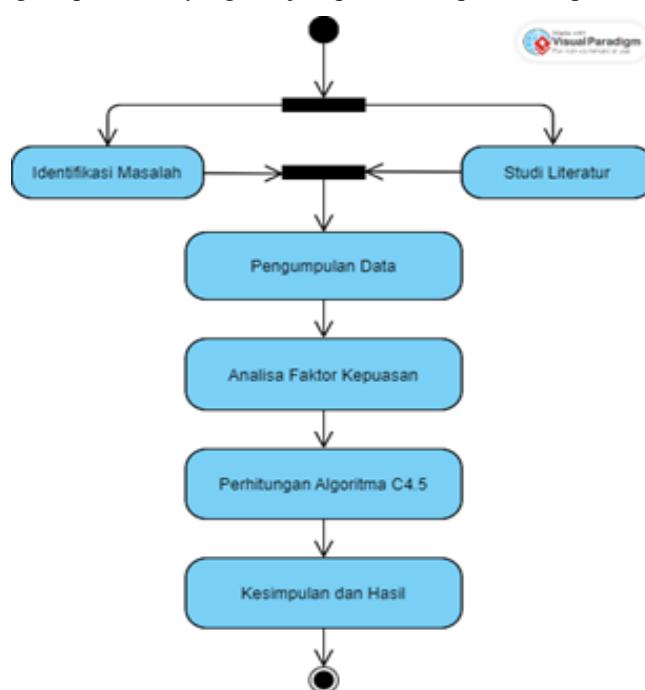
Responden adalah orang tua atau wali siswa dari SMK Letris Indonesia 2. Variabel dependen adalah perilaku orang tua atau wali siswa, sedangkan variabel independen mencakup indikator kualitas layanan. Pengembangan indikator ini didasarkan pada pengetahuan tentang faktor-faktor yang memengaruhi efektivitas layanan pendidikan (Efendi, Hapsari, Wulandari, Inayati, & Rahmi, 2025) (Kurnia & Ningtias, 2025).

Langkah – Langkah Perhitungan Data Mining Algoritma C4.5

- a. Langkah pertama pada penelitian ini adalah menghitung nilai *entropy* total dari seluruh data menggunakan persamaan (2.1).
- b. Kemudian menghitung nilai *entropy* dari setiap nilai pada atribut menggunakan persamaan (2.1).
- c. Setelah nilai *entropy* dihitung, maka hitung nilai gain dari setiap atribut menggunakan persamaan (2.2).
- d. Membuat akar *decision tree* menggunakan atribut yang memiliki nilai *gain* tertinggi.
- e. Atribut yang memiliki nilai *gain* tertinggi kemudian dihitung nilai *entropy* totalnya. Misalnya pada atribut pelayanan admin, terdapat tiga nilai yaitu baik, cukup baik dan kurang baik. Maka dihitung nilai *entropy* total untuk kasus baik, cukup baik dan kurang baik menggunakan persamaan (2.3).
- f. Setelah itu dihitung nilai *gain* dari setiap atribut dengan kasus baik, cukup baik dan kurang baik menggunakan persamaan (2.3). Maka akan didapatkan nilai gain tertinggi dari kasus baik, cukup baik, dan kurang baik dari perhitungan nilai gain tersebut yang kemudian dijadikan sebagai akar simpul pertama pada *decision tree*.
- g. Melakukan proses pada poin ke-5 sampai semua atribut masuk pada *decision tree*.
- h. Setelah *decision tree* terbentuk, maka dilakukan evaluasi akurasi untuk mengetahui keakuratan algoritma C4.5 menggunakan *confusion matrix* dan rumus *accuracy* pada persamaan (2.4).
- i. Menguji *k-fold cross validation* untuk mengetahui akurasi dari keseluruhan proses *data mining*. Pengujian *k-fold cross validation* dilakukan dengan membagi data testing menjadi beberapa bagian.
- j. Analisis hasil dari *decision tree* yang sudah terbentuk dan uji akurasi pada *confusion matrix*.
- k. Kesimpulan dan Hasil

Setelah melakukan semua tahapan, selanjutnya peneliti melakukan penarikan kesimpulan dari hasil yang didapatkan pada penelitian.

Langkah-langkah penelitian yang menjadi panduan dapat dilihat pada bagan berikut.



Gambar 1. Activity Diagram Metode Penelitian

Algoritma C4.5 merupakan algoritma yang digunakan untuk melakukan klasifikasi data dengan menggunakan teknik keputusan. Algoritma C4.5 merupakan eksistensi dari algoritma ID3

dan menggunakan prinsip *decision tree* yang mirip. Algoritma C4.5 merupakan program yang memberikan kontribusi satu set data berlabel dan menghasilkan pohon keputusan sebagai keluaran. Algoritma C4.5 membangun pohon keputusan dari serangkaian data pelatihan yang mirip dengan Algoritma ID3, dengan menggunakan konsep *entropy* informasi. Algoritma ini memiliki input berupa training samples yang merupakan data contoh yang digunakan untuk membangun sebuah *tree* yang telah diuji kebenarannya dan samples merupakan *field-field* data yang nantinya akan digunakan sebagai parameter dalam melakukan klasifikasi data.

