



Artikel

Penerapan Data Mining Penentu Minat Dan Bakat Siswa Smk Dengan Metode C4.5

Dita Merawati¹, Rino²,^{1,2} Universitas Buddhi Dharma, Teknik Informatika, Banten, Indonesia

PENGAJUAN

Received: 28 Agustus 2019
 Final Revision: 25 September 2019
 Available Online: 30 September 2019

KATA KUNCI

Data Mining, Minat dan Bakat, Metode C4.5

KORESPONDENSI

E-mail: bodhikaruna26@gmail.com
 E-mail: rino@ubd.ac.id

A B S T R A K

Tes bakat dan minat merupakan bagian dari tes inteligensi dalam tes psikologi. Dalam Kamus Psikologi J.P. Chaplin menjelaskan bahwa minat adalah suatu sikap yang berlangsung terus menerus yang mampu membuat pola perhatian seseorang sehingga membuat dirinya jadi selektif terhadap objek minatnya. Sedangkan bakat adalah suatu kondisi pada seseorang yang memungkinkan untuk mencapai suatu kecakapan, pengetahuan, dan keterampilan dengan suatu latihan khusus. Dari data tes Minat dan Bakat, dilakukan pengolahan data yaitu data mining dengan metode C4.5. Data yang didapatkan bersumber dari pengambilan data sekunder di SMK X. Kumpulan data yang didapat terbagi menjadi dua bagian yaitu data minat sebanyak 174 data dan data bakat sebanyak 97 data. Penelitian ini dilakukan untuk mengetahui tingkat keakuratan yang didapatkan berdasarkan penerapan metode yang dilakukan. Hasil dari penerapan metode C4.5 menunjukkan bahwa metode C4.5 dapat diterapkan pada pengolahan data tes minat dan bakat tersebut dengan hasil akurasi yaitu 82,65% pada tes bakat dan 90,23 pada tes minat. Penerapan metode C4.5 menghasilkan pohon keputusan. Dari hasil pohon keputusan tersebut, dilakukan penerapan algoritma C4.5 pada aplikasi yang dibuat pada penelitian ini. Aplikasi digunakan untuk mempermudah penentuan hasil minat dan bakat seseorang setelah melakukan tes secara manual karena sebelumnya penentuan hasil dilaksanakan secara manual.

I. PENDAHULUAN

Dalam Undang-Undang Republik Indonesia Nomor 20 Tahun 2003 Tentang Sistem Pendidikan Nasional disebutkan bahwa Sekolah Menengah Kejuruan (SMK) merupakan sekolah menengah lanjutan dari pendidikan dasar yaitu Sekolah Menengah Pertama (SMP) atau sekolah lanjutan yang diakui setara atau sederajat. Lulusan dari SMK dapat melanjutkan ke perguruan tinggi, menjadi wirausaha, atau menjadi pekerja pada dunia usaha dan industri (DUDI)

(KEMENPORA 2017). Penting bagi siswa SMK untuk mengetahui minat dan bakat mereka sebagai masukan setelah mereka lulus nanti, terutama bagi siswa yang berminat untuk melanjutkan ke perguruan tinggi.

Dalam Kamus Psikologi J.P. Chaplin menjelaskan bahwa minat adalah suatu sikap yang berlangsung terus menerus yang mampu membuat pola perhatian seseorang sehingga membuat dirinya jadi selektif terhadap objek minatnya. Sedangkan bakat menurut KBBI adalah

dasar seperti kepandaian, sifat, dan pembawaan yang dibawa sejak lahir. Bakat dan minat seseorang dapat diketahui dengan berbagai cara, salah satunya adalah melakukan tes minat dan bakat

Pelaksanaan tes minat dan bakat yang dilakukan seseorang membentuk data individu. Data dari dilaksanakan tes tersebut akan terus bertambah seiring berjalannya waktu dan hanya digunakan sebagai arsip atau sekedar laporan oleh lembaga penyelenggara. Dari kumpulan data tersebut dapat digali informasi baru melalui *data mining* yang dapat berguna bagi lembaga-lembaga tertentu.

