



Penerapan Algoritma C4.5 Untuk Memprediksi Kecocokan Gaya Belajar Bagi Siswa Siswi Sekolah Dasar (Studi Kasus : SD Sariputra Jambi)

Oscario¹, Jasmir², Yudi Novianto³.

^{1,2,3} STIKOM Dinamika Bangsa, Jl. Jendral Sudirman, Thehok , Jambi, 36138, Indonesia.

ABSTRACT

Data Mining was a series of process that used for discovering usefull information from a huge database. SD Sariputra is a primary school was established in 1971. On studying process at SD Sariputra Jambi, there still many students that did not pay attention or have no interest in study. Where the biggest possibility is form the incompatibility of learning style and teaching method . So, the writer to study and analyze the students data and their personality to turn it into usefull information for SD Sariputra. This information will help improve the study process at SD Sariputra. Writer use quisioner result data from grade 3 & 4 with about 155 students data and present it in arff file extention. In doing this analysis, writer use WEKA tools, Decision Tree method and C4.5 Algorithm with 11 attributes. This analyze produce 23 rules. This rules were produced by te final result of Decision Tree shapes.

Keywords: Data Mining, C4.5 Algorithm, Learning Style

ABSTRAK

Data Mining merupakan serangkaian proses yang digunakan untuk mendapatkan informasi yang berguna dari database yang besar. SD Sariputra merupakan sekolah dasar yang didirikan pada tahun 1971. Pada proses pembelajaran di SD Sariputra Jambi, masih banyak murid yang tidak fokus atau tidak memiliki keinginan untuk belajar. Dimana kemungkinan terbesar adalah ketidakcocokan gaya belajar dan metode belajar guru tersebut. Oleh karena itu, penulis mempelajari dan melakukan analisis data mining terhadap data murid dan kepribadiannya agar data tersebut dapat berubah menjadi informasi yang berharga bagi SD Sariputra. Informasi ini dapat membantu meningkatkan proses belajar mengajar di SD Sariputra. Penulis menggunakan hasil data quisioner pada siswa-siswi kelas 3 & 4 sebanyak 155 data siswa-siswi yang telah penulis sajikan dalam file ekstensi arff. Dalam melakukan analisis ini, penulis menggunakan tools WEKA. Metode yang digunakan adalah Decision Tree dengan Algoritma C4.5 dengan 11 atribut. Analisis tersebut menghasilkan 23 Rule. Rule tersebut dihasilkan dari hasil akhir bentuk Decision Tree.

Kata Kunci: Data Mining, Algoritma C4.5, Gaya Belajar

1. PENDAHULUAN

Proses belajar mengajar yang efektif tentunya sangat menentukan mutu serta kualitas dalam bidang pendidikan. Beberapa cara yang dapat digunakan untuk meningkatkan proses belajar mengajar dalam suatu institusi pendidikan, seperti menambah atau meningkatkan fasilitas belajar, peningkatan kualitas tenaga pengajar (guru), dan pembagian kelas yang terjadwal dengan baik. Permasalahan yang sering terjadi dalam pembagian kelas adalah perbedaan gaya belajar pada masing-masing siswa sehingga tenaga pengajar (guru) kesulitan dalam memilih metode yang tepat untuk digunakan[1].

Secara umum gaya belajar manusia dibedakan ke dalam tiga kelompok besar, yaitu gaya belajar visual, gaya belajar auditorial dan gaya belajar kinestetik. Dimana gaya belajar visual adalah gaya belajar dengan cara melihat, mengamati, dan memandang. Gaya belajar auditorial adalah gaya belajar dengan mendengar. Dan gaya belajar kinestetik adalah gaya belajar dengan cara bergerak, bekerja, dan menyentuh[2].

Sariputra adalah salah satu lembaga pendidikan yang memiliki jenjang pendidikan yang lengkap mulai dari TK, SD, SMP, dan SMA. Peningkatan kualitas belajar tentunya harus dimulai di saat peserta didik berada pada jenjang SD (Sekolah Dasar). Pada sistem pembagian kelas di SD Sariputra masih diacak tanpa berdasarkan kesesuaian gaya belajar, sehingga hal tersebut membuat proses belajar mengajar menjadi tidak efektif.

Dari hasil observasi peneliti di SD Sariputra, masih banyak siswa-siswi yang bermain-main atau tidak tertarik saat mengikuti proses belajar mengajar. Hal ini membuktikan bahwa siswa-siswi belum termotivasi dalam kegiatan belajar. Gaya belajar pun juga bisa mempengaruhi pribadi peserta didik dalam belajarnya, dapat dikatakan peserta didik termotivasi secara ilmiah membuat proses belajar peserta didik lebih efektif. Dengan munculnya motivasi, peserta didik akan memiliki keinginan untuk belajar.

Berdasarkan kasus diatas maka penulis tertarik untuk meneliti dan sekaligus membahas kecocokan gaya belajar bagi siswa SD ini dengan menggunakan metode prediksi dalam data mining yaitu algoritma c4.5

2. TINJAUAN PUSTAKA

2.1 Data Mining

Data mining bukanlah suatu bidang yang sama sekali baru. Salah satu kesulitan untuk mendefinisikan *data mining* adalah kenyataan bahwa *data mining* mewarisi banyak aspek dan teknik dari bidang-bidang ilmu yang sudah mapan terlebih dahulu. Berawal dari beberapa disiplin ilmu, *data mining* bertujuan untuk memperbaiki teknik tradisional sehingga bisa menangani Jumlah data yang sangat besar dimensi data yang tinggi dan data yang heterogen dan berbeda sifat. *Data Mining* merupakan proses ataupun kegiatan untuk mengumpulkan data yang berukuran besar kemudian mengekstraksi data tersebut menjadi informasi-informasi yang nantinya dapat digunakan[3].

Beberapa penelitian yang berhubungan dengan metode prediksi dalam data mining seperti yang pernah dibahas oleh Jasmir dkk membahas tentang prediksi mahasiswa dengan KNN [4][5]. Kemudian David Kamangi dkk [6], serta Angga Ginanjar dengan C4.5 [7].

2.2 Decision Tree

Pohon (*tree*) adalah sebuah struktur data yang terdiri dari simpul (*node*) dan rusuk (*edge*). Simpul pada sebuah pohon dibedakan menjadi tiga, yaitu simpul akar (*rootnode*), simpul percabangan/internal (*branch/ internal node*) dan simpul daun (*leaf node*)[8][9]

2.3 Algoritma C4.5

Algoritma C4.5 merupakan algoritma yang digunakan untuk membentuk pohon keputusan. Sedang pohon keputusan dapat diartikan suatu cara untuk memprediksi atau mengklarifikasi yang sangat kuat. Pohon keputusan dapat membagi kumpulan data yang besar menjadi himpunan-himpunan *record* yang lebih kecil dengan menerapkan serangkaian aturan keputusan. Rumus Algoritma C4.5 terbagi menjadi 2 rumus, yang pertama rumus untuk mencari nilai *gain* dan yang kedua adalah rumus untuk mencari nilai *entropy*[10].