Rumusan Algoritma C4.5 terbagi menjadi 4 rumus. Untuk menghitung *Entropy*, *Information gain*, *split info* dan *gain ratio* digunakan seperti tertera penjelasan dari pemodelan algoritma C4.5 berikut:

Perhitungan *entropy*

$$Entropy(S) = \sum_{i=1}^n -p_i * \log_2 p_i \quad (2.1)$$

Keterangan :

- S : Himpunan kasus
N : Jumlah partisi S
Pi : Proporsi dari Si terhadap S

Perhitungan nilai informasi *gain*

$$Gain(G, A) = Entropy(S) - \sum_{i=1}^n \frac{|S_i|}{|S|} * Entropy(S_i) \quad (2.2)$$

Keterangan :

- S : Himpunan Kasus
A : Atribut
N : Jumlah kasus atribut A
|Si| : Jumlah kasus pada partisi ke – i
|S| : Jumlah kasus dalam S

Perhitungan nilai *splitinfo*

$$SplitInfo_A(D) = \sum_{j=1}^D \frac{|D_j|}{|D|} * \log_2 \left(\frac{|D_j|}{|D|} \right) \quad (2.3)$$

Keterangan :

- D : Ruang Data *Sample Training*
Dj : Sample Pada Atribut J

Perhitungan nilai *gainratio*

$$GainRatio = \frac{Gain(S, A)}{SplitInfo_A(D)} \quad (2.4)$$

3. ANALISA DAN PEMBAHASAN

Jenis penelitian pada penelitian ini merupakan penelitian kuantitatif. Pada penelitian ini akan dilakukan analisis terhadap hasil berupa *decision tree* yang diperoleh dari subjek data yaitu data kepuasan para wali siswa dari bulan Januari sampai pada bulan Mei 2024. Data yang digunakan adalah data kepuasan wali siswa di google form yang didapat secara online, data yang didapatkan berupa data diskrit berdasarkan setiap atribut yang ditentukan. Faktor-faktor yang digunakan dalam penelitian ini adalah pelayanan admin, fasilitas gedung, kualitas guru dan kepuasan para wali siswa kemudian, kelas atau output yang telah ditentukan dalam penelitian ini yaitu kepuasan para wali siswa terhadap sekolah SMK Letris Indonesia 2 dengan puas dan tidak puas memperoleh hasil pengujian yang valid dengan hasil perhitungan sebelumnya yang dilakukan di Microsoft Excel.

3.1 Sumber Data dan Variabel Penelitian

Pada tahap ini dilakukan dengan melakukan penyebaran kuesioner kepada responden wali siswa SMK Letris Indonesia 2.