Data mining sendiri merupakan serangkaian proses untuk mendapatkan informasi berguna dari gudang basis data yang besar (Prasetyo 2012). Pengambilan keputusan dalam *data mining* dilakukan dengan teknik klasifikasi (*clasification*), klusterisasi (*clustering*), regresi, pencarian pola sekuensial dan kaidah asosiasi (Prasetyo 2014).

Setiap teknik dalam *data mining* memiliki metode-metode untuk mendukung dalam pengambilan keputusan. Namun, tidak semua metode dapat diterapkan dalam setiap kasus atau dalam pengambilan keputusan. Setiap metode yang diterapkan memiliki tingkat keakuratan berbeda-beda pada setiap kasus. Pada kasus penentuan minat dan bakat siswa SMK ini, diterapkan metode C4.5 sebagai metode pengambil keputusan. Penerapan metode C4.5 pada kasus tes minat dan bakat siswa SMK menghasilkan algoritma. Algoritma tersebut diterapkan pada aplikasi berbasis android untuk membantu menghitung hasil minat dan bakat setelah melakukan tes secara manual.

II. METODE

1.1 MINAT DAN BAKAT

Minat adalah suatu sikap yang berlangsung terus menerus yang mampu membuat pola perhatian seseorang sehingga membuat dirinya jadi selektif

terhadap objek minatnya. Perasaan yang menyatakan bahwa satu aktivitas pekerjaan atau objek itu berharga atau berarti bagi individu. Satu keadaan motivasi yang menuntun tingkah laku menuju arah sasaran tertentu (Chaplin, 2009).

Bakat diartikan sebagai kemampuan bawaan, sebagai potensi yang masih perlu dikembangkan dan dilatih agar terwujud (Kementrian Pendidikan dan Kebudayaan 2016). Dalam hal ini bakat memerlukan latihan dan pendidikan agar suatu tindakan dapat dilakukan dimasa yang akan datang. Kemampuan menunjukkan bahwa suatu tindakan (Performance) yang dapat dilakukan saat ini. Prestasi merupakan perwujudan dari bakat dan kemampuan.

Salah satu cara untuk mengetahui minat dan bakat seseorang adalah dengan melakukan tes minat dan bakat. Tes bakat dan minat dapat digunakan untuk memetakan seseorang dalam pemilihan bakat dan minat yang sesuai dengan kompetensi dan keterampilan yang dimiliki.

Tes minat dan bakat dilakukan agar seseorang dapat meningkatkan kelebihan dan mengatasi kelemahan yang ada pada diri mereka, sehingga Individu dapat memperoleh manfaat (Kementrian Pendidikan dan Kebudayaan 2016) :

- 1) Mengetahui dan memahami apa yang menjadi kelebihan/kekuatan Individu dan menemukan cara yang tepat untuk mengembangkan kekuatan tersebut secara optimal.
- 2) Mengetahui dan memahami apa yang menjadi kekurangan/kelemahan Individu agar dapat mengatasi dan mengkompensasi kelemahan tersebut.

Tes minat dan bakat yang disusun oleh Kementrian Pendidikan dan Kebudayaan termasuk dalam karakteristik tes bakat dan minat dalam konteks pendidikan. Tes bakat dan minat ini pada umumnya sama dengan tes bakat dan minat psikologi lainnya, mengukur minat dan bakat seseorang dengan tujuan agar dalam

menjalani sebuah pendidikan atau pekerjaan sesuai dengan minat dan bakat tersebut.

1.2 DATA MINING

Nama *data mining* mulai dikenal sejak tahun 1990, ketika pekerjaan pemanfaatan data menjadi sesuatu yang penting dari berbagai bidang, mulai dari bidang akademik, bisnis, hingga medis (Gorunescu 2011). “*Data mining* adalah campuran dari statistik, kecerdasan buatan, dan riset basis data” yang masih berkembang.

Data mining merupakan serangkaian proses untuk menggali informasi yang belum diketahui secara manual dari suatu basis data. Informasi yang dihasilkan diperoleh dengan cara mengekstrasi dan mengenali pola yang penting atau menarik dari data yang terdapat dalam basis data. *Data mining* digunakan untuk mencari pengetahuan yang terdapat dalam basis data yang besar sehingga sering disebut *Knowledge Discovery Databases (KDD)*.

1). *Data selection*

Pada tahap ini dilakukan pemisahan data yang digunakan dari data operasional.