Keakuratan prediksi yaitu kemampuan model untuk dapat memprediksi label kelas terhadap data baru atau yang belum diketahui sebelumnya dengan baik. Dalam hal kecepatan atau efisiensi waktu komputasi yang diperlukan untuk membuat dan menggunakan model. Kemampuan model untuk memprediksi dengan benar walaupun data ada nilai dari atribut yang hilang. Dan juga skalabilitas yaitu kemampuan untuk membangun model secara efisien untuk data berjumlah besar (aspek ini akan mendapatkan penekanan). Terakhir interpretabilitas yaitu model yang dihasilkan mudah dipahami.

Dalam algoritma C4.5 untuk membangun pohon keputusan hal pertama yang dilakukan yaitu memilih atribut sebagai akar. Kemudian dibuat cabang untuk tiap nilai didalam akar tersebut. Langkah berikutnya yaitu membagi kasus dalam cabang. Kemudian ulangi proses untuk setiap cabang sampai semua kasus pada cabang memiliki kelas yang sama.

Secara umum algoritma *Decision tree* C4.5 untuk membangun pohon keputusan adalah sebagai berikut (Kusrini, 2009) :

1. Pilih atribut sebagai *root*
2. Buat cabang untuk masing-masing nilai
3. Bagi atribut terpilih dalam cabang
4. Ulangi proses untuk masing-masing cabang sampai semua atribut terpilih pada cabang memiliki kelas yang sama.

Untuk memilih atribut dengan akar, didasarkan pada nilai *gain* tertinggi dari atribut-atribut yang ada. Untuk menghitung *gain* digunakan rumus seperti tertera dalam persamaan 1 berikut:

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

Keterangan:

S : Himpunan kasus

A : Atribut

N : Jumlah partisi atribut A

| Si | : Jumlah kasus pada partisi ke-i

| S | : Jumlah kasus dalam S

Sehingga akan diperoleh nilai *gain* dari atribut yang paling tertinggi. *Gain* adalah salah satu attribute *selection measure* yang digunakan untuk memilih test attribute tiap *node* pada *tree*. Atribut dengan *information gain* tertinggi dipilih sebagai test atribut dari suatu *node*. Sementara itu, penghitungan nilai entropi dapat dilihat pada persamaan 2.

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

Keterangan:

S : Himpunan kasus

A : Atribut

N : Jumlah partisi S

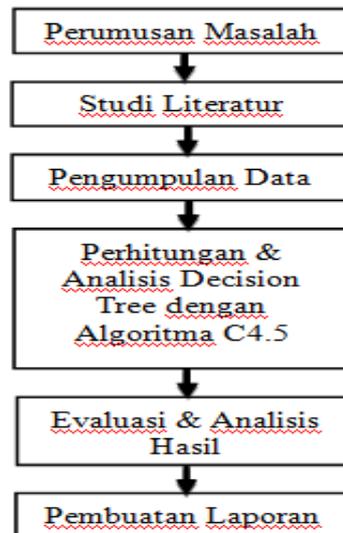
Pi : Proporsi dari Si terhadap S

Kemudian lakukan pengulangan dalam perhitungan *gain* dan *entropy* sampai semua *record* terpartisi. Proses partisi pohon keputusan akan berhenti saat:

- 1) Semua *record* dalam simpul N mendapat kelas yang sama.
- 2) Tidak ada atribut di dalam *record* yang dipartisi lagi.
- 3) Tidak ada *record* di dalam cabang yang kosong.

3. METODOLOGI PENELITIAN

Kerangka kerja penelitian dibuat agar mempermudah pencapaian hasil penelitian, dapat berjalan sesuai dengan yang diharapkan. Adapun kerangka kerja penelitian yang digunakan dapat dilihat pada gambar 1.



Gambar 1. Kerangka kerja penelitian

1. Perumusan Masalah

Pada tahap ini dilakukan peninjauan pada objek penelitian untuk melakukan observasi sehingga menemukan permasalahan yang ada pada SD Sariputra Jambi.

2. Studi Literatur

Pada tahap ini penulis melakukan pencarian terhadap landasan-landasan teori yang diperoleh dari berbagai sumber seperti buku, jurnal ilmiah dan juga referensi lainnya untuk melengkapi penelitian baik mengenai konsep dan teori sehingga memiliki acuan yang baik dan relevan.

3. Pengumpulan Data

Pada tahap ini penulis melakukan pengumpulan data dan informasi yang dibutuhkan dalam penelitian. Pengumpulan data ini dilakukan dengan beberapa metode yaitu:

a. Pengamatan (*Observasi*)

Metode ini dilakukan dengan pengamatan secara langsung pada objek yang akan diteliti. Dalam hal ini penulis mengamati proses belajar mengajar pada Sariputra.

b. Wawancara

Pada metode ini penulis melakukan wawancara singkat secara langsung dengan kepala SD Sariputra untuk memperoleh data dan informasi mengenai data siswa-siswi yang penulis butuhkan.

4. Perhitungan & Analisis Decision Tree dengan Algoritma C4.5

Pada tahapan ini penulis melakukan perhitungan *decision tree* dengan menggunakan algoritma C4.5 dengan menghitung nilai *gain* dan *entropy* dari data-data yang telah dinormalisasi dan digunakan untuk membuat hasil *decision tree*.

5. Hasil Analisis

Pada tahapan ini penulis menganalisis hasil dari perhitungan algoritma C4.5 dan pembuatan *decision tree*. Hasil yang didapat bisa digunakan untuk memprediksi gaya belajar yang cocok untuk masing-masing siswa-siswi SD Sariputra dengan melihat dari hasil *decision tree*.

6. Penulisan Laporan

Pada tahapan ini dilakukan pembuatan laporan yang disusun berdasarkan hasil penelitian dengan menggunakan teknik pengumpulan data primer dan data sekunder, sehingga menjadi laporan penelitian yang sesuai dengan analisis yang sedang diteliti.

4. ANALISIS DAN HASIL

4.1. Representasi Data

Berdasarkan hasil wawancara dan observasi yang sudah dilakukan, penulis memperoleh data-data siswa-siswi di SD Sariputra Jambi. Data yang diinputkan ±150 data quisioner siswa-siswikelas 3 dan kelas 4 SD. Berikut adalah tabel data hasil quisioner siswa-siswi kelas 3 dan kelas 4 SD. Atribut yang akan digunakan pada seluruh data siswa-siswi tersebut yaitu Jenis Kelamin, Hobi, Kelas Favorit, Kemampuan, Bacaan Favorit, Metode Belajar, beberapa pertanyaan hasil kuisioner lainnya dan Gaya Belajar. Berikut adalah tabel data training yang akan digunakan untuk perhitungan Algoritma C4.5.

