

Prediksi Siswa Putus Sekolah Menggunakan Algoritma *Decision Tree C4.5*

Yofi Lailatul Fatma¹, Naim Rochmawati²,

^{1,2} Teknik Informatika, Jurusan Teknik Informatika, Fakultas Teknik
Universitas Negeri Surabaya

yofi.18039@mhs.unesa.ac.id

naimrochmawati@unesa.ac.id

Abstrak— Pendidikan merupakan salah satu usaha sadar suatu bangsa, yang berguna dalam meningkatkan kualitas atau kemampuan sumber daya manusia (SDM), dengan melakukan aktivitas belajar yang bisa membantu siswa untuk meningkatkan kemampuan yang ada pada dirinya secara maksimal. Peningkatan kualitas pendidikan nasional yang pada saat ini sedang dilakukan oleh pemerintah, masih banyak hambatan yang dilalui di lingkungan. Besarnya jumlah siswa putus sekolah (*dropout*) adalah salah satu masalah utama yang selalu diperhatikan oleh pemerintah. Penelitian ini dimaksudkan untuk memprediksi siswa putus sekolah dengan menerapkan metode *Decision Tree C4.5*. Algoritma *Decision Tree C4.5* bisa menjalankan perkiraan kemungkinan di masa depan dengan menggunakan data yang sudah ada pada sebelumnya dan memperhitungkan sebagian variabel yang akan menjadi penentu hasil akhir suatu keputusan. Hasil dari penggunaan data mining dengan algoritma *Decision Tree C4.5* bisa diterapkan dalam melakukan prediksi jumlah siswa putus sekolah. Dari hasil pengujian algoritma *decision tree C4.5* didapatkan nilai akurasi sebesar 92%, presisi 84% dan *recall* sebesar 50%. Pada sistem menampilkan pohon keputusan dengan hasil yang diperoleh yaitu data siswa putus sekolah lebih sedikit dibandingkan dengan data siswa yang masih aktif (tidak putus sekolah).

Kata Kunci— Pendidikan, Siswa putus sekolah, Prediksi, Algoritma *Decision Tree C4.5*

I. PENDAHULUAN

Pertumbuhan dan eksploitasi ilmu pengetahuan teknologi sekarang ini dapat menciptakan inovasi-inovasi cerdas dan membawa perubahan sosial yang sangat besar. Dengan adanya teknologi saat ini telah mempengaruhi masyarakat dan lingkungan, dimana dengan adanya teknologi dapat membantu berbagai hal, seperti perbaikan ekonomi maupun pendidikan.

Pendidikan merupakan salah satu usaha sadar suatu bangsa yang berguna bagi peningkatan kualitas dan juga kemampuan sumber daya manusia (SDM). Ketidakseimbangan kualitas pendidikan antara daerah perkotaan dengan daerah terpencil, sarana dan prasarana pendidikan yang masih kurang pantas, keterbatasan jumlah guru di daerah terpencil dan masih besarnya jumlah siswa putus sekolah adalah sebagian permasalahan yang dihadapi oleh pendidikan nasional pada saat ini.

Persoalan siswa putus sekolah di Indonesia adalah permasalahan yang multidimensional. Putus sekolah merupakan persoalan pendidikan yang perlu dicari tahu akar penyebabnya. Terdapat berbagai unsur dasar alasan seorang anak bisa putus sekolah, antara lain faktor ekonomi karena orang tua tidak mampu membiayai anaknya untuk bersekolah, dapat juga karena faktor membantu usaha orang tua, ataupun faktor siswanya sendiri yang kurang mengerti akan pentingnya pendidikan, bisa juga karena faktor lingkungan dan teman-teman siswa tersebut [14].

Dalam penelitian ini menggunakan teknik data mining untuk menentukan jumlah siswa putus sekolah, yang kemudian akan dijadikan dasar untuk memprediksi jumlah siswa putus sekolah pada ajaran selanjutnya. Data mining merupakan suatu istilah yang digunakan untuk menemukan pengetahuan yang tersembunyi dalam *database*. Data mining termasuk ke dalam proses yang mempekerjakan satu atau lebih teknik machine learning untuk menganalisis dan mengekstrasikan pengetahuan secara otomatis.

