

Prosiding Seminar Nasional Informatika Dan Sistem Informasi

P-ISSN 2549-4805

Volume 8, No. 1, Juni-Agus 2024

IMPLEMENTASI METODE DECISION TREE DALAM MENENTUKAN TINGKAT PRESTASI BELAJAR SISWA DI SMP NEGERI 83 JAKARTA

Risah Subariah^{*}, Nurhidayatulloh², Sopiyan Apandi³

¹Sistem Informasi Universitas Pamulang

^{2,3}Teknik Informatika Universitas Pamulang

*e-mail: dosen02695@unpam.ac.id

ABSTRAK: Banyak siswa yang berprestasi tetapi tidak dapat memahami materi yang diajarkan, serta kebutuhan untuk meningkatkan kualitas pendidikan dengan membina siswa berprestasi dan siswa rajin tetapi tidak berprestasi. Agar siswa dapat menangkap konsep pelajaran secara efektif dalam meningkatkan kualitas pendidik sangat penting dalam membina siswa yang berprestasi dan siswa yang rajin tetapi tidak berprestasi. Penelitian ini memanfaatkan metode teknik data mining khususnya Decision Tree dengan Algoritma C4.5 untuk memprediksi prestasi siswa secara akurat. Sehingga dapat mengefektifkan proses pengambilan keputusan bagi dalam menentukan siswa berprestasi. Untuk pengumpulan data dilakukan wawancara dan observasi langsung di SMP 83 Jakarta. Analisis difokuskan pada data training yang dikumpulkan selama tahun ajaran 2020-2021. Hasil dari penelitian ini dengan 100 data didapatkan siswa berprestasi berjumlah 17 siswa dan siswa tidak berprestasi berjumlah 83 siswa diharapkan dapat menunjukkan bahwa penerapan data mining Decision Tree dengan Algoritma C4.5 berhasil memprediksi prestasi siswa dan memberikan pengambilan keputusan metode pembelajaran yang cocok untuk siswa mereka. Penelitian ini bertujuan untuk memberikan kemudahan kepada SMP 38 Jakarta dalam mengambil keputusan untuk metode yang baik dalam menerapkan metode pembelajaran yang cocok antara siswa berprestasi dan yang rajin.

Kata kunci: prediksi; decision tree; algoritma c4.5; siswa; prestasi

ABSTRACT: There are many students who excel but cannot understand the material being taught, and there is a need to improve the quality of education by developing high achieving students and diligent students who do not achieve well. So that students can grasp lesson concepts effectively in improving the quality of educators, it is very important to develop students who excel and students who are diligent but do not achieve. This research utilizes data mining techniques, especially Decision Trees with the C4.5 Algorithm to accurately predict student achievement. So that it can make the decision-making process more effective in determining outstanding students. To collect data, interviews and direct observations were carried out at SMP 83 Jakarta. The analysis focused on training data collected during the 2020-2021 school year. The results of this research with 100 data obtained from 17 students with outstanding achievements and 83 students without achievements are expected to show that the application of Decision Tree data mining with the C4.5 Algorithm is successful in predicting student achievement and providing decisions on learning methods that are suitable for their students. This research aims to make it easier for SMP 38 Jakarta in making decisions about good methods in implementing learning methods that are suitable for high achieving and diligent students.

Keywords: prediction; decision trees; c4.5 algorithm; student; performance

Prosiding Seminar Nasional Informatika Dan Sistem Informasi

P-ISSN 2549-4805

Volume 8, No. 1, Juni-Agus 2024

1. PENDAHULUAN

SMP 83 Jakarta merupakan Sekolah Menengah Pertama yang memiliki siswa-siswa berprestasi. Ada beberapa faktor yang berkontribusi dalam menentukan prestasi siswa, antara lain nilai, kerajinan, dan yang terpenting, kualitas pendidik. Sebagai seorang pendidik, sangatlah penting untuk memberikan pemahaman yang jelas terhadap materi yang diajarkan agar siswa dapat menangkap konsep secara efektif. Oleh karena itu, peningkatan kualitas pendidik sangat penting dalam membina siswa yang berprestasi. Namun, masalah lain muncul ketika siswa yang rajin namun tidak berprestasi. Oleh sebab itu, penting untuk melakukan penelitian memanfaatkan *Decision Tree* dengan Algoritma C4.5 untuk mengidentifikasi pola yang terjadi dalam menentukan prestasi siswa.