Tabel 1. Data Kuesioner

No	Nama	Pelayanan Admin	Fasilitas Gedung	Kualitas Guru	Tingkat Kepuasan	49	Fadli	Baik	Baik	Kurang Baik	Tidak Puas
						50	Rani	Baik	Baik	Kurang Baik	Tidak Puas
1	Budi Wijayanto	Cukup Baik	Baik	Baik	Puas	51	Fajar	Baik	Baik	Kurang Baik	Puas
2	Santi Indriani	Cukup Baik	Cukup Baik	Baik	Puas	52	Lia	Baik	Baik	Kurang Baik	Tidak Puas
3	Andi Kumiawan	Cukup Baik	Cukup Baik	Cukup Baik	Tidak Puas	53	Dedi	Cukup Baik	Cukup Baik	Kurang Baik	Puas
4	Rina Sulistyowati	Baik	Baik	Baik	Puas	54	Rizky	Kurang Baik	Kurang Baik	Baik	Puas
5	Adit Pratama	Kurang Baik	Cukup Baik	Cukup Baik	Tidak Puas	55	Arief	Kurang Baik	Kurang Baik	Baik	Puas
6	Dewi Astuti	Baik	Cukup Baik	Baik	Puas	56	Agus	Kurang Baik	Kurang Baik	Cukup Baik	Puas
7	Hendra Purnama	Baik	Baik	Baik	Puas	57	Tri Astuti	Baik	Baik	Baik	Puas
8	Rizal Jaelani	Kurang Baik	Kurang Baik	Kurang Baik	Tidak Puas	58	Irma	Baik	Baik	Baik	Puas
9	Ajang	Cukup Baik	Baik	Cukup Baik	Puas	59	Dewi	Kurang Baik	Kurang Baik	Baik	Puas
10	Dadang	Baik	Baik	Baik	Puas	60	Agus	Baik	Baik	Kurang Baik	Tidak Puas
11	Asep	Cukup Baik	Kurang Baik	Cukup Baik	Tidak Puas	61	Agus	Kurang Baik	Baik	Kurang Baik	Tidak Puas
12	Suganda	Baik	Kurang Baik	Baik	Tidak Puas	62	Agus	Kurang Baik	Baik	Kurang Baik	Puas
13	Suhendra	Cukup Baik	Kurang Baik	Baik	Tidak Puas	63	Nurlilah	Kurang Baik	Baik	Kurang Baik	Puas
14	Sujioni	Kurang Baik	Baik	Baik	Tidak Puas	64	Ahmad	Kurang Baik	Kurang Baik	Baik	Puas
15	Supriyatna	Baik	Cukup Baik	Baik	Puas	65	Hendrik	Kurang Baik	Kurang Baik	Cukup Baik	Puas
16	Surisna	Kurang Baik	Baik	Baik	Tidak Puas	66	Ratna	Cukup Baik	Cukup Baik	Cukup Baik	Tidak Puas
17	Surawijaya	Cukup Baik	Cukup Baik	Baik	Puas	67	Budianwan	Cukup Baik	Cukup Baik	Kurang Baik	Tidak Puas
18	Junaedi	Kurang Baik	Cukup Baik	Baik	Puas	68	Tania	Cukup Baik	Cukup Baik	Baik	Puas
19	Maman	Baik	Baik	Baik	Puas	69	Rahmat	Kurang Baik	Baik	Kurang Baik	Puas
20	Memed	Baik	Baik	Baik	Puas	70	Diana	Baik	Baik	Baik	Puas
21	andi	Baik	Baik	Baik	Puas	71	Ronald	Baik	Baik	Baik	Tidak Puas
22	siska	Baik	Baik	Baik	Puas	72	Rini	Baik	Baik	Kurang Baik	Tidak Puas
23	Rani	Baik	Baik	Baik	Tidak Puas	73	Jefri	Baik	Baik	Baik	Puas
24	Sari	Baik	Baik	Baik	Puas	74	Nina	Baik	Baik	Baik	Tidak Puas
25	Joko Purnama	Baik	Baik	Baik	Tidak Puas	75	Bima	Baik	Baik	Baik	Puas
26	Maya	Cukup Baik	Cukup Baik	Cukup Baik	Puas	76	Bintang	Baik	Baik	Baik	Puas
27	Fajar	Kurang Baik	Baik	Cukup Baik	Puas	77	Indah	Baik	Baik	Baik	Puas
28	Nina Rahayu	Baik	Baik	Baik	Tidak Puas	78	Yoga	Cukup Baik	Cukup Baik	Cukup Baik	Puas
29	Irawan	Cukup Baik	Cukup Baik	Cukup Baik	Puas	79	Denny	Cukup Baik	Cukup Baik	Kurang Baik	Puas
30	Dewi	Baik	Kurang Baik	Cukup Baik	Puas	80	Jodi	Baik	Baik	Baik	Puas
31	Rudi	Baik	Kurang Baik	Cukup Baik	Tidak Puas	81	Nayla	Cukup Baik	Cukup Baik	Cukup Baik	Puas
32	Lissa	Baik	Kurang Baik	Baik	Tidak Puas	82	Bram	Kurang Baik	Kurang Baik	Cukup Baik	Puas
33	Sinta	Baik	Kurang Baik	Cukup Baik	Puas	83	Erwin	Baik	Baik	Baik	Puas
34	Bayu	Baik	Cukup Baik	Baik	Tidak Puas	84	Rafi	Baik	Baik	Baik	Puas
35	Rina	Baik	Kurang Baik	Baik	Tidak Puas	85	Novi	Baik	Baik	Baik	Puas
36	Fitri	Baik	Kurang Baik	Cukup Baik	Tidak Puas	86	Rudi	Baik	Baik	Baik	Puas
37	Siti	Cukup Baik	Kurang Baik	Cukup Baik	Puas	87	Santi	Baik	Baik	Baik	Puas
38	Lisa Pertiwi	Cukup Baik	Kurang Baik	Cukup Baik	Tidak Puas	88	Dimas	Baik	Baik	Baik	Puas
39	Yuli	Baik	Kurang Baik	Baik	Puas	89	Rian	Baik	Baik	Baik	Puas
40	Adit	Baik	Kurang Baik	Baik	Tidak Puas	90	Tika	Baik	Baik	Baik	Puas
41	Anton	Baik	Baik	Baik	Tidak Puas	91	Johan	Baik	Baik	Baik	Puas
42	Angga	Baik	Baik	Baik	Puas	92	Rizal	Baik	Baik	Baik	Puas
43	Tri Rahayu	Baik	Baik	Baik	Puas	93	Reza	Baik	Baik	Baik	Puas
44	Rina	Baik	Baik	Baik	Puas	94	Lala	Baik	Baik	Baik	Puas
45	Fahri	Baik	Kurang Baik	Cukup Baik	Puas	95	Ara	Baik	Baik	Baik	Puas
46	Dewi	Baik	Baik	Cukup Baik	Tidak Puas	96	Devi	Baik	Baik	Baik	Puas
47	Lina	Baik	Baik	Baik	Puas	97	Fikri	Baik	Baik	Baik	Puas
48	Nina	Baik	Baik	Baik	Puas	98	Galih	Cukup Baik	Baik	Cukup Baik	Puas
						99	Tari	Kurang Baik	Baik	Kurang Baik	Puas
						100	Nindy	Cukup Baik	Baik	Kurang Baik	Puas
						101	Vino	Baik	Baik	Kurang Baik	Puas
						102	Febrianto	Baik	Baik	Baik	Puas
						103	Bagus	Baik	Baik	Baik	Puas
						104	Silvia	Baik	Baik	Kurang Baik	Puas
						105	Lola	Baik	Baik	Baik	Puas

Penelitian ini menggunakan variabel independent (bebas) dengan tiga attribut yaitu pelayanan admin, fasilitas gedung dan kualitas guru. Dan variabel dependent (terikat) yaitu label/kelas output pada penelitian ini yaitu kepuasan para wali siswa berdasarkan pilihan yang dipilihnya.

