

Prediksi Ketepatan Waktu Kelulusan Dengan Algoritma Data Mining C4.5

Indah Puji Astuti ¹⁾*

Teknik Informatika, Universitas Muhammadiyah Ponorogo
indahsan.0912@gmail.com ¹⁾*

Abstrak

Mahasiswa merupakan salah satu unsur yang ada di Universitas atau Perguruan Tinggi. Setiap mahasiswa memiliki data yang berbeda-beda, seperti informasi identitas diri seperti asal daerah, jenis sekolah asal, pekerjaan orang tua, jenis kelas tempat mahasiswa belajar, dan lain sebagainya. Mahasiswa memiliki tingkat kelulusan yang berbeda, tepat waktu dan tidak tepat waktu. Yang menjadi kendala universitas pada umumnya dan fakultas pada khususnya adalah banyaknya mahasiswa yang lulus tidak tepat waktu. Padahal jumlah kelulusan mahasiswa tiap tahun menjadi salah satu faktor yang akan dinilai ketika fakultas atau program studi suatu universitas mengajukan akreditasi. Algoritma C4.5 merupakan salah satu algoritma klasifikasi yang menggunakan pohon keputusan. Pada penelitian ini dilakukan analisis terhadap data mahasiswa program Studi Teknik Informatika Universitas Muhammadiyah Ponorogo angkatan 2012/2013. Penggunaan pohon keputusan dalam penelitian ini berguna untuk memanfaatkan data mahasiswa, dan menemukan pola-pola tertentu antara beberapa atribut input dengan sebuah atribut target. Data mahasiswa yang digunakan sebagai atribut input diantaranya adalah jenis sekolah asal, asal daerah, pekerjaan orang tua, dan kelas. Atribut output yang digunakan untuk mengklasifikasikan data mahasiswa adalah status, yang terdiri dari "lulus tepat waktu" dan "lulus tidak tepat waktu". Hasil analisis menunjukkan bahwa, kasus pada penelitian ini algoritma C4.5 dapat memprediksi dengan nilai akurasi hanya sebesar 82%.

Kata kunci: Algoritma C4.5, data mahasiswa, data mining, pohon keputusan

Abstract

[Data Mining to Predict the graduation time with C4.5 Algorithm] The student is one of unsure in University or Higher Education. The student has different of data, such as self-identity information such as address, type of school, work of parents, type of class, etc. The students whose graduation rate is different, on time and not on time. The number of students who graduate is not on time will be a problem not only for university but also for faculty. The number of students graduating each year is one of factor of assessment when faculty or study program submits accreditation. C4.5 Algorithm is one of classification algorithm with decision trees. In this study conducted an analysis of student data Engineering Studies Program University of Muhammadiyah Ponorogo 2012/2013. The decision trees in this case is useful for exploring student data, finding the hidden relationship between of input attributes with a target attribute. The input attribute consist of, the type of school, address, work of parent, and type of class. The output attribute to classify is status, which consists of "on time" and "not on time". The results from this analysis shown that in this case the C4.5 algorithm can predict with an accuracy value only 82%.

Keywords: Algoritma C4.5, student data, data mining, decision tree

1. PENDAHULUAN

Mahasiswa merupakan salah satu unsur yang ada dalam suatu Universitas atau Perguruan Tinggi. Setiap mahasiswa memiliki data yang beraneka ragam, mulai dari informasi identitas diri seperti asal daerah, jenis sekolah asal (SMA, SMK, MAN), pekerjaan orang tua, jenis kelas tempat mahasiswa belajar, dan lain sebagainya. Data akademik universitas yang telah tersimpan dapat menjadi bahan analisis, diantaranya untuk mengetahui hubungan antara prestasi dan

sekolah asal, SKS yang diambil dengan topik konsentrasi peminatan, jenis sekolah asal dengan ketepatan waktu kelulusan, dan banyak hal lainnya.