Decision tree merupakan metode klasifikasi dan prediksi yang sangat kuat dan terkenal. Metode *decision tree* mengubah fakta yang sangat besar menjadi pohon keputusan yang merepresentasikan aturan. Aturan dapat dengan mudah dipahami dengan Bahasa alami. Proses pada *decision tree* adalah mengubah bentuk data (tabel) menjadi model pohon, mengubah model pohon menjadi rule, dan menyederhanakan rule (karno & said 2019)

Berdasarkan latar belakang yang telah diuraikan, maka penelitian ini berjudul "Implementasi Metode *Decission Tree* Dalam Menentukan Tingkat Prestasi Belajar Siswa Di SMP Negeri 83 Jakarta". Tujuan dari penelitian ini diharapkan dapat memberikan kemudahan bagi pihak sekolah SMP Negeri 83 Jakarta dalam menentukan siswa yang berprestasi dan dapat mengetahui hasil prediksi dengan keakurasiannya.

2. METODA

Jenis Penelitian

Dasar dari penelitian ini bersifat deskriptif, kualitatif, dan kuantitatif. Menurut Natsir pada tahun 1999, deskriptif merupakan metode penelitian yang mengupayakan penggambaran dengan sistematis, objektif dan juga tepat. Penggunaan penelitian ini adalah membantu peneliti mendeskripsikan dengan sesuai sifat-sifat yang telah diperoleh dan berhubungan dengan parameter yang sedang dilakukan penelitian.

Lokasi Penelitian

Kegiatan penelitian ini dilakukan di Sekolah Menengah Pertama (SMP) NEGERI 83 JAKARTA, yang beralamat lengkap di Jl. Empang Bahagia IV-C, Jelambar, Kec. Grogol petamburan, Kota Jakarta Barat, Daerah Khusus Ibukota Jakarta 11460. Penentuan lokasi didasari beberapa hal, diantaranya adalah karena SMP Negeri 83 Jakarta adalah salah satu lembaga pendidikan yang pada setiap tahunnya selalu dipenuhi oleh peserta didik baru yang ingin mendaftar.

Jenis dan Sumber Data

Data yang akan diteliti adalah data primer dan sekunder. Interaksi dengan langsung pada responden dengan wawancara serta bantuan kuisioner secara langsung kepada tenaga pendidik dan peserta didik disebut dengan data primer. Lain

Prosiding Seminar Nasional Informatika Dan Sistem Informasi

P-ISSN 2549-4805

Volume 8, No. 1, Juni-Agus 2024

halnya dengan data sekunder yang didapat dari sumber pustaka atau instansi terkait. Dalam pengisian kuisioner akan ada bobot pada setiap pernyataan tingkat kepentingan pada atribut yang ada mulai dari sangat penting hingga tidak penting. Berikut ini adalah tabelnya:

Tabel 1 Tabel Penilaian Tingkat Kepentingan pada Atribut

Keterangan	Jumlah Bobot
Sangat Penting	5
Penting	4
Cukup	3
Kurang Penting	2
Tidak Penting	1

Bobot tingkat kinerja berikut ini:

Tabel 2 Penilaian Tingkat Kepuasan Pelayanan

Keterangan	Jumlah Bobot
Sangat Puas	5
Puas	4
Cukup	3
Kurang Puas	2
Tidak Puas	1

Efektivitas

Efektivitas adalah tingkat keberhasilan yang dihasilkan seseorang atau organisasi dengan cara tertentu berdasarkan tujuan yang ingin dicapai. Dengan kata lain, semakin banyak rencana diselesaikan, semakin efektif pula kegiatannya. Definisi efektivitas adalah tingkat penyelesaian pekerjaan, sejauh mana orang menghasilkan keluaran yang diharapkan. Artinya jika pekerjaan dapat diselesaikan sesuai rencana, baik waktu, biaya maupun kualitas dapat dikatakan efektif (Siti, Muis, Hastuti, 2021).