3.2 Perhitungan Rumus Algoritma C4.5

Langkah 1: Menyiapkan data lalu menghitung jumlah kasus, jumlah kasus untuk kepuasan puas, dan jumlah kasus untuk tidak puas

Langkah 2: Menghitung *Entropy* dari semua hasil penelitian dan hasil penelitian dibagi berdasarkan kelas attribut. Selanjutnya dilakukan perhitungan gain untuk masing-masing attribut. Berikut adalah perhitungan nilai *entropy*, *gain*, *SplitInfoA*, *GainRatio*.

Tabel 2. Total dari Label

Jumlah	Puas	Tidak Puas
105	70	35

$$Entropy [Total] = \left(\left(-\frac{70}{105} \right) \times \log 2 \left(\frac{70}{105} \right) + \left(-\frac{35}{105} \right) \times \log 2 \left(\frac{35}{105} \right) \right) = 0,918295834$$

Langkah 3: Setelah mendapatkan nilai entropy total maka kita menghitung entropy pada setiap attribut, lalu setelah menghitung entropy pada setiap attribut maka kita bisa menghitung gain .

Tabel 3. Total dari Semua Attribut

	Jumlah (S)	Puas (Si)	Tidak Puas (Si)
Total	105	70	35
Pelayanan Admin			
Baik	67	48	19
Cukup Baik	20	14	6
Kurang Baik	18	13	5
Fasilitas Gedung			
Baik	66	51	15
Cukup Baik	17	12	5
Kurang Baik	22	12	10
Kualitas Guru			
Baik	65	51	14
Cukup Baik	22	14	8
Kurang Baik	18	10	8

Entropy Atribut Pelayanan admin (Baik)

$$= \left(\left(-\frac{48}{67} \right) \times \log 2 \left(\frac{48}{67} \right) + \left(-\frac{19}{67} \right) \times \log 2 \left(\frac{19}{67} \right) \right) = 0,860285865$$

Entropy Atribut Pelayanan admin (Cukup Baik)

$$= \left(\left(-\frac{14}{20} \right) \times \log 2 \left(\frac{14}{20} \right) + \left(-\frac{6}{20} \right) \times \log 2 \left(\frac{6}{20} \right) \right) = 0,881290899$$

Entropy Atribut Pelayanan admin (Kurang Baik)

$$= \left(\left(-\frac{13}{18} \right) \times \log 2 \left(\frac{13}{18} \right) + \left(-\frac{5}{18} \right) \times \log 2 \left(\frac{5}{18} \right) \right) = 0,852405179$$

Tabel 4. Total Entropy pada Semua Attribut dan Entropy Total

	Jumlah (S)	Puas (Si)	Tidak Puas (Si)	Entropy
Total	105	70	35	0,918295834
Pelayanan Admin				
Baik	67	48	19	0,860285865
Cukup Baik	20	14	6	0,881290899
Kurang Baik	18	13	5	0,852405179
Fasilitas Gedung				
Baik	66	51	15	0,773226674
Cukup Baik	17	12	5	0,873981048
Kurang Baik	22	12	10	0,994030211
Kualitas Guru				
Baik	65	51	14	0,751649946
Cukup Baik	22	14	8	0,945660305
Kurang Baik	18	10	8	0,99107606

Langkah 4: Setelah semua entropy pada atribut dihitung lalu masuk ke perhitungan selanjutnya mencari information gain pada Atribut .

$$\text{Information Gain pada pelayanan admin} = (0,918295834) - \left(\left(\frac{67}{105} \times 0,860285865 \right) + \left(\frac{20}{105} \times 0,881290899 \right) + \left(\frac{20}{105} \times 0,852405179 \right) \right) = 0,055359985$$

$$\text{Information Gain pada Fasilitas Gedung} = (0,918295834) - \left(\left(\frac{66}{105} \times 0,773226674 \right) + \left(\frac{17}{105} \times 0,873981048 \right) + \left(\frac{22}{105} \times 0,994030211 \right) \right) = 0,082492949$$

$$\text{Information Gain pada Kualitas Guru} = (0,918295834) - \left(\left(\frac{65}{105} \times 0,751649946 \right) + \left(\frac{22}{105} \times 0,945660305 \right) + \left(\frac{18}{105} \times 0,99107606 \right) \right) = 0,084951622$$