Banyak mahasiswa yang tingkat kelulusannya berbeda, tepat waktu dan tidak tepat waktu. Yang menjadi kendala universitas pada umumnya dan fakultas pada khususnya adalah banyaknya mahasiswa yang lulus tidak tepat waktu. Padahal jumlah kelulusan mahasiswa tiap tahun menjadi salah satu

faktor yang akan dinilai ketika fakultas atau program studi suatu universitas mengajukan akreditasi.

Algoritma C4.5 merupakan algoritma klasifikasi yang menggunakan pohon keputusan. Pohon keputusan yaitu menginterpretasikan data yang besar kedalam bentuk yang lebih kecil menggunakan struktur atau konsep pohon keputusan [1]. Proses pada pohon keputusan adalah dengan mengubah bentuk data (tabel) menjadi model pohon, kemudian mengubahnya lagi menjadi aturan-aturan dan menyederhanakan aturan-aturan tersebut [2].

Pada penelitian ini dilakukan analisis terhadap data mahasiswa program Studi Teknik Informatika Universitas Muhammadiyah Ponorogo angkatan 2012/2013. Banyak faktor yang menentukan mahasiswa dapat lulus tepat waktu atau tidak. Faktor tersebut diantaranya adalah faktor akademis, faktor kondisi keluarga, dan dapat juga karena faktor psikologis[3]. Penggunaan atribut penelitian ini diantaranya adalah NIM, jenis sekolah asal, asal daerah, pekerjaan orang tua, dan kelas. Atribut-atribut inilah yang akan menjadi bahan analisis ketepatan waktu studi mahasiswa.

Penggunaan pohon keputusan dalam penelitian ini berguna untuk memanfaatkan data mahasiswa, menemukan pola-pola yang ada antara sejumlah atribut *input* dengan sebuah atribut target. Data mahasiswa yang digunakan sebagai atribut *input* diantaranya adalah jenis sekolah asal, asal daerah, pekerjaan orang tua, dan kelas. Atribut status menjadi atribut *output* yang akan digunakan untuk mengklasifikasikan data mahasiswa, yang terdiri dari "tepat waktu" dan "tidak tepat waktu".

Penelitian oleh Kamagi dan Hansun yang dilakukan pada tahun 2014 menggunakan algoritma data mining C4.5 untuk prediksi tingkat kelulusan mahasiswa. Atribut yang dipakai untuk menentukan prediksi diantaranya adalah nama, NIM, IP semester 1-6, jenis kelamin, asal sekolah, dan jumlah SKS. Hasil penelitian mereka menunjukkan bahwa algoritma yang digunakan dapat memprediksi tingkat kelulusan mahasiswa dengan 4 jenis klasifikasi yaitu lulus cepat, lulus tepat, lulus terlambat dan *drop out*, dimana atribut IP semester 6 adalah atribut yang paling berpengaruh [4]. Hasil prediksi kelulusan pada penelitian ini dapat dimanfaatkan oleh program studi untuk mengetahui status kelulusan mahasiswa. Hal ini dapat menjadi rekomendasi pengambilan mata kuliah bagi mahasiswa untuk semester berikutnya seperti skripsi dan magang. Dengan hal tersebut mahasiswa bisa lulus minimal tepat waktu.

Penelitian yang dilakukan oleh Mujib, dkk pada tahun 2013 mengatakan bahwa faktor yang paling berpengaruh dalam penentuan klasifikasi kinerja akademik mahasiswa adalah Indeks Prestasi Kumulatif (IPK), Indeks Prestasi Semester (IPS) semester 1, IPS semester 4 dan jenis kelamin. Pada penelitian ini digunakan algoritma data mining C4.5 untuk penentuan prediksi kelulusan berdasarkan