Probabilitas observasi

Probabilitas observasi kolektif adalah produk dari probabilitas individu. Keuntungan menggunakan Algoritma C4.5 adalah merupakan algoritma yang sangat populer yang digunakan oleh banyak peneliti di dunia, hal ini dijelaskan oleh Xindong Wu dan Vipin Kumar dalam bukunya yang berjudul The Top Ten Algorithms in Data Mining. Algoritma C4.5 merupakan pengembangan dari algoritma ID3 yang diciptakan oleh J. Rose Quinlan. Secara umum algoritma C4.5 untuk membangun

Prosiding Seminar Nasional Informatika Dan Sistem Informasi

P-ISSN 2549-4805

Volume 8, No. 1, Juni-Agus 2024

pohon keputusan adalah sebagai berikut:

- pilih atribut sebagai akar
- buat cabang untuk tiap-tiap nilai
- bagi kasus dalam cabang
- ulangi proses untuk setiap cabang sampai semua kasus pada cabang

memiliki kelas yang sama

Untuk memilih atribut sebagai akar, didasarkan pada nilai gain tertinggi dari atribut-atribut yang ada. Untuk menghitung gain digunakan persamaan

$$\text{Gain } S, A = \frac{\text{Entropy } S}{n} \quad (1)$$

$$= \sum_{i=0}^n \frac{|S_i|}{|S|} * \text{Entropy } S_i \quad (2)$$

Keterangan :

S = himpunan kasus

A = atribut

n = jumlah partisi atribut A

|S_i| = jumlah kasus pada partisi ke-i

|S| = jumlah kasus dalam S

Sementara itu, perhitungan nilai *entropy* dapat dilihat pada persamaan 2

$$\text{Entropy } S = \sum_{i=0}^n -p * \log_2 p_i \quad (3)$$

Keterangan :

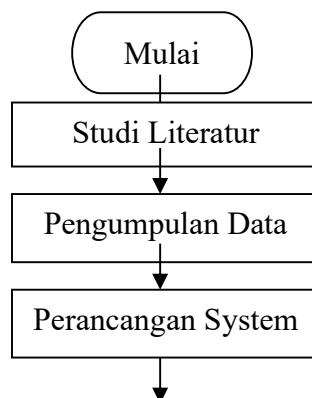
S = himpunan kasus

A = fitur

n = jumlah partisi S

p_i = proporsi dari S_i terhadap S

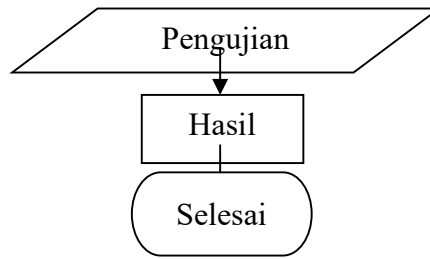
Untuk lebih jelasnya langkah-langkah dalam pembuatan pohon keputusan, berikut diberikan contoh gambar instrument penelitian sebagai berikut:



Prosiding Seminar Nasional Informatika Dan Sistem Informasi

P-ISSN 2549-4805

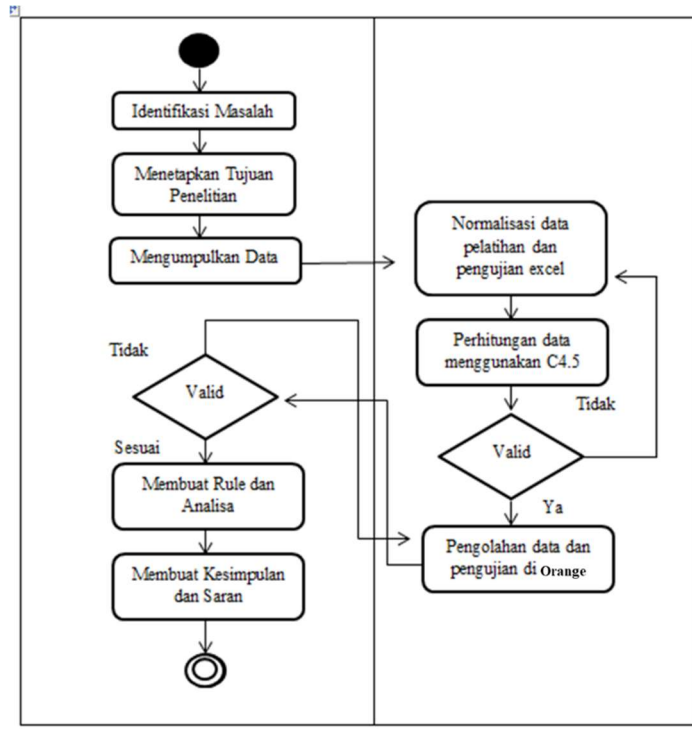
Volume 8, No. 1, Juni-Agus 2024



Gambar 1 Instrumen Penelitian

Proses penelitian diawali dengan peninjauan menyeluruh terhadap literatur yang ada, dilanjutkan dengan pengumpulan data, perancangan sistem, melakukan pengujian, dan menganalisis hasilnya.

1. Diagram Aktifitas Kerja Penelitian



Gambar 2 Diagram Aktifitas Kerja Penelitian

Gambar 2 merupakan diagram aktivitas alur kerja penelitian, pada tahap ini menjelaskan *user* atau penulis melakukan identifikasi masalah selanjutnya menetapkan tujuan penelitian, mengumpulkan data yang akan dijadikan sebagai input sistem. Kemudian dilakukan transformasi data lalu melakukan perhitungan data menggunakan *algoritma C4.5*. Setelah data valid yang selanjutnya dilakukan pengujian di *RapidMiner* Hasil dari informasi yang diperoleh dari *software RapidMiner* diperoleh *rule* keputusan yang selanjutnya di sesuaikan dengan hasil yang diperoleh dari perhitungan manual, jika data sesuai maka penulis membuat *rule* dan analisa

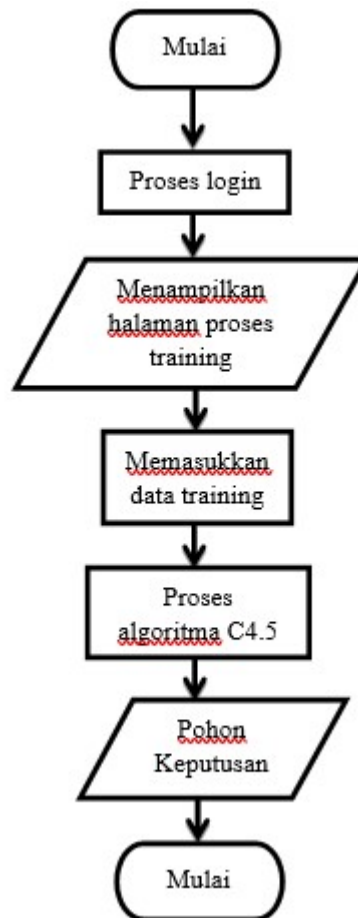
Prosiding Seminar Nasional Informatika Dan Sistem Informasi

P-ISSN 2549-4805

Volume 8, No. 1, Juni-Agus 2024

yang telah dilakukan. Dari hasil yang diperoleh penulis membuat keputusan dan membuat kesimpulan dan saran dari penelitian yang dilakukan.

2 *Pemodelan Metode*



Gambar 3 Pemodelan Metode

Gambar 3 merupakan Kerangka penelitian yang dituangkan dalam diagram alir ini. Menggambar proses penelitian yang akan ditempuh sekaligus menggambarkan penelitian secara keseluruhan.

3. HASIL DAN PEMBAHASAN

Bussiness Understanding atau Pemahaman Bisnis

Tujuan dari penelitian ini adalah memanfaatkan teknik data mining khususnya *Decision Tree* dengan Algoritma C4.5 untuk memprediksi prestasi siswa secara akurat. Sehingga dapat mengefektifkan proses pengambilan keputusan bagi sekolah dalam menentukan siswa berprestasi.

Dataset yang digunakan adalah data yang peneliti ambil secara langsung melalui pengumpulan data di sekolah SMP Negeri 83 Jakarta.