2). *Pre-processing / cleaning*

Sebelum proses *Data mining* dapat dilaksanakan, perlu dilakukan proses *cleaning* pada data yang menjadi fokus KDD. Proses *cleaning* antara lain membuang duplikasi data, memeriksa data yang inkonsisten, dan memperbaiki kesalahan pada data.

3). *Transformation*

Data diubah atau digabung ke dalam format yang sesuai sehingga data tersebut sesuai untuk proses *data mining*.

4). *Data mining*

Data mining adalah proses mencari pola atau informasi menarik dalam data

terpilih dengan menggunakan teknik, metode atau algoritma tertentu.

5). *Interpretation / evaluation*

Pola informasi yang dihasilkan dari proses *Data mining* perlu ditampilkan dalam bentuk yang mudah dimengerti oleh pihak yang berkepentingan.

1.3 METODE C4.5

Algoritma C4.5 diperkenalkan oleh Quinlan (1996) sebagai versi perbaikan dari ID3 (Prasetyo 2014). Dalam ID3 induksi *decision tree* hanya bisa dilakukan pada fitur bertipe kategorikal (nominal atau ordinal), sedangkan tipe numerik (interval atau rasio) tidak dapat digunakan. Perbaikan yang dilakukan adalah dapat menangani fitur dengan tipe numerik, melakukan pemotongan (*pruning*) *decision tree*, dan penurunan (*deriving*) rule set.

Untuk membentuk pohon keputusan dengan algoritma C4.5 secara garis besar yaitu:

- 1) **Membangun pohon keputusan** diawali dengan pembentukan akar (terletak paling atas). Kemudian data dibagi berdasarkan atribut-atribut yang cocok untuk dijadikan daun.
- 2) **Pemangkasan pohon yang telah di bentuk (*tree pruning*)** yaitu mengidentifikasi dan membuang cabang yang tidak diperlukan pada pohon yang telah terbentuk. Pemangkasan pohon dilakukan selain untuk pengurangan ukuran pohon juga bertujuan untuk mengurangi tingkat kesalahan prediksi pada kasus baru dari hasil pemecahan yang dilakukan dengan *divide and conquer*. *Pruning* ada dua pendekatan yaitu :
 - a) *Pre-pruning* yaitu menghentikan pembangunan suatu *subtree* lebih awal (dengan memutuskan untuk

tidak lebih jauh mempartisi data training). Saat seketika berhenti, maka *node* berubah menjadi *leaf* (node akhir). *Node* akhir ini menjadi kelas yang paling sering muncul di antara subset sampel.

- b) *Post-pruning* yaitu menyederhanakan *tree* dengan cara membuang beberapa cabang *subtree* setelah *tree* selesai dibangun. *Node* yang jarang dipotong akan menjadi *leaf* (node akhir) dengan kelas yang paling sering muncul.

- 3) **Pembentukan aturan keputusan** yaitu membuat aturan keputusan dari pohon yang telah dibentuk. Aturan tersebut dapat dalam bentuk *if—then* diturunkan dari pohon keputusan dengan melakukan penelusuran dari akar sampai ke daun.

Tahap dalam algoritma (Vulandari 2017) untuk membangun pohon keputusan (*decision tree*) adalah sebagai berikut:

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

Untuk memilih atribut sebagai akar harus dicari *gain* tertinggi di antara atribut – atribut yang ada, dengan rumus:

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

Dengan :

S = himpunan kasus

A = Atribut

n = jumlah partisi

|S_i| = jumlah kasus pada partisi ke-i

|S| = jumlah kasus dalam S

Sedangkan untuk perhitungan nilai entropi adalah sebagai berikut:

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

Dimana,

S = himpunan kasus

n = jumlah partisi S

p_i = proporsi dari S_i terhadap S

1.4 K-FOLD CROSS VALIDATION

K-Fold Cross Validation merupakan salah satu metode dalam menentukan data training dan data testing dari keseluruhan data. *K-fold cross validation* mengulang k-kali untuk membagi sebuah himpunan data secara acak menjadi k-subset yang saling bebas, setiap ulangan disisakan satu subset untuk pengujian dan subset lainnya untuk pelatihan. Nilai K yang disarankan adalah 10 karena lebih akurat dalam ukuran estimasinya. Dalam proses iterasi harus dilakukan sebanyak 10 kali dan dalam bagian S1 sampai S10 harus pernah dijadikan data testing sebanyak 1 kali. K bernilai 10 sehingga dalam proses iterasi dihasilkan *decision tree* sebanyak 10.