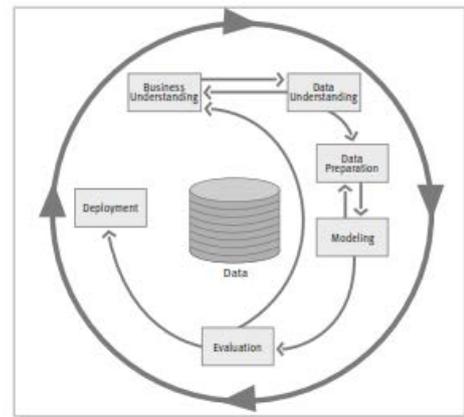
atribut jenis kelamin, asal sekolah SMA dan IPS semester 1 – 6 [5]

Penelitian yang dilakukan oleh Untari (2014) yaitu penggunaan metode *Decision Tree* C4.5 untuk menganalisa prediksi mahasiswa yang berpotensi non-aktif. Atribut yang dipakai untuk menentukan prediksi diantaranya adalah NIM, umur, tahun masuk, marital, status, kota asal, asal sekolah jalur masuk, IP semester 1 – 4, dan jumlah SKS yang diambil dari semester 1 – 4. Hasil pada penelitian ini atribut yang paling berpengaruh terhadap mahasiswa non-aktif adalah Indeks Prestasi Semester (IPS) dan Satuan Kredit Semester (SKS) yang diambil. [6]

Prediksi ketepatan waktu studi diharapkan dapat membantu program studi khususnya bagi para dosen pembimbing akademik untuk lebih intensif mengarahkan mahasiswa bimbingannya dalam belajar sehingga dapat menekan angka ketidakhadiran.

2. BAHAN DAN METODE

Bahan utama yang digunakan dalam penelitian ini, adalah data mahasiswa angkatan 2012/2013 Fakultas Teknik, Prodi Teknik Informatika pada Sistem Informasi Akademik (SIMTIK) Universitas Muhammadiyah Ponorogo. Metode penelitian yang digunakan sesuai dengan alur pada Model CRISP-DM (*Cross Industry Standard Process for Data Mining*), langkah-langkahnya adalah sebagai berikut:



Gambar 1. Fase CRISP-DM model [7]

Model CRISP-DM (*Cross Industry Standard Process for Data Mining*) terbagi dalam enam tahapan yaitu [7]:

1. *Business Understanding* (Pemahaman Bisnis)
2. *Data Understanding* (Pemahaman Data)
3. *Data Preparation* (Pengolahan Data)
4. *Modelling* (Pemodelan)
5. *Evaluation* (Evaluasi)
6. *Deployment* (Penyebaran)

Untuk penjelasan lebih lanjut tiap-tiap tahapan pada penelitian ini, akan dibahas di Hasil dan Pembahasan.

3. HASIL DAN PEMBAHASAN

Langkah-langkah penelitian ini sesuai alur pada Model CRISP-DM (*Cross Industry Standard Process for Data Mining*):

A. Pemahaman Bisnis (*Bussiness Understanding*)

Institusi perguruan tinggi baik itu lingkup Universitas atau Sekolah Tinggi sekarang dituntut untuk memiliki keunggulan dalam berdaya saing yaitu dengan meningkatkan mutu dan akreditasinya. Pencapaian akreditasi yang baik, diantaranya dipengaruhi oleh jumlah kelulusan mahasiswa per tahun. Banyak mahasiswa yang tingkat kelulusannya berbeda, tepat waktu dan tidak tepat waktu. Mahasiswa yang lulus tidak tepat waktu menjadi catatan penting untuk fakultas terutama program studi, ketika ingin mengajukan akreditasi.

B. Pemahaman Data (*Data Understanding*)

Dataset yang digunakan pada penelitian ini dikumpulkan dari Sistem Informasi Akademik (SIMTIK) Universitas Muhammadiyah Ponorogo. Data yang digunakan adalah data mahasiswa angkatan 2012/2013 Fakultas Teknik, Prodi Teknik Informatika. Tidak semua atribut dari data dari mahasiswa digunakan, hanya menggunakan 5 atribut yang terdiri dari NIM, jenis sekolah asal, asal daerah, pekerjaan orang tua, dan kelas.