Tabel 5. Total Gain dari Semua Attribute

	Jumlah (S)	Puas (Si)	Tidak Puas (Si)	Entropy	Gain
Total	105	70	35	0,918295834	
Pelayanan Admin					0,055359985
Baik	67	48	19	0,860285865	
Cukup Baik	20	14	6	0,881290899	
Kurang Baik	18	13	5	0,852405179	
Fasilitas Gedung					0,082492949
Baik	66	51	15	0,773226674	
Cukup Baik	17	12	5	0,873981048	
Kurang Baik	22	12	10	0,994030211	
Kualitas Guru					0,084951622
Baik	65	51	14	0,751649946	
Cukup Baik	22	14	8	0,945660305	
Kurang Baik	18	10	8	0,99107606	

Langkah 5: Setelah semua entropy pada atribut dihitung lalu masuk ke perhitungan mencari Splitinfo pada Atribut .

Splitinfo pada pelayanan admin (Baik, Cukup Baik, Kurang Baik)

$$= \left(- \left(\frac{67}{105} \right) x \log 2 \left(\frac{67}{105} \right) \right) + \left(- \left(\frac{20}{105} \right) x \log 2 \left(\frac{20}{105} \right) \right) + \left(- \left(\frac{18}{105} \right) x \log 2 \left(\frac{18}{105} \right) \right) \\ = 1,305434206$$

Splitinfo pada Fasilitas Gedung (Baik, Cukup Baik, Kurang Baik)

$$= \left(- \left(\frac{66}{105} \right) x \log 2 \left(\frac{66}{105} \right) \right) + \left(- \left(\frac{17}{105} \right) x \log 2 \left(\frac{17}{105} \right) \right) + \left(- \left(\frac{22}{105} \right) x \log 2 \left(\frac{22}{105} \right) \right) \\ = 1,318775272$$

Splitinfo pada Fasilitas Gedung (Baik, Cukup Baik, Kurang Baik)

$$= \left(- \left(\frac{65}{105} \right) x \log 2 \left(\frac{65}{105} \right) \right) + \left(- \left(\frac{22}{105} \right) x \log 2 \left(\frac{22}{105} \right) \right) + \left(- \left(\frac{18}{105} \right) x \log 2 \left(\frac{18}{105} \right) \right) \\ = 1,336911675$$

Tabel 6. Total Entropy, Gain dan SplitInfo

	Jumlah (S)	Puas (Si)	Tidak Puas (Si)	Entrophy	Gain
Total	105	70	35	0,918295834	
Pelayanan Admin					0,055359985
Baik	67	48	19	0,860285865	
Cukup Baik	20	14	6	0,881290899	
Kurang Baik	18	13	5	0,852405179	
Fasilitas Gedung					0,082492949
Baik	66	51	15	0,773226674	
Cukup Baik	17	12	5	0,873981048	
Kurang Baik	22	12	10	0,994030211	
Kualitas Guru					0,084951622
Baik	65	51	14	0,751649946	
Cukup Baik	22	14	8	0,945660305	
Kurang Baik	18	10	8	0,99107606	

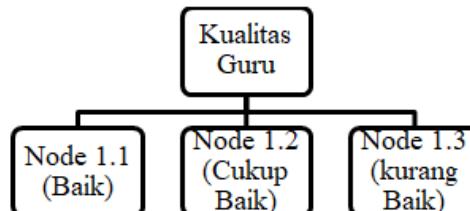
Langkah 6: Setelah semua splitinfo pada atribut dihitung lalu masuk ke perhitungan terakhir mencari GainRatio.

$$\text{Gain Ratio (Pelayanan Admin)} = \frac{0,055359985}{1,305434206} = 0,42407334$$

$$\text{Gain Ratio (Fasilitas Gedung)} = \frac{0,082492949}{1,318775272} = 0,62552696$$

$$\text{Gain Ratio (Kualitas Guru)} = \frac{0,084951622}{1,336911675} = 0,63543182$$

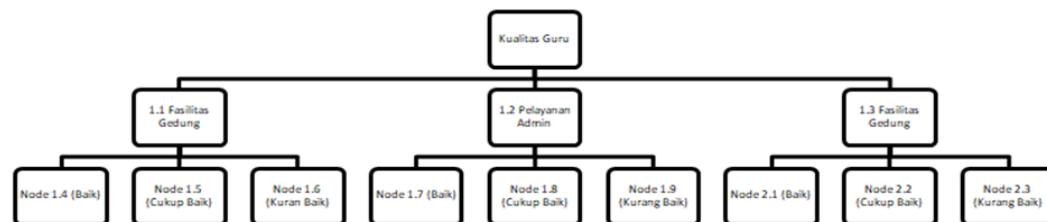
Langkah 7: Dari hasil yang telah diperoleh gain ratio tertinggi adalah kualitas guru maka sudah ditentukan kualitas guru sebagai Node akar



Gambar 2. Kualitas Guru Sebagai Node Akar

3.3 Pengulangan perhitungan pada Node 1.1 sampai Node 1.3

Hasil dari pengulangan perhitungan langkah 1 sampai langkah 7, diperoleh *gain ratio* tertinggi adalah Fasilitas Gedung, Pelayanan Admin, dan Fasilitas Gedung maka sudah ditentukan Kualitas gedung sebagai *Node 1.1*, Pelayanan Admin sebagai *Node 1.2*, dan Fasilitas Gedung sebagai *Node 1.3*.