III. HASIL

3.1 Pre-processing / cleaning

Data yang digunakan bersumber dari data primer yang didapatkan melalui pelaksanaan tes minat dan bakat dengan peserta siswa SMK “X”. Dari data primer yang didapat, dilakukan pemisahan data yang digunakan dari sekumpulan data operasional.

3.2 Transformation

Pada tahap ini, dilakukan pemilihan dataset yang digunakan dan transformasi data sehingga data set yang didapat mudah

diolah. Trasformasi dilakukan tanpa mengubah hasil data.

Hasil dari transformasi data yang dilakukan terdapat pada tabel 3.1 dan 3.2

Tabel 3.1 Tabel Data Set Minat

Id	Bidang	Minat	Hasil
1	Administrasi	Marketing	tata niaga
2	Administrasi	Kasir	administrasi
3	Administrasi	Akuntan	keuangan
4	Administrasi	Sekretaris	adminitrasi
5	Administrasi	Administrasi	adminitrasi
6	Administrasi	Akuntan	keuangan
7	Administrasi	Akuntan	keuangan
8	Administrasi	Administrasi	adminitrasi
9	Administrasi	Kasir	administrasi
10	Administrasi	Sekretaris	adminitrasi

Sumber : Penulis

3.2 Tabel Data Set Bakat

N O	N a m a	A V	P V	F A	K K	P U	P N	P S	P T	S	Bakat
1	L H H	1	2	2	1	1	2	1	1	1	Ekonomi
2	V D P	1	1	2	1	1	2	2	1	1	Teknologi Tekstil
3	K W I	1	2	2	1	1	2	1	1	1	Ekonomi
4	W P S	1	1	2	1	1	2	2	1	1	Teknologi Tekstil
5	K P I	2	1	1	1	1	2	2	1	1	IT Komputer
6	E T Y	2	2	1	1	1	1	2	1	1	Tarian
7	J V N	2	2	1	1	1	2	1	1	1	Psikologi
8	T O A	2	1	1	1	1	2	2	1	1	IT Komputer
9	C L A	2	2	1	1	1	1	2	1	1	Tarian
10	C S I	1	1	2	1	1	2	1	1	2	Akuntansi dan Keuangan

Sumber : Penulis

Keterangan :

PV:Penalaran Visual
PN:Penalaran Numerik

AV:Analisa Verbal

PU:Penalran Urutan

PS:Penalaran Spasial

PT:Tiga Dimensi

S:Sistematisasi

K:Kosa Kata

Fig:Figural Angka

1 : TIDAK

2 : YA

3.3 Data mining

Dari data yang didapatkan, dibentuk *decision tree* dengan metode C4.5 untuk membuat alur keputusan.

Langkah pertama adalah menentukan atribut sebagai akar dengan menghitung *gain information*

$$Entropy \text{ (total)} = -\left(\left(\frac{3}{96} \times \log_2 \frac{3}{96}\right) + \left(\frac{4}{96} \times \log_2 \frac{4}{96}\right) + \left(\frac{6}{96} \times \log_2 \frac{6}{96}\right) + \left(\frac{4}{96} \times \log_2 \frac{4}{96}\right) + \left(\frac{10}{96} \times \log_2 \frac{10}{96}\right) + \left(\frac{6}{96} \times \log_2 \frac{6}{96}\right) + \left(\frac{1}{96} \times \log_2 \frac{1}{96}\right) + \left(\frac{21}{96} \times \log_2 \frac{21}{96}\right) + \left(\frac{2}{96} \times \log_2 \frac{2}{96}\right) + \left(\frac{2}{96} \times \log_2 \frac{2}{96}\right) + \left(\frac{2}{96} \times \log_2 \frac{2}{96}\right) + \left(\frac{6}{96} \times \log_2 \frac{6}{96}\right) + \left(\frac{8}{96} \times \log_2 \frac{8}{96}\right) + \left(\frac{2}{96} \times \log_2 \frac{2}{96}\right) + \left(\frac{3}{96} \times \log_2 \frac{3}{96}\right) + \left(\frac{3}{96} \times \log_2 \frac{3}{96}\right) + \left(\frac{3}{96} \times \log_2 \frac{3}{96}\right) + \left(\frac{4}{96} \times \log_2 \frac{4}{96}\right) + \left(\frac{2}{96} \times \log_2 \frac{2}{96}\right) + \left(\frac{4}{96} \times \log_2 \frac{4}{96}\right) + \left(\frac{3}{96} \times \log_2 \frac{3}{96}\right)\right) = 3,970308106$$