Atribut-atribut tersebut diambil berdasarkan data diri mahasiswa yang ada di SIMTIK. Pemakaian ke 5 atribut tersebut diambil berdasarkan referensi penelitian sebelumnya yang dilakukan oleh Untari (2014) yang berjumlah 9 atribut. Atribut umur, tahun masuk, marital dan IP tidak digunakan dalam penelitian ini karena data yang tidak ada (tidak lengkap). Sehingga setelah melewati *pra processing* data, dalam penelitian ini hanya menggunakan 5 atribut.

Tabel 1. Seleksi Atribut

No	Atribut	Keterangan
1	NIM	ID
2	Jenis sekolah asal	√
3	Asal daerah	√
4	Pekerjaan orang tua	√
5	Kelas	√

C. Pengolahan Data (*Data Preparation*)

Seluruh data yang berhasil dikumpulkan sebanyak 150 data. Data terlebih dahulu dilakukan *pra processing* data. Ditahapan ini data dibersihkan. Atribut-atribut yang sekiranya tidak penting untuk digunakan dalam proses analisis, (contoh : atribut nama mahasiswa), maka atribut tersebut akan dihilangkan. Atau jika ada data mahasiswa yang kurang lengkap dalam pengisian, misalnya mahasiswa tidak menuliskan asal daerah atau jenis sekolah asal, maka data tersebut akan dihapus/tidak digunakan. Hasil dari *pra processing* data ini, jumlah data set nya menjadi 100 data.

Pada tahap ini adalah pemilihan atribut data yang akan digunakan untuk pemodelan. Tabel 1. menjelaskan mengenai atribut-atribut yang akan

digunakan dalam analisis. Berikut adalah perincian dari keseluruhan atribut:

1. NIM

Atribut NIM merupakan ID.

2. Jenis sekolah asal

Atribut jenis sekolah asal terdiri dari 3 macam, yaitu "SMA", "MAN", dan "SMK".

3. Asal daerah

Atribut asal daerah terdiri dari 2 macam, yaitu "Ponorogo" dan "Luar Ponorogo".

4. Pekerjaan orang tua

Atribut pekerjaan orang tua terdiri dari 5 macam jenis pekerjaan, yaitu PNS, wiraswasta, petani, guru, dan polisi.

5. Kelas

Atribut kelas terdiri dari 2 macam, yaitu digunakan untuk membedakan bahwa mahasiswa berasal dari kelas Reguler atau kelas Khusus.

D. Pemodelan (*Modelling*)

Metode yang akan digunakan adalah Algoritma C4.5 untuk membentuk *decision tree* nya. Salah satu kelas klasifikasi pada aplikasi data mining WEKA yang mengimplementasikan algoritma C4.5 adalah Algoritma J48 [8].

Algoritma C4.5 akan menghasilkan sebuah pohon keputusan. Cara menentukan atribut mana yang menjadi akar, dilihat pada nilai *gain* tertinggi dari atribut-atribut yang ada. Perhitungan *node* untuk data set, dapat dilihat pada Tabel 2.

Tabel 2. Tabel Perhitungan *Node*

Node			Jml data	T P	T T P	Entropy	Gain
	Total			100	83	17	0.657704779
Jenis sekolah asal	SMK		46	40	6	0.558629373	0.007785513
	SMA		46	37	9	0.713146749	
	MAN		8	6	2	0.811278124	
Asal daerah	Ponorogo		76	66	10	0.561752608	0.021765324
	Luar Ponorogo		24	17	7	0.870864469	
Pekerjaan orang tua	PNS		22	20	2	0.439496987	0.095583014
	Wiraswasta		50	43	7	0.584238812	
	Petani		23	19	4	0.666578358	
	Guru		3	0	3	0	
	Polisi		2	1	1	1	
Kelas	Reguler		87	71	16	0.688552168	0.00780273
	Khusus		13	12	1	0.391243564	