Tabel 1. Data Siswa Siswi Kelas 3 & 4

No.	JK	Hobi	Kelas Favorit	Kemampuan	Metode Belajar	...	Gaya Belajar
1	L	Musik	Musik	Bernyanyi	Tanya Jawab	...	Auditorial
2	P	Musik	Seni	Menggambar	Membaca Catatan	...	Visual
3	L	Musik	Olahraga	Olahraga	Menulis Ulang	...	Kinestetik
4	L	Olahraga	Olahraga	Menggambar	Tanya Jawab	...	Kinestetik
...
...
153	P	Musik	Seni	Bernyanyi	Tanya Jawab	...	Auditorial
154	P	Seni	Seni	Menggambar	Membaca Catatan	...	Visual
155	P	Olahraga	Olahraga	Menggambar	Membaca Catatan	...	Kinestetik

4.2. Perhitungan Manual Decision Tree

Penulis melakukan perhitungan dengan menggunakan rumus *Gain* dan *Entropy* berikut:

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

Keterangan:

S : Himpunan kasus

A : Atribut

N : Jumlah partisi S

Pi : Proporsi dari Si terhadap S

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

Keterangan:

S : Himpunan kasus

A : Atribut

N : Jumlah partisi atribut A

|Si| : Jumlah kasus pada partisi ke-i

|S| : Jumlah kasus dalam S

Untuk melakukan perhitungan ini, penulis mengambil 15 sample data siswa-siswi yang diambil secara acak yang akan digunakan untuk contoh perhitungan *decision tree* secara manual:

Tabel 2. Data Testing

No.	JK	Hobi	Kelas Favorit	Kemampuan	Metode Belajar	...	Gaya Belajar
1	L	Musik	Seni	Olahraga	Menulis Ulang	...	Kinestetik
3	L	Musik	Musik	Olahraga	Tanya Jawab	...	Auditorial
3	L	Seni	Olahraga	Bernyanyi	Menulis Ulang	...	Auditorial
6	L	Musik	Musik	Menggambar	Membaca Catatan	...	Visual
9	L	Musik	Seni	Bernyanyi	Membaca Catatan	...	Auditorial
11	P	Olahraga	Musik	Olahraga	Menulis Ulang	...	Kinestetik
13	P	Olahraga	Musik	Bernyanyi	Menulis Ulang	...	Kinestetik
15	L	Musik	Olahraga	Bernyanyi	Tanya Jawab	...	Auditorial
15	P	Seni	Musik	Menggambar	Membaca Catatan	...	Visual
16	P	Musik	Olahraga	Olahraga	Membaca Catatan	...	Visual
21	L	Olahraga	Musik	Olahraga	Membaca Catatan	...	Kinestetik
23	P	Seni	Musik	Menggambar	Tanya Jawab	...	Auditorial
24	L	Seni	Seni	Menggambar	Tanya Jawab	...	Visual
26	P	Seni	Musik	Bernyanyi	Membaca Catatan	...	Auditorial
27	P	Olahraga	Olahraga	Menggambar	Tanya Jawab	...	Kinestetik

Alur kerja algoritma *decision tree* yaitu:

Alur pertama kerja algoritma C4.5 yaitu menghitung Node 1 / *Root Node* dengan membuat tabel *Gain* dan *Entropy*. Berdasarkan data yang sudah dinormalisasi pada tabel 2. maka dibuatlah tabel Perhitungan *Gain* dan *Entropy* untuk Node 1 / *Root Node* seperti tabel 3. di bawah ini.

Tabel 3. Tabel Node 1

Attribut	Jml Kasus (S)	Visual (S1)	Auditorial (S2)	Kinestetik (S3)	Entropy	Gain
Total	15	4	6	6		
Jenis Kelamin						
L	8	2	4	2		
P	7	2	2	3		
Hobi						
Seni	5	2	3	0		
Musik	6	2	3	1		
Olahraga	4	0	0	4		
Kelas Favorit						
Seni	3	1	1	1		
Musik	8	2	3	3		
Olahraga	4	1	2	1		
Kemampuan						
Menggambar	5	3	1	1		
Bernyanyi	5	0	4	1		
Olahraga	5	1	1	3		
Metode Belajar						
Membaca Catatan	6	3	2	1		
Tanya Jawab	5	1	3	1		
Menulis Ulang	4	0	1	3		

Angka-angka pada tabel Node 1 di atas tertera sesuai dengan data-data pada tabel 3. dimana Kolom Jumlah Kasus (S) merupakan jumlah data yang dimasukkan yaitu total 15 data dan Kolom Visual (S1), Auditorial (S2) dan Kinestetik (S3) merupakan hasil keputusan yang disesuaikan dengan atribut yang memiliki hasil keputusan Gaya Belajar sesuai pada tabel 4.7 dimana hasil keputusan Visual bertotal 4 buah, Auditorial bertotal 6 buah, dan Kinestetik bertotal 6 buah dari seluruh data yang digunakan.

Pada baris Atribut "Jenis Kelamin", terdapat sub atribut berupa "L" dan "P". Total masing-masing sub atribut dimasukkan pada kolom Jumlah Kasus (S) yang berarti Jumlah kasus sub atribut "L" terdapat 8 buah, dan sub atribut "P" terdapat 7 buah. Pada kolom Keputusan S1 - S3 diisi sesuai dengan jumlah keputusan sesuai dengan sub atribut yang tersedia yang berarti jumlah keputusan Visual terdapat 2 buah, Auditorial terdapat 4 buah, Kinestetik terdapat 2 buah untuk sub atribut "L" dari atribut "Jenis Kelamin". Pengisian dilakukan sampai ke sub atribut terakhir.

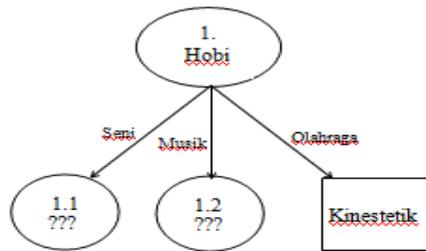
Perhitungan *Entropy* total dan masing-masing sub atribut dilakukan terlebih dahulu karena *Entropy* tersebut akan digunakan untuk mencari nilai *Gain* dari atribut.

Setelah didapatkan semua nilai *Entropy* dan *Gain*nya maka akan dipilih atribut dengan nilai *Gain* terbesar untuk Node 1 / *Root Node*

Tabel 4. Tabel Node 1 setelah mendapatkan *Gain* dan *Entropy*

Attribut	Jml Kasus (S)	Visual (S1)	Auditorial (S2)	Kinestetik (S3)	Entropy	Gain
Total	15	4	6	6	0.7881	
Jenis Kelamin						
L	8	2	4	2	0.75	
P	7	2	2	3	0.7869	0.020877
Hobi						
Seni	5	2	3	0	0	
Musik	6	2	3	1	0.7437	
Olahraga	4	0	0	4	0	0.490609
Kelas Favorit						
Seni	3	1	1	1	0.8074	
Musik	8	2	3	3	0.7437	
Olahraga	4	1	2	1	0.75	0.009773
Kemampuan						
Menggambar	5	3	1	1	0.6578	
Bernyanyi	5	0	4	1	0	
Olahraga	5	1	1	3	0.6697	0.345583
Metode Belajar						
Membaca Catatan	6	3	2	1	0.7276	
Tanya Jawab	5	1	3	1	0.6697	
Menulis Ulang	4	0	1	3	0	0.273815

Yang berarti atribut “Hobi” akan dijadikan Node 1. atau *Root Node*, kemudian digambarkan *decision tree* dengan Hobi sebagai akarnya, dan masing-masing sub atribut akan digambarkan pada gambar 3.