Salah satu metode yang digunakan untuk memprediksi jumlah siswa putus sekolah ini adalah dengan menggunakan metode *Decision Tree C4.5*. Metode ini memiliki *inputan* berupa *training samples* dan *samples*. *Training samples* merupakan contoh data yang diperlukan dalam membangun sebuah pohon keputusan yang sudah diuji kebenarannya. Sementara, *samples* adalah bidang data yang kemudian akan

digunakan sebagai unsur-unsur untuk menjalankan klasifikasi data. Metode C4.5 merupakan salah satu algoritma yang digunakan untuk membuat pohon keputusan yang didasarkan pada *training* data yang sudah disiapkan. Metode *Decision Tree C4.5* adalah penjabaran dari metode ID3. Berikut merupakan sebagian dari penjabaran yang dilakukan algoritma *Decision Tree C4.5* yaitu : dapat mengatasi *missing value*, dapat mengatasi *continue* data, dan *praining* [13].

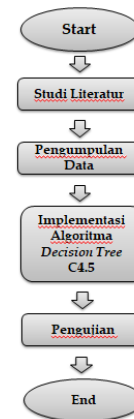
Algoritma *Decision Tree C4.5* terdapat beberapa kelebihan dibandingkan dengan ID3 yaitu, mampu tidak memberikan batasan pada cabang yang berbentuk biner, *Decision Tree C4.5* secara *default* dapat memisahkan cabang pada masing-masing nilai ke dalam atribut kategori. Supaya, klasifikasi yang dihasilkan lebih terkelompokkan jika dibandingkan dengan ekspektasi. Algoritma *Decision tree* mempunyai tingkat akurasi yang lumayan tinggi terhadap berbagai persoalan [6]. Selain itu, metode *Decision Tree C4.5* juga mempunyai kelebihan yang lainnya antara lain : penggabungan data yang tidak sulit ke dalam sistem basis data, bisa menemukan perpaduan data yang tidak disangka, hasil yang sebelumnya rumit dan sangat signifikan bisa dibuat lebih gampang, lebih khusus(kas), dan bisa menghapus hitungan yang tidak digunakan. Disamping itu, Algoritma *Decision Tree C4.5* memiliki kekurangan yakni, keseluruhan akumulasi kesalahan dari setiap tingkatan dalam *decision tree* sangat tinggi, sulitnya merancang *decision tree* yang paling baik, mutu yang didapat sangat bergantung kepada bagaimana *decision tree* itu dirancang [6].

Penelitian *Decision Tree* telah dilakukan oleh beberapa peneliti antara lain, Penerapan Algoritma C4.5 Untuk Penentuan Penerima Beasiswa Mahasiswa [7], Aplikasi Penerapan Algoritma C4.5 Untuk Memprediksi Kelulusan Mahasiswa Berbasis *Web* pernah dilakukan [10], dan Metode *Decision Tree* Algoritma C.45 Dalam Mengklasifikasi Data Penjualan Bisnis Gerai Makanan Cepat Saji ini pernah dilakukan pada sebelumnya [13].

Untuk itu penelitian kali ini, akan mencoba memprediksi siswa putus sekolah dan memanfaatkan metode *Decision Tree C4.5*. Dalam penelitian ini memilih metode *Decision Tree* karena didasarkan pada kelebihan yang dimiliki oleh algoritma *Decision Tree C4.5* yang bisa mengelola berbagai data dan memiliki tingkatan akurasi yang begitu tinggi. Algoritma C4.5 dan *Decision Tree* adalah dua model algoritma yang tidak dapat dipisahkan, oleh sebab itu, guna membuat sebuah pohon keputusan, diperlukan algoritma *Decision Tree C4.5*. Alur dibuatnya penelitian ini berguna untuk memanfaatkan AI (metode *decision tree*) untuk memprediksi perkiraan siswa putus sekolah.

II. METODE PENELITIAN

Untuk melakukan suatu penelitian diperlukan sebuah petunjuk berupa langkah-langkah atau tahapan ilmiah yang terstruktur supaya mendapatkan hasil sesuai dengan yang diinginkan. Alur dalam penelitian ini bisa dilihat pada Gbr. 1 berikut :



Gbr. 1 Alur Penelitian

Dari Gbr. 1 bisa diketahui bahwa dalam melaksanakan suatu penelitian, terdapat beberapa langkah-langkah untuk jalannya penelitian tersebut. Pada setiap langkah akan dijelaskan sebagai berikut :