Prosiding Seminar Nasional Informatika Dan Sistem Informasi

P-ISSN 2549-4805

Volume 8, No. 1, Juni-Agus 2024

Tabel 4 Class Label

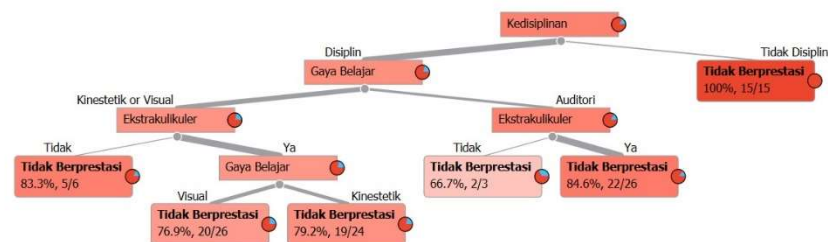
Class Label	Data Training	Data Testing
Prestasi	66%	34%

Data Preparation atau Persiapan Data

Dalam Penelitian ini, data yang digunakan adalah data primer dimana proses pengumpulan data dilakukan melalui proses wawancara dan observasi langsung di SMP Negeri 83 Jakarta. Analisis difokuskan pada data *training* yang dikumpulkan selama tahun ajaran 2020-2021. Atribut-atribut yang diperoleh dari proses pengumpulan data dapat dilihat pada tabel 4.2 dibawah ini:

Penelitian Menggunakan *Decision Tree* Dengan Algoritma C4.5

Pada *decision tree* dengan algoritma C4.5, data record yang sudah di import ke aplikasi Orange digunakan untuk menghasilkan pola pada pohon keputusan. Adapun pola pohon keputusan yang dihasilkan data data yang telah di import sebelumnya sebagai berikut.



Gambar 4 Desain model dan pola pohon Keputusan

Setelah didapatkan pola pohon keputusan berdasarkan data training, maka akan dimodelkan untuk meningkatkan tingkat akurasi.

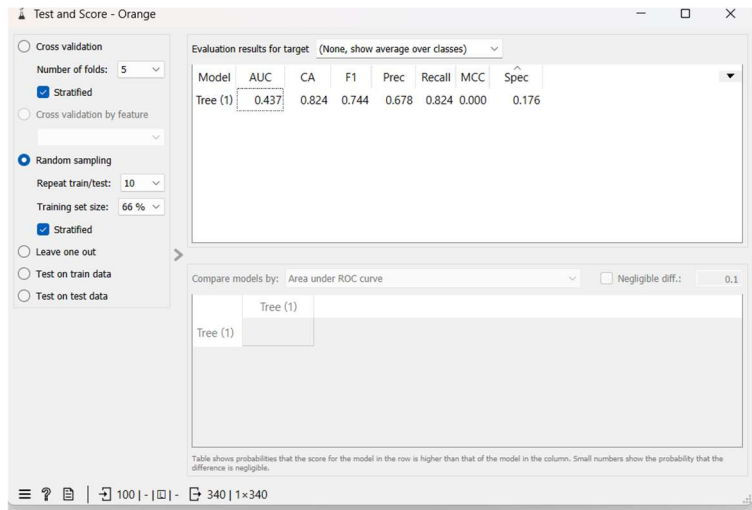
Penelitian Menggunakan *Decision Tree* Dengan Algoritma C4.5

Penerapan data pada orange untuk klasifikasi siswa berprestasi pada SMP Negeri 83 Jakarta menggunakan *decision tree* dengan algoritma C4.5 ditunjukkan pada gambar dibawah ini:

Prosiding Seminar Nasional Informatika Dan Sistem Informasi

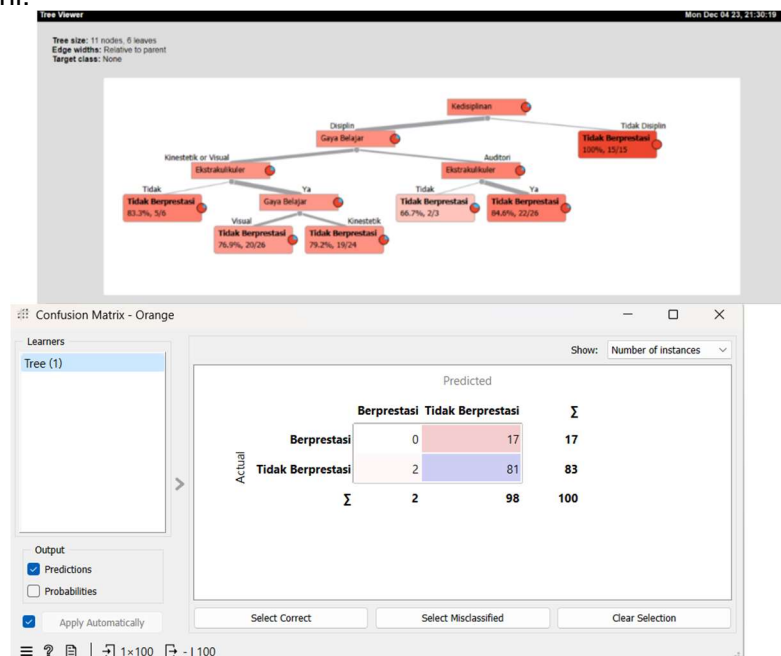
P-ISSN 2549-4805

Volume 8, No. 1, Juni-Agus 2024



Gambar 5 Penerapan data klasifikasi siswa berprestasi menggunakan decision tree dengan algoritma C4.5

Pada gambar 5 data yang sudah disiapkan diterapkan pada aplikasi orange dan data yang digunakan adalah data primer menggunakan algoritma klasifikasi dari *decision tree* dengan algoritma C4.5. Hasil eksperimen dapat kita lihat pada gambar 4.3 dibawah ini:



Gambar 6 Confusion Matrix Hasil Klasifikasi Siswa Berprestasi Menggunakan Algoritma Decision Tree pada Orange

Gambar diatas adalah confusion Matrix yang menunjukkan hasil eksperimen dari *decision tree* dengan algoritma C4.5, didalam confusion matrix dapat kita lihat hasil predictions. Dari 100 data yang ada diperoleh 17 siswa berprestasi dan 83 data siswa

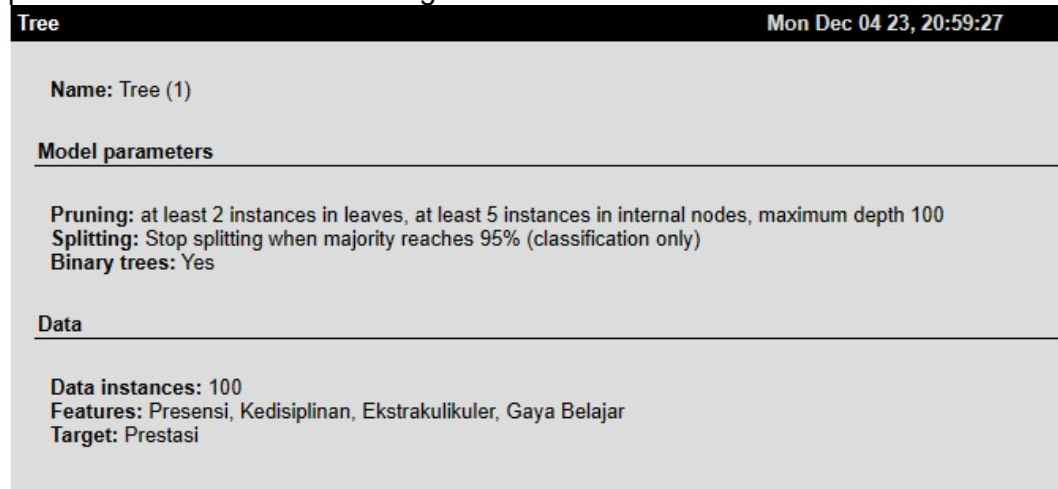
Prosiding Seminar Nasional Informatika Dan Sistem Informasi

P-ISSN 2549-4805

Volume 8, No. 1, Juni-Agus 2024

tidak berprestasi.