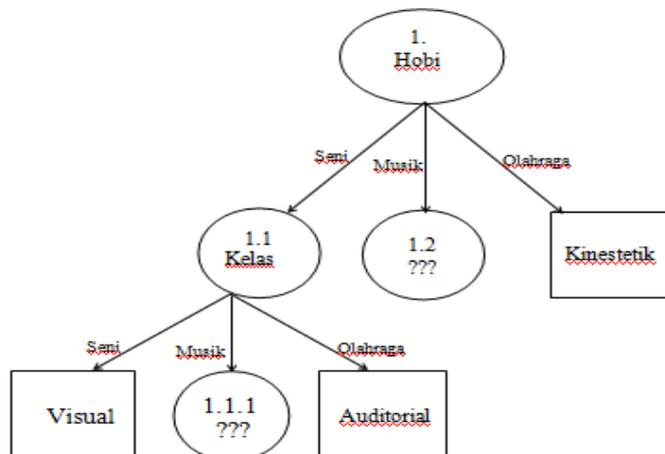
Gambar 2. *Decision Tree* Node 1.

Pada sub atribut “Olahraga” langsung mendapatkan hasil keputusan Kinestetik dikarenakan pada sub atribut “Olahraga” hanya memiliki 1 jenis keputusan sehingga dapat langsung digambarkan leaf nodenya. Selanjutnya akan dilakukan perhitungan ulang *Gain* dan *Entropy* untuk sub atribut “Seni” dari atribut “Hobi” dengan menggambar tabel 5.

Tabel 5. Tabel Node 1.1

Atribut	Jml Kasus (S)	Visual (S1)	Auditorial (S2)	Kinestetik (S3)	Entropy	Gain
Hobi (Seni)	5	2	3	0	0	
Jenis Kelamin						0.788103
L	2	1	1	0	0	
P	3	1	2	0	0	
Kelas Favorit						0.788103
Seni	1	1	0	0	0	
Musik	3	1	2	0	0	
Olahraga	1	0	1	0	0	
Kemampuan						0.788103
Menggambar	3	2	1	0	0	
Bernyanyi	2	0	2	0	0	
Olahraga	0	0	0	0	0	
Metode Belajar						0.788103
Membaca Catatan	1	1	1	0	0	
Tanya Jawab	2	1	1	0	0	
Menulis Ulang	1	0	1	0	0	

Dengan melihat terdapatnya angka 0 pada tiap sub atribut, dapat dipastikan seluruh nilai *Entropy* nya adalah 0 dan memiliki nilai *Gain* yang sama sehingga tidak diperlukan untuk melakukan perhitungan ulang untuk *Gain* dan *Entropy* masing-masing atribut. Untuk melanjutkan gambar *decision tree* maka dipilihlah salah satu atribut di atas, yaitu atribut “Kelas Favorit” untuk menjadi node 1.1 pada gambar *decision tree* seperti gambar 4.



Gambar 3. *Decision Tree* Node 1.1

Setelah menemukan node 1.1 maka akan dicari lagi node 1.2 dari *decision tree* yang berarti mencari data-data yang memiliki atribut “Hobi” dengan sub atribut “Musik” dengan menggambar tabel node 1.2.

Tabel 6. Tabel Node 1.2

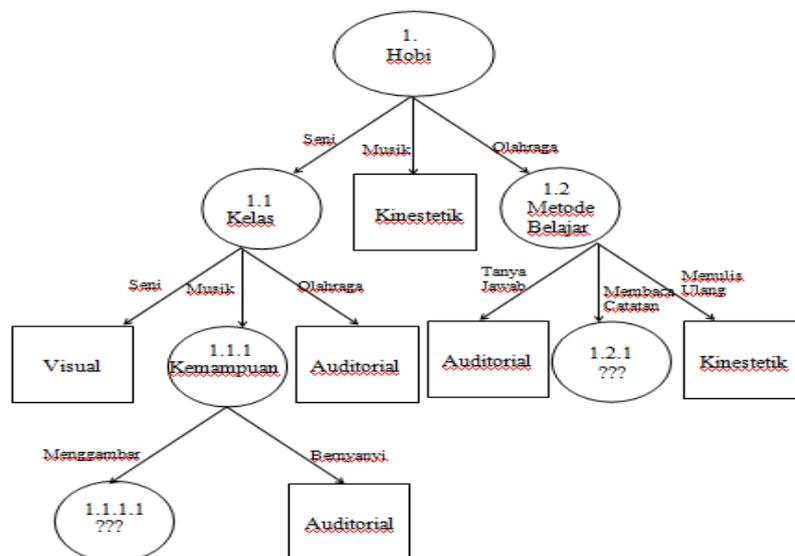
Attribut	Jml Kasus (S)	Visual (S1)	Auditorial (S2)	Kinestetik (S3)	Entropy	Gain
Hobi (Musik)	6	2	3	1	0.7437	
Jenis Kelamin						
L	5	1	3	1		
P	1	0	1	0	0	
Kelas Favorit						0.788103
Seni	2	0	1	1	0	
Musik	2	1	1	0	0	
Olahraga	2	1	1	0	0	
Kemampuan						
Menggambar	1	2	1	0	0	
Bernyanyi	2	0	2	0	0	
Olahraga	3	1	1	1		
Metode Belajar						0.788103
Membaca Catatan	3	2	1	0	0	
Tanya Jawab	2	0	2	0	0	
Menulis Ulang	1	0	0	1	0	

Tidak seperti tabel 4.10, tidak semua *Entropy* sub atribut di tabel node 1.2 bernilai 0, sehingga dilakukan perhitungan *Entropy* dan *Gain* ulang dengan menjadikan atribut “Hobi” dengan sub atribut “Musik” sebagai pengganti total data. Setelah menghitung *Gain* dan *Entropy* untuk node 1.2 maka akan dipilih atribut sebagai node 1.2 dengan nilai *Gain* tertinggi.

Tabel 7. Tabel Node 1.2 setelah mendapatkan *Gain* dan *Entropy*

Attribut	Jml Kasus (S)	Visual (S1)	Auditorial (S2)	Kinestetik (S3)	Entropy	Gain
Hobi (Musik)	6	2	3	1	0.7437	
Jenis Kelamin						0.564861
L	5	1	3	1	0.6697	
P	1	0	1	0	0	
Kelas Favorit						0.788103
Seni	2	0	1	1	0	
Musik	2	1	1	0	0	
Olahraga	2	1	1	0	0	
Kemampuan						0.788103
Menggambar	1	2	1	0	0	
Bernyanyi	2	0	2	0	0	
Olahraga	3	1	1	1	0.8074	
Metode Belajar						0.788103
Membaca Catatan	3	2	1	0	0	
Tanya Jawab	2	0	2	0	0	
Menulis Ulang	1	0	0	1	0	

Dikarenakan terdapat 3 buah *Gain* yang berjumlah sama, maka dipilihlah salah 1 atribut yaitu atribut “Metode Belajar” sebagai node 1.2. Dengan demikian gambar *decision tree* pun menjadi seperti pada gambar 5.