Gambar 3. Node 1.1 adalah Fasilitas Gedung, Node 1.2 adalah Pelayanan Admin, dan Node 1.3 adalah Fasilitas Gedung.

3.4 Pengulangan perhitungan pada Node 1.4 sampai Node 2.3

Langkah 1: Pengulangan perhitungan pada *Node 1.4* untuk menghitung Kualitas Guru (Baik, Fasilitas Gedung Baik)

Langkah 2: Menghitung *Entropy* dari semua hasil penelitian dan hasil penelitian dibagi berdasarkan kelas atribut. Selanjutnya dilakukan perhitungan *gain* untuk masing – masing atribut. Berikut adalah perhitungan nilai *entropy*, *gain*, *SplitInfoA*, *GainRatio*.

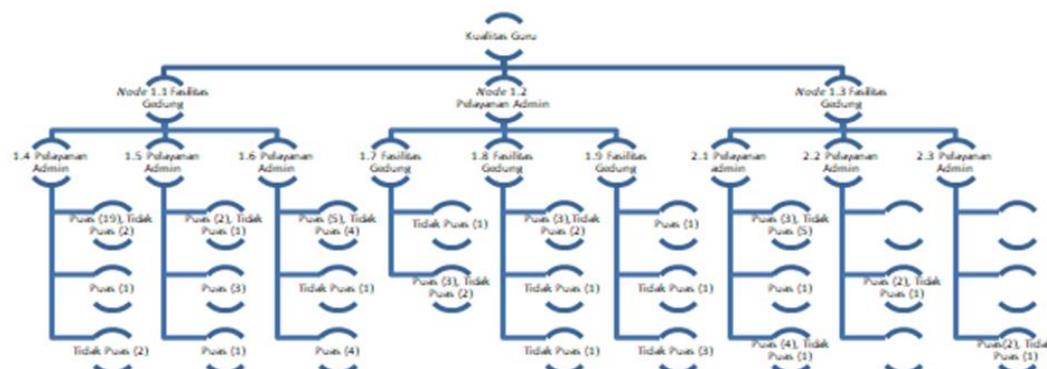
Tabel 7. Total dari Label

Jumlah	Puas	Tidak Puas
48	40	8

Tabel 8. Total Hasil Pencarian Terakhir

	Jumlah (S)	Puas (Si)	Tidak Puas (Si)
Kualitas Guru (Baik) > Fasilitas Gedung (Baik)	48	40	8
Pelayanan Admin			
Baik	45	39	6
Cukup Baik	1	1	0
Kurang Baik	2	0	2

Langkah 3: Karena sudah pencarian terakhir maka tidak perlu menghitung *gain rationya* lagi.



Gambar 4. Hasil Pencarian Terakhir Node 1.4 sampai Node 2.3

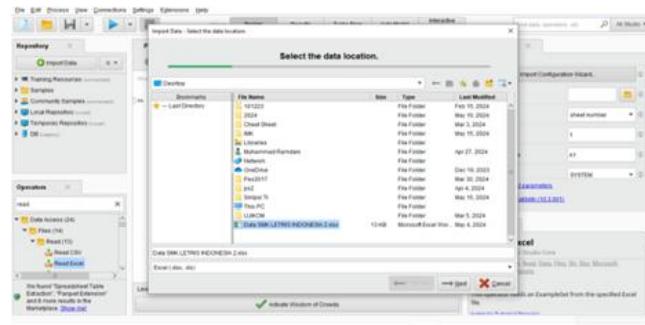
3.5 Penerapan RapidMiner

Tahapan yang dilakukan selanjutnya adalah melakukan proses dari data kuesioner di implementasikan ke *software* RapidMiner. Pada tahap ini juga sebagai pembuktian bahwa perhitungan menggunakan Microsoft Excel secara manual dan menggunakan RapidMiner mempunyai pohon keputusan yang sama.

Proses Perhitungan Algoritma C4.5 menggunakan RapidMiner:

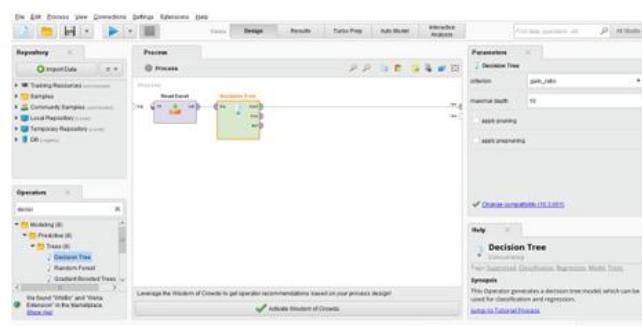
Langkah 1: Menyiapkan data hasil kueioner yang akan dimasukan ke software RapidMiner 10.3.1

Langkah 2: Import data hasil kuesioner dengan pengujian data sebesar 74,29% % yaitu 78 data kedalam software RapidMiner 10.3.1, Seperti gambar dibawah ini.