Berdasarkan model proses yang telah berhasil dilakukan, maka diperoleh data seleksi pembobotan bisa dilihat melalui gambar 8 berikut ini:



Gambar 7 Model parameters dan Data

Dari Gambar diatas kita bisa lihat terdapat atribut-atribut yang dijadikan acuan untuk pembobotan, terdapat 5 atribut yaitu presensi, ekstrakurikuler, gaya belajar, kedisiplinan, dan prestasi sebagai label. Hasil pengolahan data menggunakan *decision tree* dengan algoritma C4.5 untuk menentukan siswa berprestasi dari total 100 data yang diolah dinyatakan data siswa berprestasi berjumlah 17 siswa dan siswa tidak berprestasi berjumlah 83 siswa berdasarkan hasil pembobotan yang dibuat menggunakan aplikasi orange. Dari Model parameters didapat hasil Splitting berjumlah 95%.

4. SIMPULAN DAN SARAN

Berdasarkan hasil analisa implementasi decision tree dengan algoritma C4.5 dalam memprediksi prestasi siswa, maka dapat disimpulkan sebagai berikut:

- a) Penggunaan model Decision Tree dengan Algoritma C4.5 dapat diterapkan untuk mengidentifikasi siswa berprestasi di SMP 83 Jakarta. Metode ini melibatkan analisis berbagai faktor dan kriteria untuk secara akurat memprediksi siswa mana yang menunjukkan prestasi dan potensi luar biasa di sekolah. Dengan memanfaatkan metode ini, dapat memudahkan pihak sekolah untuk memprediksi siswa berprestasi pada SMP Negeri 83 Jakarta.
- b) Pemanfaatan aplikasi Orange dalam mengimplementasikan Decision Tree dengan Algoritma C4.5 untuk keperluan analisis dan klasifikasi data dapat diterapkan dalam mengidentifikasi siswa berprestasi di SMP 83 Jakarta

DAFTAR PUSTAKA

Adi, Koesdijati T, Utomo Y. (2020). Pengendalian Kualitas Statistik. Surabaya: Scopindo

Prosiding Seminar Nasional Informatika Dan Sistem Informasi

P-ISSN 2549-4805

Volume 8, No. 1, Juni-Agus 2024

- Ardi Ramadhan Sukma, Riqadri Halfis, Ady Hermawan. 2019. "Klasifikasi Channel Youtube Indonesia Menggunakan Algoritma C4.5." *Jurnal Teknik Komputer*, 21–28. <https://doi.org/10.31294/jtk.v4i2>
- Arsyad, Azhar. 2019. *Media Pembelajaran*. Jakarta : PT. Raja Grafindo Persada
- Fatmawati, k., & windarto, a. p. (2018). data mining: penerapan rapidminer dengan k-means cluster pada daerah terjangkau demam berdarah dengue (dbd)berdasarkan provinsi. *computer engineering, science and system journal*, 3(2), 173. <https://doi.org/10.24114/cess.v3i2.9661>
- Helmawati. (2018). *Mendidik Anak Berprestasi Melalui 10 Kecerdasan*. Bandung: PT Remaja Rosdakarya.
- Karno Ganjar Prasetyo dan Said Mirza Pahlevi. 2019. *ANALISIS PERBANDINGAN ALGORITMA DECISION TREE DENGAN SUPPORT VECTOR MACHINE UNTUK MENDETEKSI KOMPETENSI MAHASISWA KONSENTRASI INFORMATIKA KOMPUTER STUDI KASUS : POLITEKNIK LP3I JAKARTA, KAMPUS DEPOK*. Jakarta: Jurnal Lentera ICT
- Rosyid, Moh. Zaiful, dkk. 2019. *Prestasi Belajar*. Jawa Timur : Literasi Nusantara
- Slameto. 2010. *Belajar dan faktor-faktor yang Mempengaruhinya*. Jakarta: PT.Rineka Cipta
- Suntoro, J. (2019). *Data Mining : Algoritma dan Implementasi dengan pemrograman PHP*. Jakarta: PT Elex Media Komputindo
- Susanti, L. (2019). *Prestasi Belajar Akademik & Non Akademik*. Malang: Literasi Nusantara Abadi.
- Waliansyah, R. R., Fitriyah, C., "Perbandingan Akurasi Klasifikasi Citra Kayu Jati Menggunakan Metode Naive Bayes dan k-Nearest Neighbor (k-NN)". *Jurnal Edukasi dan Penelitian Informatika (Jepin)* Vol. 5 no. 2, 2019.