

IMPLEMENTASI DATA MINING UNTUK MEMPREDIKSI TINGKAT KELULUSAN
MAHASISWA DENGAN MENGGUNAKAN ALGORITMA *DECISION TREE C5.0*

Fitria Hardianti Puspitaningrum*, **Andy Haryoko***, **Andik Adi Suryanto***, **Alfianisa Hanny Saputri***

* Universitas PGRI Ronggolawe

Correspondence Author: andyharyoko@gmail.com

Info Artikel :	ABSTRACT (in English)
Sejarah Artikel : Menerima : Revisi : Diterima : Online : Keyword : Classification, Method C5.0	<p><i>The development of information technology is increasingly rapidly affecting human life, so it demands high quality human resources. The education sector is one that is affected by the development of information technology and plays an important role in preparing quality human resources. The use of information technology can be utilized in monitoring student performance or academic achievement continuously, accurately and maximally, however monitoring academic achievement is currently still done manually. Then an automatic system was created to classify academic achievement using the C5.0 method. This research will explain how the C5.0 method accurately classifies student academic achievement. This study uses data from students of the Faculty of Engineering, PGRI Ronggolawe University Tuban. The research resulted in a desktop-based student academic achievement classification system using the VB programming language. Net and MySQL databases and also shows that the resulting academic achievement classification results are more accurate.</i></p>
	INTISARI (in Indonesia)
Kata Kunci : Klasifikasi, C5.0	<p>Perkembangan teknologi informasi yang semakin pesat sangat mempengaruhi kehidupan manusia, sehingga menuntut adanya sumber daya manusia yang berkualitas tinggi. Sektor pendidikan adalah salah satu yang mendapat pengaruh perkembangan teknologi informasi dan memegang peranan penting dalam menyiapkan sumber daya manusia yang berkualitas. Pemakaian teknologi informasi dapat dimanfaatkan dalam memantau kinerja atau prestasi akademik mahasiswa secara berkesinambungan, akurat dan juga maksimal namun pemantauan prestasi akademik saat ini masih dilakukan secara manual. Maka dibuatlah sistem otomatis untuk mengklasifikasi prestasi akademik dengan menggunakan metode C5.0. Pada penelitian akan dijelaskan bagaimana metode C5.0 mengklasifikasi prestasi akademik mahasiswa dengan akurat. Penelitian ini menggunakan data mahasiswa Fakultas Teknik Universitas PGRI Ronggolawe Tuban. Pada penelitian menghasilkan sistem klasifikasi prestasi akademik mahasiswa</p>

	berbasis desktop dengan menggunakan bahasa pemrograman VB. Net dan <i>database MySQL</i> dan juga menunjukkan bahwa hasil klasifikasi prestasi akademik yang dihasilkan lebih akurat.
--	---

1. PENDAHULUAN

Berdasarkan data dari *Forlap* Dikti, Fakultas Teknik Universitas PGRI Ronggolawe (Unirow) Tuban memiliki jumlah mahasiswa aktif 291 mahasiswa. Salah satu cabang ilmu komputer yang bisa melakukan prediksi adalah data *mining* karena manajemen membutuhkan aplikasi yang dapat memprediksi kelulusan mahasiswanya. Data *mining* adalah teknik untuk menemukan pola- pola yang ada dalam data sebagai sebuah alat untuk membantu menjelaskan data tersebut dan membuat prakiraan dari data itu (Witten & Eibe Frank, 2005). Pada kasus ini penggunaan ilmu data *mining* dapat digunakan sebagai pertimbangan dalam mengambil keputusan lebih lanjut tentang faktor yang mempengaruhi kelulusan khususnya faktor dalam data induk mahasiswa.

Pada penelitian menggunakan metode *Decision Tree*. Yang mana pohon keputusan merupakan struktur yang dapat digunakan untuk membagi kumpulan data menjadi himpunan-himpunan *record* yang lebih kecil. Data biasanya dinyatakan dalam bentuk tabel dan *record*. Pohon keputusan juga dapat digunakan untuk mengestimasi nilai dari variabel *continue*. Banyak algoritma yang dapat dipakai dalam pembentukan pohon keputusan, antara lain ID3, CART, C4.5 dan C5.0 (Larose, 2005). Algoritma C5.0 adalah algoritma yang terdapat dalam klasifikasi data *mining* di samping algoritma CART, yang khususnya diterapkan pada teknik *decision tree*.

Pada implementasi data *mining*, peneliti menggunakan algoritma C5.0 dalam menentukan prediksi kelulusan mahasiswa Fakultas Teknik Universitas PGRI Ronggolawe (Unirow) Tuban.