Setelah menghitung entropy total, hitung nilai entropy dari masing-masing atribut dengan rumus yang sama.

$$Entropy \text{ (AV,tidak)} = -\left(\left(\frac{3}{96} \times \log_2 \frac{3}{96}\right) + \left(\frac{4}{96} \times \log_2 \frac{4}{96}\right) + 0 + 0 + \left(\frac{10}{96} \times \log_2 \frac{10}{96}\right) + \left(\frac{6}{96} \times \log_2 \frac{6}{96}\right) + 0 + 0 + 0 + 0 + \left(\frac{2}{96} \times \log_2 \frac{2}{96}\right) + 0 + \left(\frac{8}{96} \times \log_2 \frac{8}{96}\right) + \left(\frac{2}{96} \times \log_2 \frac{2}{96}\right) + \left(\frac{1}{96} \times \log_2 \frac{1}{96}\right) + \left(\frac{3}{96} \times \log_2 \frac{3}{96}\right) + \left(\frac{3}{96} \times \log_2 \frac{3}{96}\right) + \left(\frac{4}{96} \times \log_2 \frac{4}{96}\right) + \left(\frac{2}{96} \times \log_2 \frac{2}{96}\right) + \left(\frac{4}{96} \times \log_2 \frac{4}{96}\right) + 0\right) = 3,450560517$$

$$Entropy \text{ (AV,ya)} = -\left(\left(0 + 0 + \left(\frac{6}{96} \times \log_2 \frac{6}{96}\right) + \left(\frac{4}{96} \times \log_2 \frac{4}{96}\right) + 0 + 0 + \left(\frac{1}{96} \times \log_2 \frac{1}{96}\right) + \left(\frac{21}{96} \times \log_2 \frac{21}{96}\right) + \left(\frac{2}{96} \times \log_2 \frac{2}{96}\right) + \left(\frac{2}{96} \times \log_2 \frac{2}{96}\right) + 0 + \left(\frac{3}{96} \times \log_2 \frac{3}{96}\right) + 0 + 0 + \left(\frac{2}{96} \times \log_2 \frac{2}{96}\right) + 0 + 0 + 0 + 0 + 0 + \left(\frac{3}{96} \times \log_2 \frac{3}{96}\right)\right) = 2,476290348$$

Setelah dilakukan penghitungan nilai entropy, hitung nilai gain dari masing-masing atribut dengan rumus :

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

* $Entropy(S_i)$

Berikut merupakan penerapan rumus gain pada atribut AV:

$$Gain(AV) = 3,970308106 - \left(\left(\frac{52}{96} \times 3,450560517 \right) + \left(\frac{44}{96} \times 2,476290348 \right) \right) = 0,966288083$$

Atribut dengan *information gain* tertinggi digunakan sebagai akar untuk membentuk *decision tree*.

3.4 Interpretation / evaluation

Pada tahap ini, dibentuk pola informasi yang dihasilkan dari proses data mining perlu ditampilkan dalam bentuk yang mudah dimengerti.

Pola informasi yang dihasilkan adalah sebagai berikut:

a. Bakat

```

Visual = 1
|   Numerik = 1
|   |   Angka = 1
|   |   |   Spasial = 1: Sastra
|   |   |   Spasial = 2
|   |   |   |   Verbal = 1: Katring
|   |   |   |   Verbal = 2: Sejarah Seni
|   |   |   Angka = 2
|   |   |   |   Verbal = 1: Manajemen Hotel
|   |   |   |   Verbal = 2: Tata Bahasa
|   |   Numerik = 2
|   |   |   Sistematis = 1
|   |   |   |   Dimensi = 1
|   |   |   |   Verbal = 1: Teknologi
Tekstil
|   |   |   |   Verbal = 2: Ekonomi
|   |   |   |   Dimensi = 2
|   |   |   |   Angka = 1: Arsitektur
|   |   |   |   Angka = 2: Matematika
Statistiska
|   |   |   Sistematis = 2: Akuntansi dan
Keuangan
Visual = 2
|   Verbal = 1
|   |   Numerik = 1
|   |   |   Dimensi = 1
|   |   |   |   Angka = 1: Geografi
|   |   |   |   Angka = 2: Olah Raga dan
Penelitian Rekreasi
|   |   |   |   Dimensi = 2: Desain Seni
|   |   |   |   Numerik = 2: IT Komputer
|   |   Verbal = 2
|   |   |   Spasial = 1
|   |   |   |   Kosa Kata = 1: Psikologi
    
```

```

|   |   |   Kosa Kata = 2: Politik dan
Hubungan Internasional
|   |   Spasial = 2: Media dan Komunikasi
    
```

b. Minat

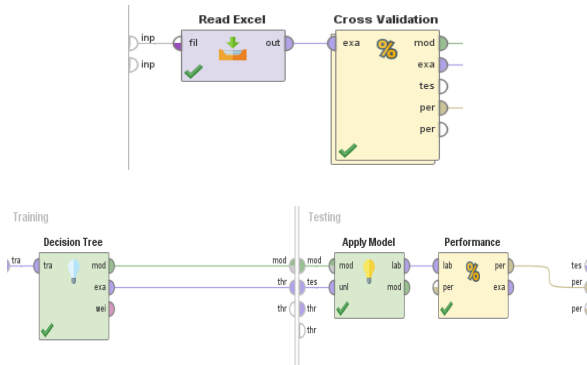
```

Minat = Administrasi: adminitrasi
Minat = Akuntan: keuangan
Minat = Apoteker: kesehatan
Minat = Audio Visual: teknik komputer dan
informatika
Minat = Bartender: tata boga
Minat = Desainer
|   Bidang = Pariwisata: tata busana
|   Bidang = Teknik: teknologi tekstil
Minat = Dokter: kesehatan
Minat = Dokter Hewan: kesehatan hewan
Minat = Drama: seni teater
Minat = Editing: broadcasting
Minat = Farmasi: kesehatan
Minat = Fotografer: bradcasting
Minat = Insinyur: teknik energi terbarukan
Minat = Jaringan Internet: telekomunikasi
Minat = Juru Masak: tata boga
Minat = Kameramen: broadcasting
Minat = Kasir: administrasi
Minat = Komponen Listrik: teknik tenaga
listrikan
Minat = Konstruksi Kapal: teknik perkapalan
Minat = Konstruksi Kayu: teknik furniture
Minat = Kontraktor: teknik bangunan
Minat = Make Up Artis: tata kecantikan
Minat = Marketing: tata niaga
Minat = Mekanik: otomotif
Minat = Mekanik Pesawat: teknik pesawat udara
Minat = Model Pakaian: tata busana
Minat = Musisi: seni musik
Minat = Nahkoda: teknologi penangkapan ikan
Minat = Operator Pabrik: teknik mesin
Minat = Panti Asuhan: pekerja sosial
Minat = Pegawai Tambang: teknik geologi
pertambangan
Minat = Pekerja Sosial: pekerja sosial
Minat = Pelukis: seni rupa
Minat = Pemandu Wisata: kepariwisataan
Minat = Pemasaran: tata niaga
Minat = Penari: seni tari
Minat = Pengawas Hasil Pertanian dan
Perikanan: agrobisnis pengolahan hasil
pertanian dan perikanan
Minat = Perawat: kesehatan
Minat = Perawatan Mesin: teknik industry
Minat = Pertambangan dan Minyak: teknik
perminyakan
    sosial
Minat = Petugas Perkebunan: agrobisnis
produksi tanam
Minat = Petugas Pertanahan: geomatika
Minat = Produksi Hasil Pertanian: mekanisme
pertanian
Minat = Programmer: teknik komputer dan
informatika
Minat = Psikologi: pekerja sosial
Minat = Salon: tata kecantikan
Minat = Sekertaris: adminitrasi
Minat = Teknisi Instrumen: teknik
instrumentasi industry
Minat = Travelling: kepariwisataan
Minat = kameramen: broadcasting
    
```