TP = Lulus Tepat Waktu

TTP = Lulus Tidak Tepat Waktu

Perhitungan untuk mendapatkan nilai *entropy* dan *gain* seperti terlihat pada tabel 2, yaitu dengan menggunakan rumus berikut [9]:

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

Menghitung *Entropi* Total:

$$Entropi (Total) = \sum_{i=1}^n \left(-\frac{83}{100} * \log_2 \left(\frac{83}{100} \right) \right) + \left(-\frac{17}{100} * \log_2 \left(\frac{17}{100} \right) \right) = 0.657704779$$

Menghitung *Entropi* Jenis Sekolah Asal:

$$Entropi (Jenis Sekolah Asal SMK) = \sum_{i=1}^n \left(-\frac{40}{46} * \log_2 \left(\frac{40}{46} \right) \right) + \left(-\frac{6}{46} * \log_2 \left(\frac{6}{46} \right) \right) = 0.558629373$$

$$Entropi (Jenis Sekolah Asal SMA) = \sum_{i=1}^n \left(-\frac{37}{46} * \log_2 \left(\frac{37}{46} \right) \right) + \left(-\frac{9}{46} * \log_2 \left(\frac{9}{46} \right) \right) = 0.713146749$$

$$Entropi (Jenis Sekolah Asal MAN) = \sum_{i=1}^n \left(-\frac{6}{8} * \log_2 \left(\frac{6}{8} \right) \right) + \left(-\frac{2}{8} * \log_2 \left(\frac{2}{8} \right) \right) = 0.811278124$$

Sementara itu, nilai *Gain* pada atribut jenis sekolah asal dihitung dengan menggunakan formula *gain* sebagai berikut:

$$Gain (S,A) = Entropi (S) - \sum_{i=1}^n \frac{|S_i|}{|S|} * Entropy S_i \dots (2)$$

Menghitung *Gain* Jenis Sekolah Asal

$$Gain (Total, Jenis Asal Sekolah) = 0.657704779 - \left(\left(\frac{46}{100} * 0.558629373 \right) + \left(\frac{46}{100} * 0.713146749 \right) + \left(\frac{8}{100} * 0.811278124 \right) \right) = 0.007785513$$

Dari Tabel 2 atribut dengan nilai *gain* tertinggi akan dipilih sebagai *node* pertama pada pohon keputusan. Pada *node* selanjutnya akan diisi oleh atribut-atribut yang bernilai *gain* lebih rendah dan begitu seterusnya. *Node* yang tidak memiliki percabangan akan menunjukkan *output* dari setiap cabangnya yang dikenal dengan nama *leaf* atau daun.

Pada tabel 2 terlihat bahwa nilai atribut pekerjaan orangtua memiliki nilai *gain* tertinggi, maka atribut ini menjadi atribut *root* pada pohon keputusan, kemudian dilanjutkan dengan atribut daerah, dan jenis sekolah asal. Diakhiri oleh status yang menyatakan keterangan tepat dan tidak tepat yang berfungsi sebagai *leaf*. Maka dapat dikatakan bahwa, kasus pada data penelitian ini parameter penentu pertama tingkat kelulusan seorang mahasiswa pada waktu yang akan datang dilihat dari pekerjaan orang tua, daerah, dan jenis sekolah asal mahasiswa tersebut.