A. Study Literature

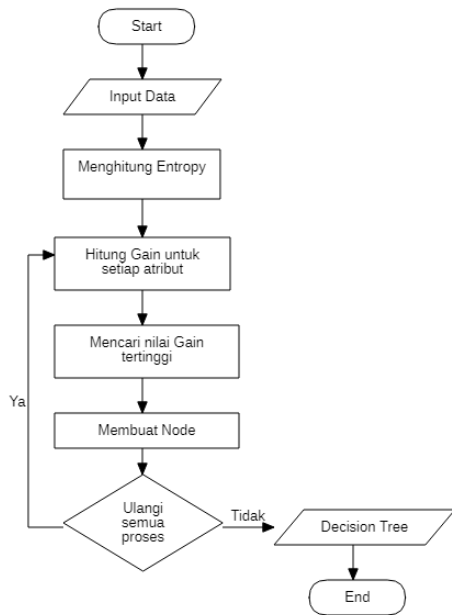
Tahapan yang pertama untuk melaksanakan penelitian ini yaitu *study literature*. Tahapan ini dilakukan dengan cara memahami penelitian-penelitian yang sama dan sudah pernah dilakukan pada sebelumnya agar mendapatkan teori yang mendukung dan pengetahuan yang sesuai dengan penelitian yang dilakukan. *Literature* yang dipelajari yakni mengenai prediksi terhadap jumlah siswa putus sekolah menggunakan algoritma *Decision Tree C4.5*.

B. Pengumpulan Data

Data yang nantinya akan dimanfaatkan pada penelitian ini dibuat dengan menggunakan data umum dan tidak mengacu pada sekolah manapun. Sebelum melakukan klasifikasi, data yang digunakan akan melalui proses *Pre-Processing* terlebih dahulu. Karena data yang didapat masih berupa data tekstual yang tidak terstruktur, maka diperlukan proses pengubahan data menjadi data terstruktur agar mendapatkan data yang siap untuk digunakan. Pada data yang didapatkan terdapat label yang digunakan untuk prosesnya yaitu, label atribut dan label

target. Label atribut pada data siswa yaitu, ID, Jenis Kelamin, Umur, Penghasilan Orang Tua, dan Tanggungan Orang Tua. Untuk label targetnya yaitu Status.

C. Implementasi Algoritma Decision Tree C4.5



Gbr. 2 Flowchart Decision Tree C4.5

Gbr. 2 adalah *flowchart* algoritma *Decision Tree C4.5*, dapat dijelaskan bahwa langkah-langkah metode *Decision Tree C4.5* di bawah ini :

- Pada tahap pertama yaitu input data jumlah siswa putus sekolah dengan atributnya.
- Menghitung nilai *Entropy*. Entropi merupakan informasi yang menyatakan ukuran nilai belum pasti (*impurity*) dari atribut suatu kumpulan data. Nilai entropi dapat dihitung menggunakan rumus (7) berikut :

$$Entropy(X) = \sum_{j=1}^n -p_j * \log_2 p_j \quad (7)$$
 Dimana X merupakan himpunan kasus dengan nilai n adalah jumlah partisi yang ada dan p_j merupakan proporsi kuantitas kasus pada partisi ke-i terhadap kuantitas kasus keseluruhan.
- Menghitung nilai Gain. Gain merupakan ukuran efektifitas suatu atribut yang digunakan untuk klasifikasi data. Nilai gain dihitung pada setiap atributnya menggunakan rumus (8) di bawah ini :

$$Gain(X, A) = Entropy(X) - \sum_{j=1}^n \frac{|X_j|}{|X|} * Entropy(X_j) \quad (8)$$

Dimana A merupakan atribut dari data siswa putus sekolah, X adalah himpunan kasus yang ada

pada data, n adalah kuantitas partisi atribut A, X_i adalah kuantitas kasus pada partisi ke-i.

- Setelah menghitung nilai Gain, langkah selanjutnya yaitu mencari nilai gain tertinggi untuk menentukan akar (*node*).
- Membuat node dari nilai tertinggi yang didapat pada perhitungan gain dan menentukan cabang untuk masing-masing atribut.
- Decision Tree C4.5* dapat direpresentasikan dalam bentuk struktur percabangan, dimana masing-masing cabang yang terbentuk mendeskripsikan *output* dari perolehan nilai yang telah diujikan.

D. Pengujian

Dalam penelitian ini, akan membuat pengujian perhitungan dengan cara membandingkan perolehan perhitungan manual metode *Decision Tree C4.5* dengan perolehan perhitungan pada sistem. Pengujian ini juga menguji perbandingan akurasi perhitungan antara perhitungan manual dan sistem.

E. Analisis Hasil

Analisis hasil adalah langkah yang digunakan untuk menganalisa hasil dari sebuah penelitian. Dalam penelitian ini menerapkan metode *Decision Tree C4.5* untuk menentukan prediksi siswa putus sekolah. Analisis ini juga menganalisa pengujian perhitungan metode *Decision Tree C4.5* yang terdapat pada penelitian ini.