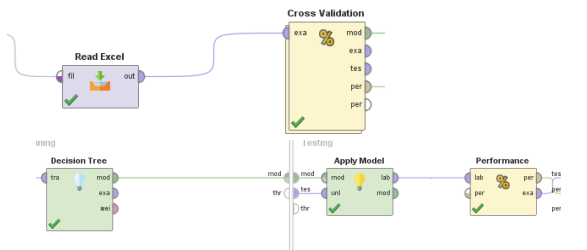
IV. PEMBAHASAN

4.1 Pengujian Pada *Rapidminer*

Penerapan dan pengujian data set juga dilakukan pada *rapidminer*. Rancangan layar dapat dilihat pada gambar 4.1 dan 4.2, sedangkan *data testing* yang digunakan sebanyak 10 data.



Gambar 4.1 Perancangan Penerapan C4.5 pada Rapidminer (Kasus : Minat siswa)
Sumber : Rapidminer



Gambar 4.2 Perancangan Penerapan C4.5 pada Rapidminer (Kasus : Bakat siswa)
Sumber : Rapidminer

Pengujian *K-Fold Cross Validation* metode 4.5 pada kasus tes minat dan bakat siswa SMK menggunakan *rapidminer* menghasilkan keakuratan sebagai berikut:

a. Bakat

accuracy: 82.78% +/- 11.45% (micro average: 82.65%)

Metode C4.5 pada tes bakat memiliki keakuratan sebesar 82,78%.

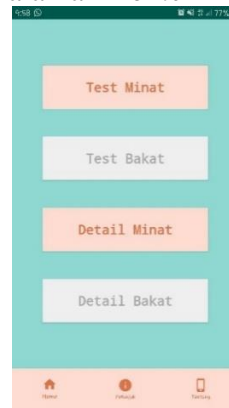
b. Minat

accuracy: 90.23% +/- 5.40% (micro average: 90.23%)

Metode C4.5 pada tes bakat memiliki keakuratan sebesar 90,23%.

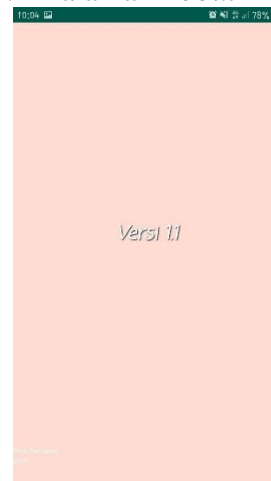
4.2 Tampilan Program

a. Halaman Home



Gambar IV.1 Tampilan halaman home

b. Halaman About



Gambar IV.2 Tampilan halaman about

c. Halaman Petunjuk



Gambar IV.3 Tampilan halaman petunjuk

d. Halaman Input Aspek Minat



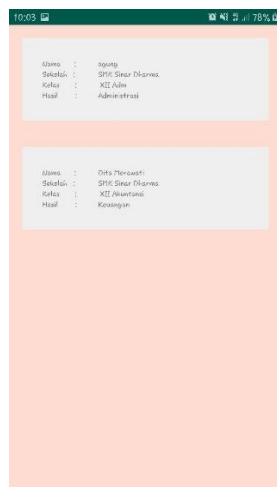
Gambar IV.4 Tampilan halaman input aspek minat

e. Halaman Input Aspek Bakat



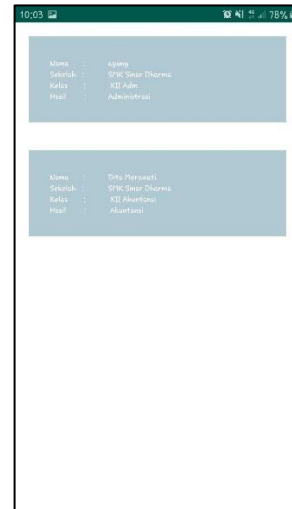
Gambar IV.5 Tampilan halaman input aspek bakat

f. Halaman Daftar Hasil Minat



Gambar IV.6 Tampilan halaman daftar hasil minat

g. Halaman Daftar Hasil Bakat



Gambar IV.6 Tampilan halaman daftar hasil bakat

V. KESIMPULAN DAN SARAN

5.1 Kesimpulan

Berdasarkan penelitian yang dilakukan, penulis menyimpulkan bahwa:

- a. Metode C4.5 dapat digunakan sebagai salah satu metode pada kasus penentuan minat dan bakat siswa SMK karena tingkat keakuratan yang didapatkan pada tes bakat mencapai 82,65% dan pada tes minat mencapai 90,23%.
- b. Pembuatan aplikasi dengan menerapkan alur *decission tree* hasil dari penelitian sebagai algoritma, dapat digunakan untuk menentukan hasil minat dan bakat siswa SMK berdasarkan data yang dimasukkan.