Dengan membaca hasil pohon keputusan yang ada pada Gambar 3 terlihat bahwa atribut kelas tidak ada atau mengalami pemangkasan (*pruned*) dan tidak dipilih sebagai atribut untuk pohon keputusan. Namun demikian, seluruh informasi yang ada pada data set sudah bisa dibaca atau ditelusuri. Gambar 2 berikut ini merupakan hasil klasifikasi data menggunakan WEKA Classifier.

```
Classifier output
Relation: cobadatal
Instances: 100
Attributes: 5
    Jenis sekolah asal
    daerah
    pekerjaan orangtua
    KELAS
    status
Test mode:3-fold cross-validation

=== Classifier model (full training set) ===

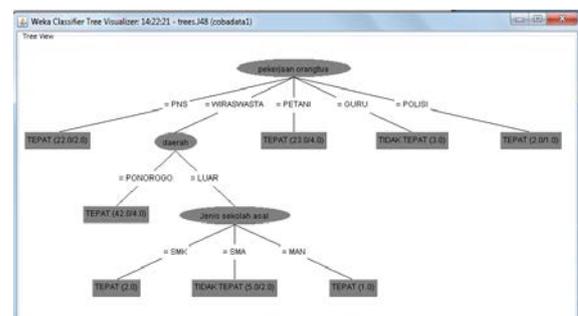
J48 unpruned tree
-----
pekerjaan orangtua = PNS: TEPAT (22.0/2.0)
pekerjaan orangtua = WIRASWASTA
| daerah = PONOROGO: TEPAT (42.0/4.0)
| daerah = LUAR
| | Jenis sekolah asal = SMK: TEPAT (2.0)
| | Jenis sekolah asal = SMA: TIDAK TEPAT (5.0/2.0)
| | Jenis sekolah asal = MAN: TEPAT (1.0)
pekerjaan orangtua = PETANI: TEPAT (23.0/4.0)
pekerjaan orangtua = GURU: TIDAK TEPAT (3.0)
pekerjaan orangtua = POLISI: TEPAT (2.0/1.0)

Number of Leaves : 8
Size of the tree : 11
```

Gambar 2. Hasil Klasifikasi menggunakan WEKA

Pada gambar 2 diatas terlihat bahwa WEKA Classifier hanya memilih atribut pekerjaan orang tua, daerah, dan jenis sekolah asal sebagai atribut dalam pohon keputusan, sedangkan atribut kelas langsung terpangkas dari pohon keputusan. Dapat disimpulkan bahwa dengan jumlah dan jenis data yang ada hanya dibutuhkan beberapa atribut untuk mendapatkan kelas *output* dari dataset tersebut [8]. Pohon keputusan dapat dilihat pada Gambar 3.

Gambar 3. Pohon Keputusan hasil Analisis Menggunakan WEKA



E. Evaluasi

Dalam tahap ini akan dilakukan pengukuran keakuratan hasil yang telah dicapai. Berdasarkan uji coba pembacaan pohon keputusan dengan sinkronisasi data set, pada kasus ini algoritma C4.5 memiliki tingkat akurasi kurang dari 100%.

Akurasi klasifikasi didapatkan berdasarkan tabel *confusion matrix*. *Confusion matrix* dari data *testing* yang digunakan dengan *output* Tepat dan Tidak Tepat dapat dilihat pada Gambar 4.

```

=== Confusion Matrix ===
  a  b  <-- classified as
79  4 | a = TEPAT
14  3 | b = TIDAK TEPAT
    
```

Gambar 4. Confution Matrix

Berdasarkan *confusion matrix* di atas didapatkan tingkat akurasi klasifikasi algoritma C4.5 sebesar 82%.

Accuracy adalah persentase dari total data uji coba yang dinyatakan benar ketika dilakukan proses identifikasi [10] Perhitungan *Accuracy* dapat dilihat pada rumus di bawah ini.

Tabel 3. Tabel Rumus Penilaian *accuracy* dengan konsep *confution matrix* [10]

	Diidentifikasi lulus tepat	Diidentifikasi lulus tidak tepat
Tepat	a	b
Tidak tepat	c	d

$$Accuracy = \frac{(a+d)}{(totalsample)} \times 100\% \dots\dots\dots (3)$$

$$Accuracy = \frac{79+3}{100} \times 100\%$$

$$= 0,82$$

Dengan perhitungan akurasi di atas menyatakan bahwa dalam kasus ini algoritma C4.5 dapat memprediksi ketepatan kelulusan dengan nilai akurasi 0,82% .