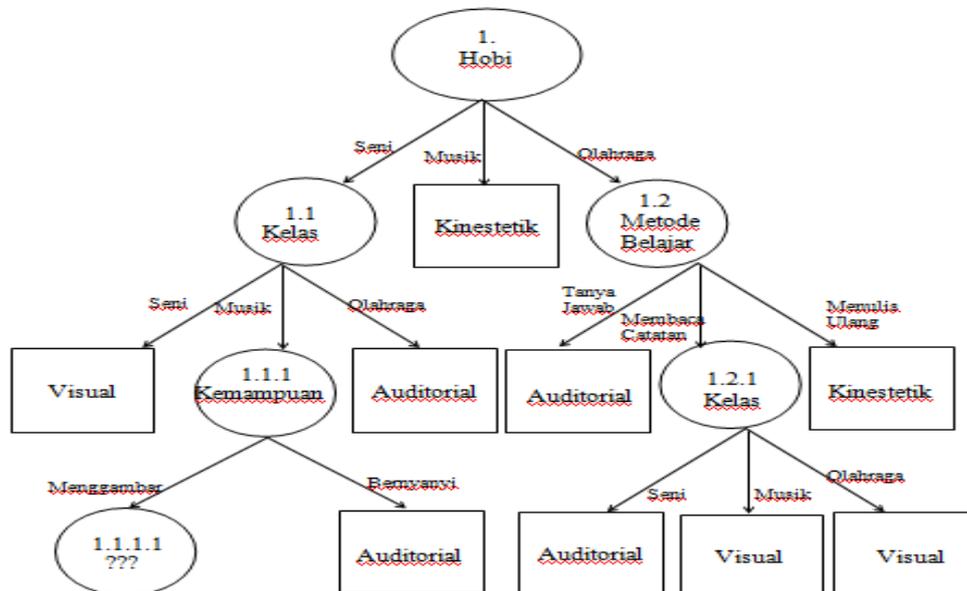
Gambar 4. *Decision Tree* Node 1.2.1

Langkah berikutnya adalah mencari node 1.2.1 dengan menggunakan atribut “Metode Belajar” dengan sub atribut “Membaca Catatan” sebagai pengganti total data.

Tabel 8. Tabel Node 1.2.1

Attribut	Jml Kasus (S)	Visual (S1)	Auditorial (S2)	Kinestetik (S3)	Entropy	Gain
Hobi (Musik), Metode (Membaca Catatan)	3	2	1	0	0	
Jenis Kelamin						0.788103
L	2	1	1	0	0	
P	1	1	0	0	0	
Kelas Favorit						0.788103
Seni	1	0	1	0	0	
Musik	1	1	0	0	0	
Olahraga	1	1	0	0	0	
Kemampuan						0.788103
Menggambar	1	1	0	0	0	
Bernyanyi	1	0	1	0	0	
Olahraga	1	1	0	0	0	

Seperti tabel node sebelumnya, dengan semua *Entropy* atribut bernilai 0 maka langsung dipilih atribut untuk mengisi node 1.2.1. Dengan demikian atribut “Kelas Favorit” terpilih untuk node 1.2.1 sehingga gambar decision tree bertambah lagi seperti pada gambar 6.



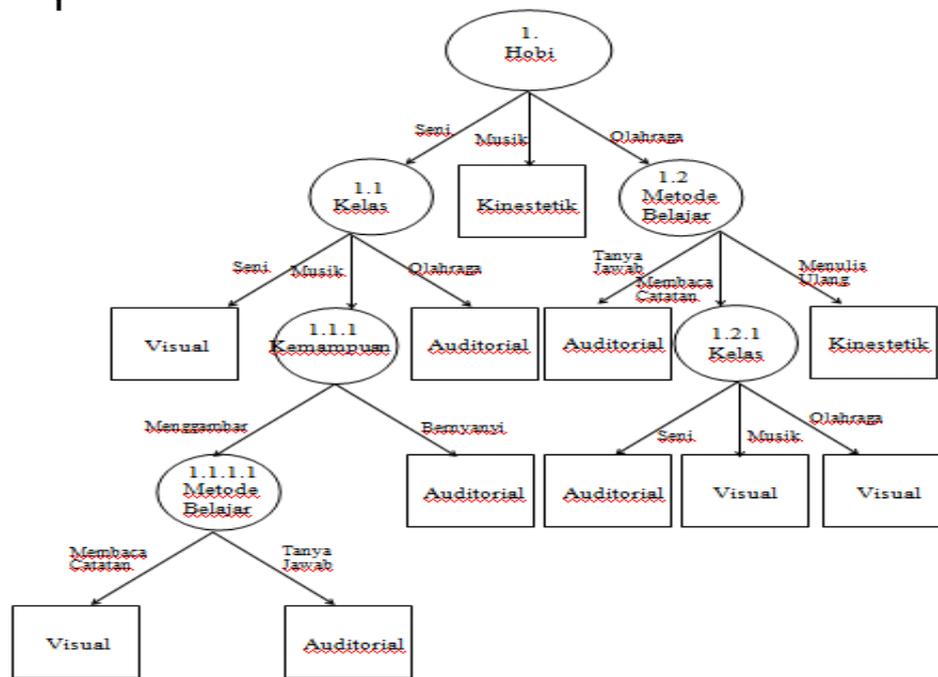
Gambar 5. Decision Tree Node 1.2.1

Setelah menemukan node 1.2.1 maka langkah selanjutnya adalah mencari node 1.1.1.1 dengan berdasarkan atribut “Kemampuan” sub atribut “Menggambar” dari atribut “Kelas” sub atribut “Musik” dari atribut “Hobi” sub atribut “Seni” dengan menggunakan tabel node 1.1.1.1.

Tabel 9. Tabel Node 1.1.1.1

Attribut	Jml Kasus (S)	Visual (S1)	Auditorial (S2)	Kinestetik (S3)	Entropy	Gain
Hobi (Seni), Kelas (Musik)	2	1	1	0	0	
Jenis Kelamin						0.788103
L	0	0	0	0	0	
P	2	1	1	0	0	
Metode Belajar						0.788103
Membaca Catatan	1	1	0	0	0	
Tanya Jawab	1	0	1	0	0	
Menulis Ulang	0	0	0	0	0	

Sesuai dengan aturan sebelumnya, bila *Entropy* seluruh atribut bernilai 0 maka salah satu atribut tersebut akan dipilih karena memiliki nilai *Gain* yang setara. Dengan demikian, atribut metode digunakan untuk mengisi node 1.1.1.1.