III. HASIL DAN PEMBAHASAN

Hasil yang diperoleh dalam penelitian ini berupa akurasi algoritma *decision tree c4.5* untuk memprediksi siswa putus sekolah yang menggunakan bahasa pemrograman *Python*.

A. Algoritma Decision Tree C4.5

Pada tahap pertama yaitu menentukan *Learning Dataset*. Data siswa yang sudah dikumpulkan, selanjutnya akan dilakukan seleksi kriteria dan diubah menjadi 6 kriteria yang kemudian akan dideskripsikan pada tabel 4.1, yaitu ID, Jenis Kelamin, Umur, Penghasilan Orang Tua, Tanggungan Orang Tua, dan Status berisi keterangan putus sekolah (1) atau tidak (0). *Dataset* yang sudah siap digunakan bisa dilihat pada Tabel I di bawah :

TABEL I
DATA SISWA

ID	Jenis Kelamin	Umur	Penghasilan Orang Tua	Tanggungjawab Orang Tua	Status
1	P	18	8000000	4	0
2	P	16	5000000	2	0
3	L	19	2000000	1	1
4	L	18	4000000	2	0
5	P	16	3500000	2	0
6	L	17	3800000	3	0
7	P	18	3000000	1	0
8	P	18	3300000	2	0

Kemudian langkah berikutnya menghitung nilai entropis dari keseluruhan kasus berdasarkan atribut “Umur”, “Penghasilan Orang Tua”, dan “Tanggungjawab Orang Tua” dengan rumus (13) berikut :

$$Entropy(X) = \sum_{j=1}^n p_j * \log_2 \frac{1}{p_j} = \sum_{j=1}^n -p_j * \log_2 p_j \quad (13)$$

Keterangan :

X = Kumpulan kasus

n = Total partisi X

p_j = Proporsi X_i terhadap X

Sehingga didapat nilai entropisnya sebagai berikut :

$$Entropy(Total) = \left(-\left(\frac{953}{1025}\right) \times \log_2 \left(\frac{953}{1025}\right)\right) + \left(-\left(\frac{72}{1025}\right) \times \log_2 \left(\frac{72}{1025}\right)\right) = 1.297$$

TABEL II
NILAI ENTROPY

Total Kasus	SUM (Tidak Putus)	SUM (Putus)	Entropis Total
1025	953	72	1.297

Tabel II merupakan hasil perhitungan nilai entropi dari keseluruhan atributnya. Untuk selanjutnya, menghitung nilai entropi pada tiap atributnya :

$$Entropy(Penghasilan) = \left(-\left(\frac{762}{820}\right) \times \log_2 \left(\frac{762}{820}\right)\right) + \left(-\left(\frac{58}{820}\right) \times \log_2 \left(\frac{58}{820}\right)\right) = 0.369$$

TABEL III
NILAI ENTROPY PENGHASILAN

Total Kasus	SUM (Tidak Putus)	SUM (Putus)	Entropis Total
820	762	58	0.369

Tabel III adalah tabel perhitungan nilai entropi atribut penghasilan orang tua. Dari entropi penghasilan orang tua selanjutnya, mencari nilai entropi tanggungan orang tua :

$$Entropy(Tanggungjawab) = \left(-\left(\frac{280}{329}\right) \times \log_2 \left(\frac{280}{329}\right)\right) + \left(-\left(\frac{49}{329}\right) \times \log_2 \left(\frac{49}{329}\right)\right) = 0.607$$

TABEL IV
NILAI ENTROPY TANGGUNGAN

Total Kasus	SUM (Tidak Putus)	SUM (Putus)	Entropis Total
329	280	49	0.607

Pada Tabel IV merupakan hasil perhitungan nilai entropi tanggungan orang tua. Tahap berikutnya menghitung nilai Gain pada tiap atributnya menggunakan rumus (14) berikut :

$$Gain = Entropy(X) - \sum_{j=1}^n \frac{|X_i|}{|X|} * Entropy(X_i) \quad (14)$$

Keterangan :

X = Kumpulan kasus

A = Atribut

N = Total partisi atribut A

|X_i| = Total kasus pada partisi ke-i

|X| = Total kasus dalam X

Tabel V adalah hasil nilai gain dari entropi yang telah dihitung sebelumnya yaitu :

TABEL V
PERHITUNGAN GAIN

Atribut	Jumlah	SUM (Total)	SUM (Putus)	SUM (Tidak)	Entropis
Penghasilan	<= 2.950.00 0	820	58	762	0.369
$Gain(Penghasilan) = 1.297 - \left(\left(\frac{820}{1025}\right) * 0.369\right) = 1.0018$					
Tanggungjawab	<= 2.5	329	49	280	0.607

$$Gain(Tanggung) = 1.297 - \left(\left(\frac{329}{1025} * 0.607 \right) \right) = 1.10217$$