2. TINJAUAN PUSTAKA

Data *mining* adalah proses yang mempekerjakan satu atau lebih teknik pembelajaran komputer untuk menganalisis dan mengekstrak pengetahuan secara otomatis atau serangkaian proses untuk menggali nilai tambah dari suatu kumpulan data berupa pengetahuan yang selama ini tidak diketahui secara manual (Pramudiono, 2007). Metode pohon keputusan mengubah fakta yang sangat besar menjadi pohon keputusan yang merepresentasikan aturan. Dapat diekspresikan dalam bentuk bahasa basis data seperti *Structured Query Language* untuk mencari *record* pada kategori tertentu.

Algoritma C5.0 adalah salah satu algoritma yang terdapat dalam klasifikasi *data mining* di samping algoritma CART, yang khususnya diterapkan pada teknik *decision tree*. Secara heuristik akan dipilih atribut yang menghasilkan simpul yang paling bersih (*purest*).

Kelulusan mahasiswa adalah hasil akhir suatu proses akademik yang merupakan suatu keputusan di mana seorang mahasiswa itu dinyatakan telah memenuhi syarat dari berbagai macam persyaratan

seperti: persyaratan akademik dan administrasi yang diwajibkan / telah lunas semuanya dan berhak memperoleh gelar kesarjanaan.

3. METODE PENELITIAN

Sistem yang dikembangkan dalam Aplikasi Data Mining berbasis web. Pengguna dalam aplikasi adalah pihak penentu kebijakan dalam analisis tingkat kelulusan mahasiswa sehingga dapat diambil langkah-langkah strategis guna meningkatkan tingkat kelulusan dan juga mahasiswa sendiri yang dapat mengetahui prediksi kelulusannya secara dini agar lebih giat menentukan strategi kelulusannya. Untuk lebih memahami alur dari proses perhitungan data *mining* algoritma *decision tree* C5.0 dapat digambarkan dengan *flowchart* yang dapat dilihat pada Gambar 1.



Gambar 1. *Flowchart* metode / algoritma C5.0

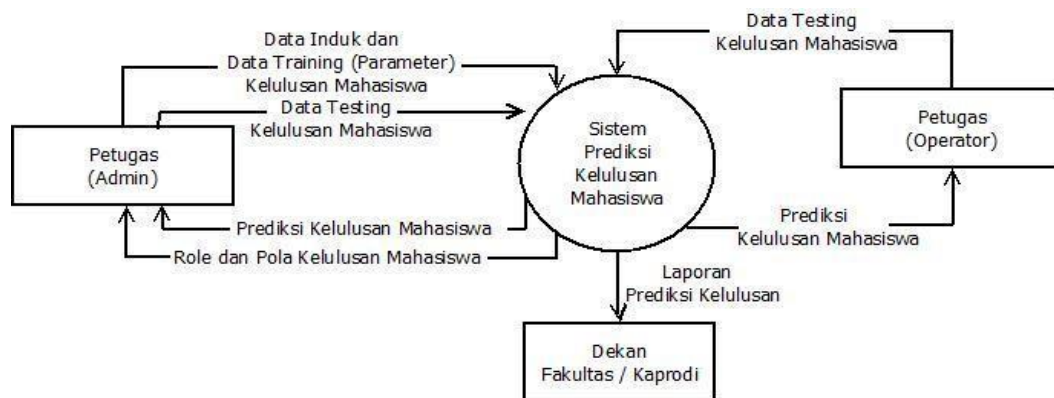
Data yang digunakan dalam penulisan penelitian terdiri dari beberapa *atribut* data dari mahasiswa yang dapat diolah untuk mengetahui pola – pola kelulusan. Data diperoleh di bagian akademik Fakultas Teknik Universitas PGRI Ronggolawe Tuban, pada program studi Teknik Industri dan

Teknik Informatika. Tabel 1. adalah beberapa contoh data menggunakan beberapa *attribut* yang mewakili penentuan kelulusan mahasiswa.

Tabel 1. Sample Data Mentah

NPM	Nama	Tahun Masuk	Tahun Lulus	IPK
1411070016	Yusdhitira Christiono	2007	2011	3.71
1411070012	Masrukan	2007	2012	2.99
1411080010	Khoirul Afandi	2008	2012	3.05
1411080035	Zainul Mustofa	2008	2012	2.89
1411080041	Ernawati	2008	2012	3.3
1412070065	Alfian Ardhianta	2007	2013	2.98
1412070076	Ali Usbah	2007	2013	3.1

Konteks diagram merupakan suatu teknik yang dapat digunakan untuk perancangan *output* yang ingin ditujukan pada aplikasi / sistem. Diagram konteks terdiri dari sebuah proses yang diberi label 0, sehingga dapat disebut dengan DFD Level 0 dan berinteraksi dengan beberapa *entiti* / entitas. Desain konteks diagram untuk aplikasi prediksi kelulusan mahasiswa yang akan dirancang pada Gambar 2.