5.2 Saran

Dalam pelaksanaan penelitian ini, penulis tidak luput dari kekurangan. Penulis mengharapkan penelitian selanjutnya dapat memperbaiki beberapa kekurangan yang dapat dikembangkan seperti :

- a. Pengembangan aplikasi menjadi aplikasi tes minat bakat berbasis

mobile sehingga tidak menggunakan tes manual atau tertulis lagi.

- b. Pengujian menggunakan penggabungan beberapa metode untuk membandingkan tingkat keakuratan.
- c. Perbandingan beberapa metode untuk mengetahui, metode mana yang lebih baik pada kasus yang sama.
- d. Penyelenggaraan tes minat dan bakat siswa SMK untuk memperoleh data primer dilakukan tiga kali atau lebih dengan kurun waktu yang berbeda agar mendukung keakuratan dari hasil tes.

DAFTAR PUSTAKA

- [1] Eka Sabna, Muhardi. 2016. “Penerapan Data Mining Untuk Memprediksi Prestasi Akademik Mahasiswa Berdasarkan Dosen, Motivasi, Kedisiplinan, Ekonomi, dan Hasil Belajar.” *Jurnal CoreIT* 41-44.
- [2] Gorunescu, Florin. 2011. *Data Mining Concepts, Models, and Techniques*. Berlin: Spring-Verlag.
- [3] Hermawati, Fajar Astuti. 2013. *Data Mining*. Yogyakarta: Andi.
- [4] KEMENPORA. 2017. *Revitalisasi SMK Untuk Produktivitas dan Daya Saing Bangsa*. 22 Mei. Diakses April 8, 2019. <http://www.kemempora.go.id/index/preview/pengumuman/432>.
- [5] Kementerian Pendidikan dan Kebudayaan. 2016. *Buku Panduan Bakat & Minat SMK*. Jakarta.
- [6] Prasetyo, Eko. 2012. *Data Mining - Konsep dan Aplikasi Menggunakan MATLAB*. Yogyakarta: C.V Andi Offset.
- [7] —. 2014. *Data Mining Mengolah Data Menjadi Informasi Menggunakan Matlab*. Yogyakarta: Andi.
- [8] RapidMiner. 2018. *About Us Learn More About RapidMiner*. Diakses November 08, 2018. <https://rapidminer.com/us/>.
- [9] RI, Kementerian kesehatan. 2014. “Infodatin Pusat Data dan Informasi Kementerian Kesehatan RI.” *Situasi dan Analisis Diabetes*, 14 November: 8.
- [10] SMK, Direktorat Pembinaan. t.thn. Diakses Mei 12, 2019. <http://eksis.ditpsmk.net/artikel/definisi-smk-sekolah-menengah-kejuruan>.
- [11] Vulandari, Retno Tri. 2017. *Data Mining Teori dan Aplikasi Rapid Miner*. Yogyakarta: Gava Media.

BIOGRAFI

Dita Merawati, S.Kom, dilahirkan di Temanggung, 25 Juni 1997. Sekolah Dasar dilaksanakan di SDN 3 Getas, Kaloran, Temanggung, SMPN 2 Kaloran, Temanggung, SMAS Sinar Dharma, Jakarta Barat. Pendidikan S-1 (Teknik Informatika – Universitas Buddhi Dharma) lulus tahun 2019.

Rino M.Kom, dilahirkan di Bogor, 20 Mei 1985. Pendidikan Sarjana Komputer – Konsentrasi Jaringan di STMIK Buddhi, Tangerang, pada tahun 2004-2008 dan Pendidikan Magister Komputer – Konsentrasi Data Mining di STMIK Eresha, Tangerang pada tahun 2010-2012.