Pada Tabel V adalah hasil perhitungan nilai gain yang didapat dari perhitungan *entropy*. Pada Tabel V dapat diketahui nilai gain penghasilan sebesar 1.008 dan nilai gain tanggungan sebesar 1.10217. *Decision Tree* yang dihasilkan dalam perhitungan manual dapat dilihat pada gambar berikut :



Gbr. 3 Pohon Keputusan Hitungan Manual

Gbr. 3 di atas merupakan pohon keputusan hasil akhir yang tercipta dari implementasi metode *Decision Tree* C4.5 dari data *learning* pada Tabel V. Dari gambar tersebut dapat diketahui jika jumlah tanggungan orang tua yang paling berpengaruh terhadap siswa putus sekolah.

B. Hasil Implementasi

```

C:\Windows\System32\cmd.exe
id = 5 depth =2
id = 10 depth =3
id = 9 depth =3
id = 3 depth =2

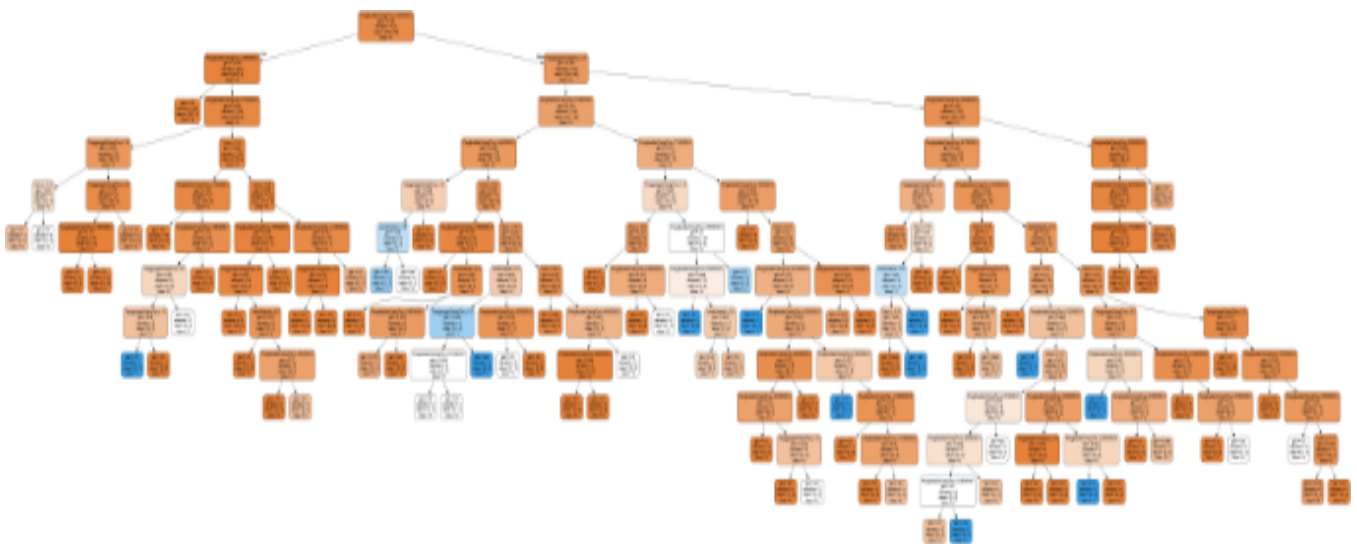
Non-leaf nodes: *****
id = 0 depth =0
id = 2 depth =1
id = 6 depth =2
id = 14 depth =3
id = 1 depth =1
id = 4 depth =2

Tree before pruning with accuracy: 97.0%
Pruned node_id: 2 to achieve accuracy: 97.0%
Pruned node_id: 6 to achieve accuracy: 97.0%
Pruned node_id: 14 to achieve accuracy: 97.0%
Pruned node_id: 1 to achieve accuracy: 98.0%
Pruned node_id: 4 to achieve accuracy: 97.0%

Pruning strategy did'nt increased accuracy
***** Final Tree with accuracy: 97.0% *****
    
```