Gambar 6. Gambar Akhir *Decision Tree*

Dengan tidak adanya node selanjutnya, semua node telah memiliki keputusan dan semua kasus sudah masuk ke dalam class, maka decision tree dapat diselesaikan. Dan terbentuklah sejumlah rules dalam decision tree tersebut, seperti:

1. Jika Hobi = Olahraga maka class = Kinestetik
Jika Atribut Hobi dari siswa-siswi adalah Olahraga maka Gaya Belajar siswa siswi tersebut adalah Kinestetik.
2. Jika Hobi = Seni DAN Kelas = Seni maka class = Visual
Jika Atribut Hobi adalah Seni dan Atribut Kelas adalah Favorit dari siswa-siswi adalah Seni maka Gaya Belajar siswa siswi tersebut adalah Visual.
3. Jika Hobi = Seni DAN Kelas = Olahraga maka class = Auditorial
Jika Atribut Hobi adalah Seni dan Atribut Kelas adalah Olahraga dari siswa-siswi maka Gaya Belajar siswa siswi tersebut adalah Auditorial.
4. Jika Hobi = Musik DAN Metode Belajar = Tanya Jawab maka class = Auditorial
Jika Atribut Hobi adalah Musik dan Atribut Metode Belajar adalah Tanya Jawab dari siswa-siswi maka Gaya Belajar siswa siswi tersebut adalah Auditorial.
5. Jika Hobi = Musik DAN Metode Belajar = Menulis Ulang maka class = Kinestetik
Jika Atribut Hobi adalah Musik dan Atribut Metode Belajar adalah Menulis Ulang dari siswa-siswi maka Gaya Belajar siswa siswi tersebut adalah Kinestetik.
6. Jika Hobi = Seni DAN Kelas = Musik DAN Kemampuan = Bernyanyi maka class = Auditorial
Jika Atribut Hobi adalah Seni dan Atribut Kelas adalah Musik dan Kemampuan adalah Bernyanyi dari siswa-siswi maka Gaya Belajar siswa siswi tersebut adalah Auditorial.
7. Jika Hobi = Musik DAN Metode Belajar = Membaca Catatan DAN Kelas = Seni maka class = Auditorial
Jika Atribut Hobi adalah Musik dan Atribut Metode Belajar adalah Membaca Catatan dan Kelas Favorit adalah Seni dari siswa-siswi maka Gaya Belajar siswa siswi tersebut adalah Auditorial.
8. Jika Hobi = Musik DAN Metode Belajar = Membaca Catatan DAN Kelas = Musik maka class = Visual
Jika Atribut Hobi adalah Musik dan Atribut Metode Belajar adalah Membaca Catatan dan Kelas Favorit adalah Musik dari siswa-siswi maka Gaya Belajar siswa siswi tersebut adalah Visual.
9. Jika Hobi = Musik DAN Metode Belajar = Membaca Catatan DAN Kelas = Olahraga maka class = Visual
Jika Atribut Hobi adalah Musik dan Atribut Metode Belajar adalah Membaca Catatan dan Kelas Favorit adalah Olahraga dari siswa-siswi maka Gaya Belajar siswa siswi tersebut adalah Visual.
10. Jika Hobi = Seni DAN Kelas = Musik DAN Kemampuan = Menggambar DAN Metode Belajar = Membaca Catatan maka class = Visual
Jika Atribut Hobi adalah Seni dan Atribut Kelas adalah Musik dan Kemampuan adalah Menggambar dan Atribut Metode Belajar adalah Membaca Catatan dari siswa-siswi maka Gaya Belajar siswa-siswi tersebut adalah Visual.
11. Jika Hobi = Seni DAN Kelas = Musik DAN Kemampuan = Menggambar DAN Metode Belajar = Tanya Jawab maka class = Auditorial
Jika Atribut Hobi adalah Seni dan Atribut Kelas adalah Musik dan Kemampuan adalah Menggambar dan Atribut Metode Belajar adalah Tanya Jawab dari siswa-siswi maka Gaya Belajar siswa siswi tersebut adalah Auditorial

5.1. Hasil Perbandingan Evaluasi Akurasi

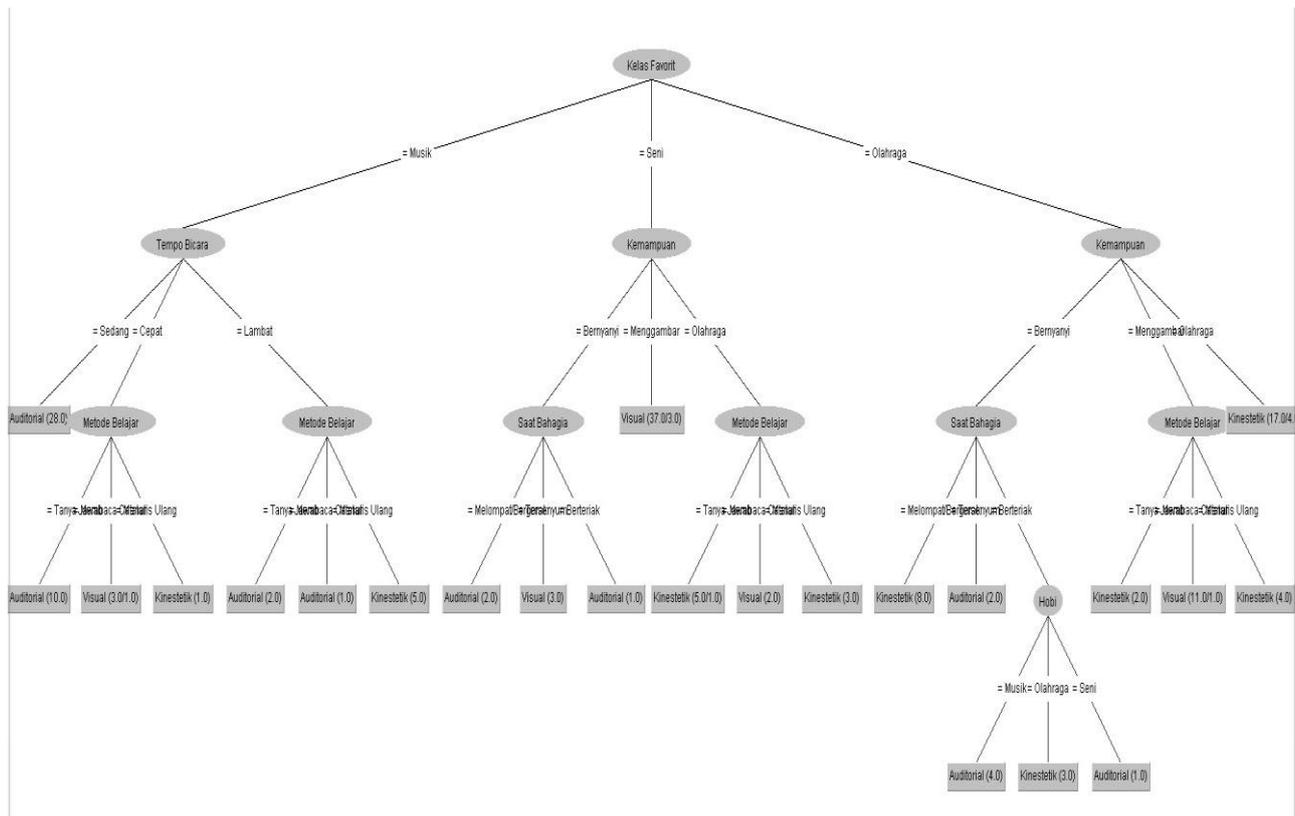
Setelah dilakukan analisis klasifikasi decision tree pada tools WEKA menggunakan *Use Training Set*, *10-Fold Cross Validation*, *5-Fold Cross Validation* maka didapatkan akurasi tertinggi yaitu dengan menggunakan *Use Training Set* dengan persentasi akurasi

yaitu 93.5484% untuk *Correctly Classified Instances* dan 6.4516 %. Untuk *Incorrectly Classified Instances*. Perbandingan hasil analisis dapat dilihat pada tabel 10.