Gambar 5. Tampilan File yang Akan di Import

Langkah 3: *Drag* dan *drop* *decision tree* pada area *training* ,Sebelum dihubungkan ubah parameter menjadi *gain ratio*, kemudian hubungkan setiap operator, Seperti pada gambar dibawah ini.



Gambar 6. Tampilan Menghubungkan Hasil Kuesioner dan *Decision Tree*

Decision Tree digunakan untuk melihat pohon keputusan pada data yang telah dipilih pada RapidMiner.



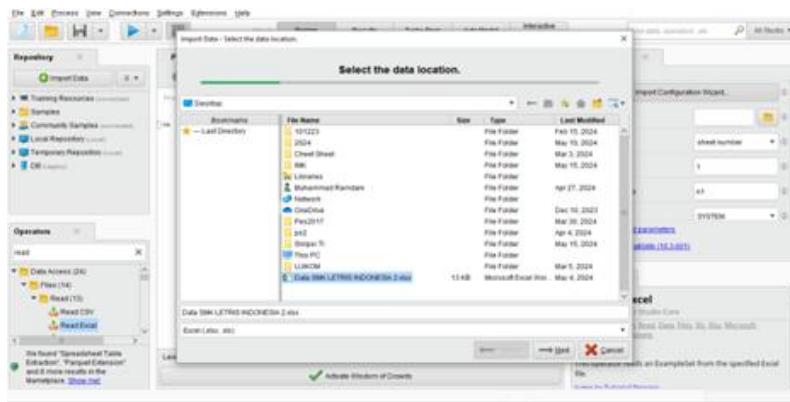
Gambar 7. Tampilan Hasil Pohon Keputusan (*Decision Tree*)

4. IMPLEMENTASI

Pengujian disini yang dilakukan adalah melakukan pengujian dimana untuk menentukan performance accuracy dari data pengujian sebesar 74,29% yaitu 78 dari 105.

Langkah 1: Langkah 1: Menyiapkan data hasil kueioner yang akan dimasukan ke dalam software RapidMiner 10.3.1

Langkah 2: Langkah 2: Import data hasil kuesioner ke software RapidMiner 10.3.1, Seperti gambar dibawah ini.



Gambar 8. Tampilan File Pengujian yang akan di Import

Langkah 3: Drag and drop Multiply pada area training, untuk menduplikasi data yang akan berguna pada perhitungan ini.

Langkah 4: Klik kanan lalu insert building block pilih nominal Cross Validation pada area training, untuk membandingkan tingkat accuracy.

Langkah 5: Sambungkan performance agar bisa menjalankan pengujian dan mengetahui hasil dari tingkat accuracy.



Gambar 9. Tampilan Accuracy Pada

Hasil pengujian dari metode Algoritma C4.5 mendapatkan *rule* sebanyak 18 rule dan mendapatkan tingkat *accuracy* sebanyak 73.08 % lalu untuk predikat tidak puas mendapatkan hasil 56.52%, Predikat Puas mendapatkan hasil 80.00 % dan untuk *true* tidak puas mendapatkan hasil 54.17 %, *true* puas mendapatkan 81.48 %.

1. Accurration

$$= (TP + TN) / (TP + TN + FP + FN) \times 100\%$$

$$= (13 + 44) / (13 + 44 + 11 + 10) \times 100\%$$

$$= 57/78 \times 100\%$$

$$= 73.08\%$$

2. Class Precession	=TP/(TP+FP) 100%	=TP/(TP+FP) 100%													
	=13/(13+10) 100%	=44/(44+11) 100%													
	=13/23 100%	=44/55 100%													
	=56.52%	= 80.00%													
3. Class Recall	=TP/(TP+FN) 100%	=TP/(TP+FN) 100%													
	=13/(13+11) 100%	=44/(44+10) 100%													
	=44/55 100%	=44/54 100%													
	= 54.17%	= 81.48%													
	<table border="1"><thead><tr><th>Uji</th><th>Accuration</th><th>Precision</th><th>Recall</th></tr></thead><tbody><tr><td>Tingkat Akurasi</td><td>73.08%</td><td>56.52%</td><td>54.17%</td></tr><tr><td></td><td>80.00 %.</td><td>80.00 %.</td><td>81.40 %</td></tr></tbody></table>			Uji	Accuration	Precision	Recall	Tingkat Akurasi	73.08%	56.52%	54.17%		80.00 %.	80.00 %.	81.40 %
Uji	Accuration	Precision	Recall												
Tingkat Akurasi	73.08%	56.52%	54.17%												
	80.00 %.	80.00 %.	81.40 %												

Gambar 10. Contoh Tingkat *Accuracy*

5. KESIMPULAN

Dari hasil analisa dan pembahasan pada bab sebelumnya maka dapat disimpulkan sebagai berikut:

1. Faktor kualitas guru menunjukkan pengaruh yang signifikan terhadap puasnya wali murid.
2. Pada penelitian ini menentukan faktor kepuasan wali murid pada SMK Letris Indonesia 2 yang diselesaikan menggunakan metode algoritma C4.5 dengan menggunakan RapidMiner sebagai pengujian mendapatkan hasil dengan 18 rules dan mendapatkan tingkat akurasi 73,08 % lalu untuk predikat tidak puas mendapatkan hasil 56.52%, Predikat Puas mendapatkan hasil 80.00 % dan untuk *true* tidak puas mendapatkan hasil 54.17 %, *true* puas mendapatkan 81.48 %.
3. Berdasarkan perhitungan menggunakan RapidMiner maka disimpulkan bahwa data diatas 70.00% akan sangat berpengaruh dan setelah melakukan pengujian peneliti mendapatkan hasil 73,08%.

Adapun saran untuk pengembangan lebih lanjut terhadap penelitian, sebagai berikut:

1. Pada penelitian ini penulis hanya menggunakan 1 metode saja yaitu algoritma C4.5 yang menghasilkan pohon keputusan (*decision tree*), maka dari itu untuk pengembangan lebih lanjut disarankan menggunakan lebih dari 1 metode teknik data mining agar hasil lebih optimal.
2. Adanya pembandingan software *data mining* lainnya, Karena penelitian ini hanya menggunakan *software data mining* RapidMiner.
3. Adanya perancangan aplikasi dan *database* agar data lebih aman dari yang namanya perubahan yang tidak diinginkan oleh pihak yang tidak bertanggung jawab.

REFERENCES

- Aini, N., & Nasution, M. I. (2024). Akurasi Kualitas Data Informasi pada Sistem Manajemen. *Vol. 2*(No. 1).
- Damanik, S. F., Wanto, A., & Gunawan, I. (2022). Penerapan Algoritma Decision Tree C4.5 untuk Klasifikasi Tingkat Kesejahteraan Keluarga pada Desa Tiga Dolok. *Jurnal Krisnadana*, 1(2), 21-32.
- Efendi, Y., Hapsari, R. F., Wulandari, D. K., Inayati, D. N., & Rahmi, R. (2025). ANALYSIS OF THE CORRELATION BETWEEN THE QUALITY OF SCHOOL SERVICES AND

- PARENTAL SATISFACTION AT SDN SERUA INDAH 02, TANGERANG SELATAN. *Jurnal Cahaya Pendidikan*, 11(1), 32-42.
- Galih, G. (2019). Data Mining di Bidang Pendidikan untuk Analisa Prediksi Kinerja Mahasiswa dengan Komparasi 2 Model Klasifikasi pada STMIK Jabar. *JTSIA: Jurnal Teknologi Sistem Informasi dan Aplikasi*, 2(1), 23-30.
- Girsang, R., Ginting, E. F., & Hustasuhut, M. (2022). Penerapan Algoritma C4.5 Pada Penentuan Penerimaan Program Bantuan Pemerintah Daerah. *Jurnal Sistem Informasi TGD*, 359-449.
- Kementrian Pendidikan dan Kebudayaan. (2020). *Revitalisasi SMK*. Jakarta: Kemendikbud.
- Kharis, S. A., & Zili, A. H. (2022). Learning Analytics dan Educational Data Mining pada Data Pendidikan. *Jurnal Riset Pembelajaran Matematika Sekolah*, 6(1), 12-20.
- Kurnia, S. H., & Ningtias, S. A. (2025, Januari). Evaluasi Kepuasan Pelanggan di Madrasah Aliyah: Studi Tentang Persepsi Orang Tua Terhadap Kualitas Sekolah. *Jurnal Pendidikan Indonesia: Teori, Penelitian dan Inovasi*, 5(1), 155-160.
- Marisa, F., Maukar, A. L., & Akhriza, T. M. (2021). *Data Mining Konsep dan Penerapannya*. Yogyakarta: Deepublish.
- Maryani, E. (2020). *Peningkatan kualitas pendidikan melalui penguatan kompetensi guru*. Bandung: PT Remaja Rosdakarya.
- Nento, F., & Manto, R. (2023). Peran Teknologi Dalam Dunia Pendidikan. *E Tech : Jurnal Ilmiah Teknologi Pendidikan*, Vol. 11(No. 1), 1-5.
- Rahmayanti, A., & Rusdiana, L. (2022, Agustus 4). Perbandingan Metode Algoritma C4.5 dan Naive Bayes Untuk Memprediksi Kelulusan Mahasiswa. *Walisongo Journal of Information Technology*, Vol. 4(No. 1), 11-22.
- Sari, P. W., Giwangga, G., Fitriyanti, N. D., Putra, A. A., & Irawan, H. (2024). Implementasi Algoritma C4.5 Untuk Klasifikasi Data Insentif Karyawan PT. Adhi Cakra Utama Mulia. *Jurnal Komputer Antartika*, 2(3), 95-101.
- Tilaar, H. (2012). *Perubahan sosial dan pendidikan: pengantar pedagogik transformatif untuk Indonesia*. Jakarta: Penerbit Rineka Cipta.