F. Penyebaran (Deployment)

Hasil dari penelitian ini adalah berupa analisa yang mengarah ke pengambilan keputusan. Hasil analisa diharapkan dapat dimanfaatkan oleh institusi perguruan tinggi sebagai bahan pertimbangan dalam menentukan langkah guna mengatasi permasalahan mahasiswa yang lulus tidak tepat waktu dimasa yang akan datang.

4. KESIMPULAN

Data mining dapat digunakan salah satunya untuk memprediksi ketepatan kelulusan mahasiswa. Pada penelitian ini, nilai *gain* tertinggi terletak pada pekerjaan orang tua, sehingga yang menjadi *root* pada pohon keputusan di penelitian ini adalah pekerjaan orang tua, kemudian dilanjutkan dengan atribut daerah, dan jenis sekolah asal.

Dari pengukuran kinerja algoritma yang telah dilakukan, dapat disimpulkan untuk kasus ini algoritma C4.5 memiliki nilai akurasi sebesar 82%. Dalam hal ini berarti penggunaan algoritma masih ditemukan tingkat kesalahan/*error* sebesar 18%.

Faktor yang menentukan kelulusan mahasiswa tidak hanya dilihat dari data diri mahasiswa saja, masih banyak faktor lain diantaranya adalah faktor akademis, faktor kondisi keluarga, dan dapat juga karena faktor psikologis. Jadi hasil analisis pada penelitian ini tidak dapat langsung dijadikan

keputusan untuk penerimaan calon mahasiswa baru bagi prodi Teknik Informatika Universitas Muhammadiyah Ponorogo.

5. DAFTAR PUSTAKA

[1] M. J. A. Berry and G. S. Linoff, *Data Mining Techniques For Marketing, Sales, Customer Relationship Management.*, Second Edi. United States of America: Wiley Computer Publishing.

[2] A. Basuki and I. Syarif, "Decision Tree," 2003.

[3] S. Soeparman, "Faktor-faktor yang Mempengaruhi Keberhasilan Studi Mahasiswa Penyandang Disabilitas," no. 19, pp. 12–19, 2011.

[4] D. H. Kamagi and S. Hansun, "Implementasi Data Mining dengan Algoritma C4 . 5 untuk Memprediksi Tingkat Kelulusan Mahasiswa," vol. VI, no. 1, pp. 15–20, 2014.

[5] M. Ridwan, H. Suyono, and M. Sarosa, "Penerapan Data Mining Untuk Evaluasi Kinerja Akademik Mahasiswa Menggunakan Algoritma Naive Bayes Classifier," vol. 7, no. 1, pp. 59–64, 2013.

[6] D. U. A, "Data Mining Untuk Menganalisa Prediksi Mahasiswa Berpotensi Non-Aktif Menggunakan Metode Decision Tree C4.5," 2010.

[7] P. C. Ncr *et al.*, "Crisp-dm 1.0," 2000.

[8] Y. N. Kunang *et al.*, "IMPLEMENTASI TEKNIK DATA MINING UNTUK MEMPREDIKSI TINGKAT KELULUSAN MAHASISWA PADA UNIVERSITAS BINA DARMA PALEMBANG," vol. 2013, no. semnasIF, pp. 1–8, 2013.

[9] -----, "ALGORITMA C4.5." .

[10] R. K. Amin, D. Indwiarti, M. Si, Y. Sibaroni, and S. Si, "IMPLEMENTASI KLASIFIKASI DECISION TREE DENGAN ALGORITMA C4 . 5 DALAM PENGAMBILAN KEPUTUSAN PERMOHONAN KREDIT OLEH DEBITUR Prodi Ilmu Komputasi Fakultas Informatika," vol. 2, no. 1, 2015.