Tabel 10. Perbandingan Evaluasi Akurasi

MODEL EVALUASI	Akurasi	Jumlah Kelas	Persentase	Satuan
Use Training Set	<i>Correctly Classified Instances</i>	145	93.5484	%
	<i>Incorrectly Classified Instances</i>	10	6.4516	%
10 Fold Cross Validation	<i>Correctly Classified Instances</i>	120	77.4194	%
	<i>Incorrectly Classified Instances</i>	35	22.5806	%
5 Fold Cross Validation	<i>Correctly Classified Instances</i>	129	83.2258	%
	<i>Incorrectly Classified Instances</i>	26	26.7742	%

Hasil Analisis Classify Dengan Tools Weka



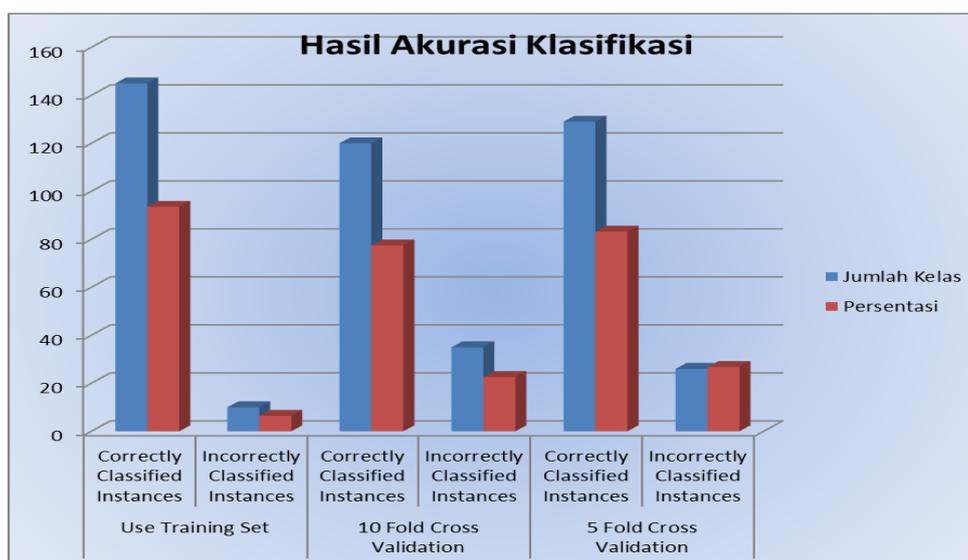
Gambar 7. Hasil Decision Tree J48 menggunakan Tools WEKA

Decision Tree yang terdapat pada gambar 8. didapatkan dengan menggunakan tools WEKA melalui *Classify J4.8 Tree* dimana algoritma J4.8 merupakan algoritma C4.5 versi WEKA. Terdapatnya 23 leaf nodes yang berarti menghasilkan 23 buah rules seperti yang telah di jelaskan pada Bab IV sebelumnya. Rules tersebut berupa :

1. Jika *attribut Kelas* adalah **Musik** dan *attribut Tempo Bicara* adalah **Sedang** maka *class Gaya Belajarnya* adalah **Auditoria**
2. Jika *attribut Kelas* adalah **Musik** dan *attribut Tempo Bicara* adalah **Cepat** dan *attribut Metode Belajar* adalah **Tanya Jawab** maka *class Gaya Belajarnya* adalah **Auditorial**
3. Jika *attribut Kelas* adalah **Musik** dan *attribut Tempo Bicara* adalah **Cepat** dan *attribut Metode Belajar* adalah **Membaca Catatan** maka *class Gaya Belajarnya* adalah **Visual**
4. Jika *attribut Kelas* adalah **Musik** dan *attribut Tempo Bicara* adalah **Cepat** dan *attribut Metode Belajar* adalah **Menulis Ulang** maka *class Gaya Belajarnya* adalah **Kinestetik**
5. Jika *attribut Kelas* adalah **Musik** dan *attribut Tempo Bicara* adalah **Lambat** dan *attribut Metode Belajar* adalah **Tanya Jawab** maka *class Gaya Belajarnya* adalah **Auditorial**
6. Jika *attribut Kelas* adalah **Musik** dan *attribut Tempo Bicara* adalah **Lambat** dan *attribut Metode Belajar* adalah **Membaca Catatan** maka *class Gaya Belajarnya* adalah **Auditorial**
7. Jika *attribut Kelas* adalah **Musik** dan *attribut Tempo Bicara* adalah **Cepat** dan *attribut Metode Belajar* adalah **Menulis Ulang** maka *class Gaya Belajarnya* adalah **Kinestetik**
8. Jika *attribut Kelas* adalah **Seni** dan *attribut Kemampuan* adalah **Menggambar** maka *class Gaya Belajarnya* adalah **Visual**
9. Jika *attribut Kelas* adalah **Seni** dan *attribut Kemampuan* adalah **Bernyanyi** dan *attribut Saat Bahagia* adalah **Melompat/Bergerak** maka *class Gaya Belajarnya* adalah **Auditorial**
10. Jika *attribut Kelas* adalah **Seni** dan *attribut Kemampuan* adalah **Bernyanyi** dan *attribut Saat Bahagia* adalah **Tersenyum** maka *class Gaya Belajarnya* adalah **Visual**
11. Jika *attribut Kelas* adalah **Seni** dan *attribut Kemampuan* adalah **Bernyanyi** dan *attribut Saat Bahagia* adalah **Berteriak** maka *class Gaya Belajarnya* adalah **Auditorial**