Gbr. 4 Hasil

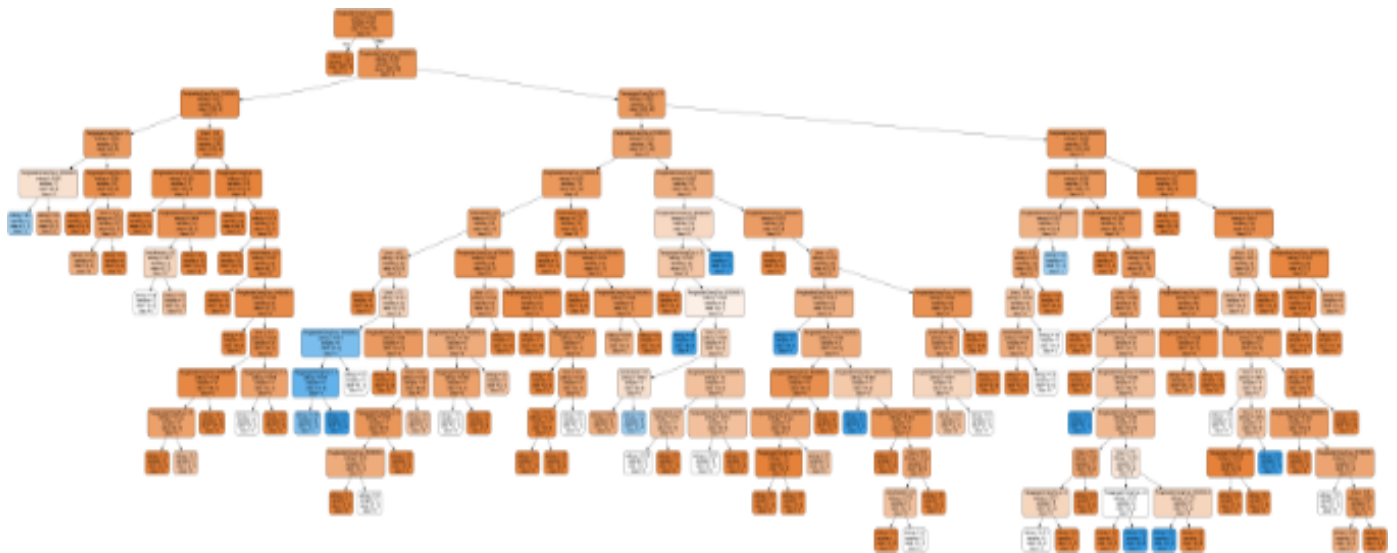
Gbr. 4 adalah hasil dari *running* sistem guna menentukan calon cabang dari data yang di ambil untuk menentukan cabang pada pohon keputusan. Id merupakan atribut pendukung untuk membedakan cabangnya. Sedangkan, *depth* merupakan kedalaman dari cabang tersebut.



Gbr. 5 Pohon keputusan Index Gini

Gbr. 5 merupakan pohon keputusan sebelum dipangkas. Dari hasil tersebut dapat dilihat bahwa penghasilan orang tua kurang dari sama dengan 4.350.000 dengan *index gini* sebesar

0.131 dengan jumlah data latih sebanyak 820. Maka didapat 762 data siswa tidak putus sekolah dan 58 data siswa diprediksi putus sekolah.



Gbr. 6 Pohon Keputusan Akhir

Gbr. 6 merupakan penentuan cabang dari pohon keputusan untuk menghasilkan nilai akurasi. Pada gambar di atas terdapat beberapa nilai prediksi dan masing-masing prediksi memiliki kedalaman yang berbeda. Dari hasil tersebut dapat dilihat bahwa tanggungan orang tua kurang dari sama dengan 2 anak, dengan nilai *entropy* sebesar 0.607 dengan jumlah data latih sebanyak 329 data. Maka didapat 280 data siswa tidak putus sekolah dan 49 data siswa diprediksi putus sekolah.

Accuracy score is 92.20%
 Tree already generate

Gbr. 7 Hasil Akurasi

Gbr. 7 merupakan hasil akurasi yang didapat dari *running* sistem. Dari gambar di atas dapat dilihat hasil akurasinya sebesar 92.20%.

C. Analisis Hasil

Didasarkan dari hasil penelitian yang sudah dilakukan pada data dengan menggunakan perhitungan metode *Decision Tree C4.5* dan manual mendapatkan hasil akurasi sebesar 92.20% dan untuk Hasil Analisa bisa dilihat pada Tabel VI berikut :

a. Uji Coba ke-1

Uji coba yang pertama algoritma *Decision Tree C4.5* menggunakan 1025 data dengan perbandingan 60% data *testing* dan 40% data *training* dari data keseluruhan.

b. Uji Coba ke-2

Uji coba yang kedua algoritma *Decision Tree C4.5* menggunakan 1025 data dengan perbandingan 50% data *testing* dan 50% data *training* dari data keseluruhan.

c. Uji Coba ke-3

Uji coba yang ketiga algoritma *Decision Tree C4.5* menggunakan 1025 data dengan perbandingan 30% data *testing* dan 70% data *training* dari data keseluruhan.

d. Uji Coba ke-4

Uji coba yang keempat algoritma *Decision Tree C4.5* menggunakan 1025 data dengan perbandingan 20% data *testing* dan 80% data *training* dari data keseluruhan.