12. Jika *attribut Kelas* adalah **Seni** dan *attribut Kemampuan* adalah **Olahraga** dan *attribut Metode Belajar* adalah **Tanya Jawab** maka *class Gaya Belajarnya* adalah **Kinestetik**
13. Jika *attribut Kelas* adalah **Seni** dan *attribut Kemampuan* adalah **Olahraga** dan *attribut Metode Belajar* adalah **Membaca Catatan** maka *class Gaya Belajarnya* adalah **Visual**
14. Jika *attribut Kelas* adalah **Seni** dan *attribut Kemampuan* adalah **Olahraga** dan *attribut Metode Belajar* adalah **Menulis Ulang** maka *class Gaya Belajarnya* adalah **Kinestetik**
15. Jika *attribut Kelas* adalah **Olahraga** dan *attribut Kemampuan* adalah **Bernyanyi** dan *attribut Saat Bahagia* adalah **Melompat/Bergerak** maka *class Gaya Belajarnya* adalah **Kinestetik**
16. Jika *attribut Kelas* adalah **Olahraga** dan *attribut Kemampuan* adalah **Bernyanyi** dan *attribut Saat Bahagia* adalah **Tersenyum** maka *class Gaya Belajarnya* adalah **Auditorial**
17. Jika *attribut Kelas* adalah **Olahraga** dan *attribut Kemampuan* adalah **Bernyanyi** dan *attribut Saat Bahagia* adalah **Berteriak** dan *attribut Hobi* adalah **Musik** maka *class Gaya Belajarnya* adalah **Auditorial**
18. Jika *attribut Kelas* adalah **Olahraga** dan *attribut Kemampuan* adalah **Bernyanyi** dan *attribut Saat Bahagia* adalah **Berteriak** dan *attribut Hobi* adalah **Olahraga** maka *class Gaya Belajarnya* adalah **Kinestetik**
19. Jika *attribut Kelas* adalah **Olahraga** dan *attribut Kemampuan* adalah **Bernyanyi** dan *attribut Saat Bahagia* adalah **Berteriak** dan *attribut Hobi* adalah **Seni** maka *class Gaya Belajarnya* adalah **Auditorial**
20. Jika *attribut Kelas* adalah **Olahraga** dan *attribut Kemampuan* adalah **Menggambar** dan *attribut Metode Belajar* adalah **Tanya Jawab** maka *class Gaya Belajarnya* adalah **Kinestetik**
21. Jika *attribut Kelas* adalah **Olahraga** dan *attribut Kemampuan* adalah **Menggambar** dan *attribut Metode Belajar* adalah **Membaca Catatan** maka *class Gaya Belajarnya* adalah **Visual**
22. Jika *attribut Kelas* adalah **Olahraga** dan *attribut Kemampuan* adalah **Menggambar** dan *attribut Metode Belajar* adalah **Menulis Ulang** maka *class Gaya Belajarnya* adalah **Kinestetik**
23. Jika *attribut Kelas* adalah **Olahraga** dan *attribut Kemampuan* adalah **Olahraga** maka *class Gaya Belajarnya* adalah **Kinestetik**

Hasil Akurasi Klasifikasi Dengan *Decision Tree*:



Gambar 8. Hasil Akurasi Klasifikasi Dengan *Decision Tree*

Gambar 9. merupakan Grafik persentasi hasil klasifikasi dengan menggunakan *Use Training Set* dengan jumlah kelas *Correctly Classified Instances* 145, *Incorrectly Classified Instances* 10, dan persentasi akurasi *Correctly Classified Instances* sebesar 93.5484%, *Incorrectly Classified Instances* 6.4516 %. Menggunakan *5-cross validation* *Correctly* dengan jumlah kelas *Classified Instances* 120, *Incorrectly Classified Instances* 35, dan persentasi akurasi *Correctly Classified Instances* sebesar 77.4194%, *Incorrectly Classified Instances* 22.5806%. Menggunakan *10-Fold Cross Validation* dengan jumlah kelas *Correctly Classified Instances* 129, *Incorrectly Classified Instances* 26, dan persentasi akurasi *Correctly Classified Instances* sebesar 83.2258%, *Incorrectly Classified Instances* 26.7742%.

6. KESIMPULAN

Kesimpulan yang dapat di ambil dari hasil yang penelitian ini adalah bahwa Penelitian prediksi gaya belajar siswa siswi ini menggunakan data siswa-siswi di SD Sariputra tahun ajaran 2018 dan 2019 dengan jumlah data 155 dan menggunakan *attribut* sejumlah 11 *attribut* untuk membantu proses prediksi yaitu, Hobi, Kemampuan, Kelas Favorit, Metode Belajar dan lain-lain. Hasil prediksi Gaya Belajar terdapat 3 jenis Gaya Belajar yaitu, Visual, Auditorial dan Kinestetik. Hasil prediksi metode *decision tree* dengan menggunakan algoritma C4.5 menghasilkan 23 *rules* untuk memprediksi gaya belajar dari masing-masing siswa-siswi SD. Dimana hasil dari *rules decision tree* tersebut dapat digunakan untuk memprediksi masing-masing gaya belajar siswa-siswi tingkat SD. Persentasi hasil akurasi *decision tree* dengan menggunakan *Use Training Set* dengan persentasi akurasi *Correctly Classified Instances* sebesar 93.5484 % dan *Incorrectly Classified Instances* 6.4516 %. menggunakan *5-cross validation* *Correctly* dengan persentasi akurasi *Correctly Classified Instances* sebesar 83.2258 % dan *Incorrectly Classified Instances* 26.7742 %. Menggunakan *10-Fold Cross Validation* dengan persentasi akurasi *Correctly Classified Instances* sebesar 77.4194 %, *Incorrectly Classified Instances* 22.5806 %.

DAFTAR PUSTAKA

- [1] S. S. K. M S Hasibuan, LE Nugroho, P I Santosa, "A Proposed Model for Detecting Learning Styles Based on Agent Learning," *Int. J. Emerg. Technol. Learn.*, vol. 11, no. 10, pp. 65–69, 2016.
- [2] A. L. Bire, U. Geradus, and J. Bire, "Pengaruh Gaya Belajar Visual, Auditorial, Dan Kinestetik Terhadap Prestasi Belajar Siswa," *J. Kependidikan*, 2014.
- [3] V. S. Moertini, "Data Mining Sebagai Solusi Bisnis," *Integral*, vol. 7, no. 1, pp. 44–56, 2002.
- [4] J. Jasmir, D. Z. Abidin, S. Nurmaini, and R. F. Malik, "Penerapan Metode K-Nearest Neighbor dalam Memprediksi Masa Studi Mahasiswa (Studi Kasus : Mahasiswa STIKOM Dinamika Bangsa)," *Pros. Annu. Res. Semin. 2017*, vol. 3, no. 1, pp. 133–138, 2017.
- [5] J. Jasmir, D. Z. Abidin, E. Rasywir, and P. A. Jusia, "Prediksi Mahasiswa Drop Out dengan menggunakan Algoritma Klasifikasi Data Mining," *Pros. Annu. Res. Semin. 2018*, vol. 4, no. 1, pp. 978–979, 2018.
- [6] D. H. Kamagi and S. Hansun, "Implementasi Data Mining dengan Algoritma C4 . 5 untuk Memprediksi Tingkat Kelulusan Mahasiswa," *ULTIMATICS*, vol. VI, no. 1, pp. 15–20, 2014.
- [7] A. G. Mabur and R. Lubis, "Penerapan Data Mining untuk Memprediksi Kriteria Nasabah Kredit," *J. Komput. dan Inform.*, vol. 1, 2012.
- [8] A. Dipaloka, "Penerapan Decision Tree untuk Penentuan Pola Data Penerimaan Mahasiswa Baru," *J. Penelit. Sitrotika*, no. September, 2016.
- [9] J. Kozak, "Ant Colony Decision Trees – A New Method for Constructing Decision Trees Based on Ant Colony Optimization Ant Colony Decision Trees – A New Method," no. March, 2014.
- [10] F. F. Harryanto, S. Hansun, U. M. Nusantara, G. Serpong, and C. Pegawai, "Penerapan Algoritma C4 . 5 untuk Memprediksi Penerimaan Calon Pegawai Baru di PT WISE," *J. Tek. Inform. dan Sist. Inf.*, vol. 3, no. 2, pp. 95–103, 2017.