TABEL VI
 HASIL ANALISA

Percobaan	Accuracy	Recall	Precision
1	86.50%	50%	50.90%
2	87.72%	43.90%	67.92%
3	90.91%	43%	84.31%
4	92.20%	50%	84%

Dari Tabel VI dapat diketahui jika jumlah data *training* bertambah banyak maka, nilai *entropy* yang dihasilkan akan semakin tinggi dan nilai akurasi yang didapat juga tinggi. Dari

data yang sudah ada dan menggunakan jumlah data *training* sebanyak 80% dari jumlah data, akan menghasilkan nilai *entropy* sebesar 0.607 dan tingkat akurasi sebesar 92.20%, *recall* 50%, dan presisi 84%. Bahkan, jika menggunakan data *training* yang sedikit maka, nilai *entropy* yang didapat semakin menurun yaitu menghasilkan nilai sebesar 0.529 dan tingkat akurasi menjadi rendah. Dari data yang sudah ada dan menggunakan jumlah data *training* sebanyak 30% dari jumlah data, akan menghasilkan akurasi sebesar 86.50%, presisi 50%, dan *recall* 50.90%.

Dari hasil analisa yang didapat bisa diketahui jika nilai akurasi yang diperoleh semakin tinggi, maka jumlah siswa putus sekolah semakin sedikit. Begitu pula sebaliknya, jika nilai akurasi yang diperoleh semakin rendah, maka jumlah siswa putus sekolah semakin tinggi.

IV. KESIMPULAN

Didasarkan dari hasil penelitian dan pembahasan bisa dirangkum, bahwa penggunaan *data mining* dengan metode *Decision Tree C4.5* dapat digunakan dalam melakukan prediksi siswa putus sekolah, berdasarkan hasil pengujian perhitungan secara manual didapat nilai *entropy* dan *gain* tertinggi yaitu 0.607 dan 1.10217, algoritma *Decision Tree* mendapat akurasi sebesar 92%, presisi sebesar 84% dan *recall* sebesar 50%. Sedangkan, pada sistem diperoleh hasil pohon keputusan seperti pada Gbr. 6, dimana hasil yang diperoleh yaitu data siswa putus sekolah lebih sedikit dibandingkan dengan data siswa yang masih aktif (tidak putus sekolah).

V. SARAN

Didasarkan dari hasil pengujian terhadap model yang digunakan memiliki hasil yang masih kurang bagus. Maka dari itu, peneliti mempunyai sedikit saran untuk penelitian berikutnya yang berkesinambungan dengan prediksi jumlah siswa putus sekolah, dengan harapan hasil prediksi lebih akurat. Pada penelitian selanjutnya untu bisa menambahkan data lebih banyak lagi, agar mendapatkan hasil akurasi yang lebih tinggi dan akurat. Demi mendapat hasil yang lebih nyata maka, dibutuhkan penataan data sesuai dengan metode yang digunakan. Prediksi jumlah siswa putus sekolah bisa dikembangkan dengan menggunakan algoritma yang lain supaya bisa dibandingkan dengan algoritma mana yang memiliki ketetapan lebih baik.

UCAPAN TERIMAKASIH

Syukur alhamdulillah, penulis ucapkan kepada Allah SWT atas rahmat, ridho, dan hidayah-Nya hingga penulis bisa

menamatkan penelitian ini dengan baik. Tak lupa, sholawat dan salam terus tucurahkan pada Nabi agung Muhammad SAW. yang telah menuntun penulis supaya menjadi manusia yang lebih baik. Terimakasih yang tulus penulis sampaikan pada kedua orang tua serta orang-orang terdekat penulis yang selalu memberi motivasi dalam menyelesaikan tugas yang seharusnya diselesaikan. Saya ucapkan terimakasih yang sebesar-besarnya kepada dosen pembimbing akademik sekaligus dosen pembimbing skripsi saya yaitu Ibu Naim Rochmawati, S.Kom., M.T. yang telah bersedia memberikan bimbingan dan arahan dalam melaksanakan penelitian ini . Tak lupa, saya ucapkan Terimakasih pada rekan-rekan dan seluruh pihak yang mendukung dan memberikan semangat kepada saya hingga penelitian ini selesai. Yang terakhir saya ucapkan Terimakasih kepada diri saya sendiri yang terus semangat dan bisa bertahan sampai tugas akhir ini bisa terselesaikan.

REFERENSI

- [1] V. Matzavela and E. Alepis, "Decision tree learning through a Predictive Model for Student Academic Performance in Intelligent M-Learning environments," *Comput. Educ. Artif. Intell.*, vol. 2, p. 100035, 2021, doi: 10.1016/j.caeai.2021.100035.
- [2] I. D. Mienye, Y. Sun, and Z. Wang, "Prediction performance of improved decision tree-based algorithms: A review," *Procedia Manuf.*, vol. 35, pp. 698–703, 2019, doi: 10.1016/j.promfg.2019.06.011.
- [3] S. Guggari, V. Kadappa, V. Umadevi, and A. Abraham, "Music rhythm tree based partitioning approach to decision tree classifier," *J. King Saud Univ. - Comput. Inf. Sci.*, no. xxxx, 2020, doi: 10.1016/j.jksuci.2020.03.015.
- [4] M. Przybyła-Kasperek and S. Aning, "Stop criterion in building decision trees with bagging method for dispersed data," *Procedia Comput. Sci.*, vol. 192, pp. 3560–3569, 2021, doi: 10.1016/j.procs.2021.09.129.
- [5] F. Osmani and M. Ziaee, "Assessment of the risk factors for vitamin D3 deficiency in chronic hepatitis B patient using the decision tree learning algorithm in Birjand," *Informatics Med. Unlocked*, vol. 23, p. 100519, 2021, doi: 10.1016/j.imu.2021.100519.
- [6] S. Supangat, A. R. Amna, and T. Rahmawati, "Implementasi Decision Tree C4.5 Untuk Menentukan

- Status Berat Badan dan Kebutuhan Energi Pada Anak Usia 7-12 Tahun,” *Teknika*, vol. 7, no. 2, pp. 73–78, 2018, doi: 10.34148/teknika.v7i2.90.
- [7] E. Hasmin and S. Aisa, “Penerapan Algoritma C4.5 Untuk Penentuan Penerima Beasiswa Mahasiswa,” *CogITO Smart J.*, vol. 5, no. 2, p. 308, 2019, doi: 10.31154/cogito.v5i2.219.308-320.
- [8] N. A. Damayanti and R. M. Dewi, “EDUKATIF : JURNAL ILMU PENDIDIKAN Pengembangan Aplikasi Kahoot Sebagai Media Evaluasi Hasil Belajar Siswa,” vol. 3, no. 4, pp. 1647–1659, 2021.
- [9] L. Kusbudiyanto and A. I. Munandar, “Karakteristik Siswa Putus Sekolah Pada Jenjang Sekolah Menengah Kejuruan (Smk) Di Kota Bekasi,” *Sos. Horiz. J. Pendidik. Sos.*, vol. 6, no. 2, p. 298, 2020, doi: 10.31571/sosial.v6i2.1153.
- [10] F. S. Sulaeman and M. A. Rilmansyah, “Aplikasi Penerapan Algoritma C45 untuk Memprediksi Kelulusan Mahasiswa Berbasis Web,” *J. Media Tek. Sist. Ind.*, vol. 5, no. 1, pp. 41–54, 2021.
- [11] J. K. Informatika, “KLASIFIKASI SISWA SMK BERPOTENSI PUTUS SEKOLAH MENGGUNAKAN ALGORITMA DECISION TREE , SUPPORT VECTOR MACHINE DAN NAIVE BAYES,” vol. VII, no. 2, pp. 85–90, 2019.
- [12] D. Alverina, A. R. Chrismanto, and R. G. Santosa, “Perbandingan Akurasi Algoritma C4 . 5 dan CART dalam Memprediksi Kategori Indeks Prestasi Mahasiswa,” vol. 6, no. April, pp. 76–83, 2018, doi: 10.14710/jtsiskom.6.2.2018.76-83.
- [13] E. P. Cynthia and E. Ismanto, “Metode Decision Tree Algoritma C.45 Dalam Mengklasifikasi Data Penjualan Bisnis Gerai Makanan Cepat Saji,” *Jurasik (Jurnal Ris. Sist. Inf. dan Tek. Inform.)*, vol. 3, no. July, p. 1, 2018, doi: 10.30645/jurasik.v3i0.60.
- [14] A. R. Arrahimi, M. K. Ihsan, and D. Kartini, “Teknik Bagging Dan Boosting Pada Algoritma CART Untuk Klasifikasi Masa Studi Mahasiswa,” vol. 5, pp. 21–30, 2019.