Gambar 2. Konteks diagram

Perancangan *interface* berfungsi sebagai perantara antara pengguna dengan perangkat lunak. Agar perangkat lunak dapat digunakan dengan baik oleh pengguna perlu dibuat rancangan *interface* yang *user friendly*.

Tambah Data

NPM :

Nama Mahasiswa :

Tempat Lahir :

Tanggal Lahir :

Jenis Kelamin : Laki-laki Perempuan

Alamat :

Program Studi :

Angkatan :

Gambar 5. Rancangan *interface* tampilan “Tambah / Ubah Data Mahasiswa”

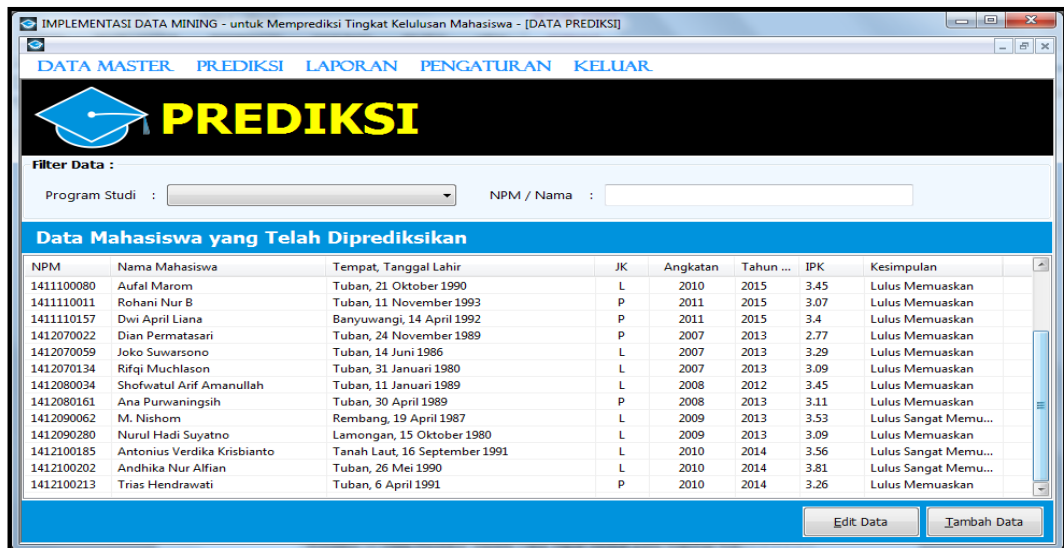
3 HASIL DAN ANALISA

Berdasarkan hasil analisis dan perancangan sistem yang telah dilakukan, maka dilakukan implementasi sistem menggunakan algoritma *decision tree* C5.0 ke dalam bentuk program dengan menggunakan bahasa pemrograman Visual Basic (VB) dengan MYSQL sebagai *datasenya*. Pada tahap ini dilakukan penerapan hasil perancangan *interface* ke dalam sistem yang dibangun dengan menggunakan bahasa pemrograman *Visual Basic (VB .Net)*.



Gambar 6. Interface tampilan awal aplikasi

Data – data hasil prediksi dapat dilihat atau ditampilkan pada menu “Prediksi > Data Prediksi” seperti yang dapat dilihat pada Gambar 7.



Gambar 7. Interface tampilan menu "Data Prediksi"

Untuk mencetak agar dapat dilakukan pelaporan klik menu Laporan -> Hasil Prediksi, kemudian pilih data dari prodi mana yang akan dicetak kemudian klik tombol cetak maka data hasil prediksi dapat dicetak maupun disimpan dalam bentuk Excel, menu cetak Laporan ini dapat dilihat pada Gambar 8. :



Gambar 8. Interface tampilan menu "Laporan"

Pada tahap pengujian ini, peneliti mendapatkan 109 (sebanyak 85 merupakan data *training*, dan 24 merupakan data *testing*) data dari Fakultas Teknik Universitas PGRI Ronggolawe Tuban berdasarkan atribut yang dibutuhkan. Menguji aplikasi dengan data *training* yang dipilih secara acak dan diperoleh data hasil pengujian yang dapat dilihat pada Tabel 2.

Tabel 2. Hasil prediksi aplikasi menggunakan 12 data *training* acak

No	NPM	Tahun Masuk	Tahun Lulus	IPK	Hasil Pengujian Aplikasi	Hasil Sebenarnya	Keterangan Prediksi
1	1411070002	2007	2011	3.19	Lulus Memuaskan	Lulus Memuaskan	Benar
2	1411070025	2007	2011	2.68	Lulus Memuaskan	Lulus Memuaskan	Benar
3	1411070036	2007	2011	3.58	Lulus Sangat Memuaskan	Lulus Sangat Memuaskan	Benar
4	1411070060	2007	2011	3.35	Lulus Memuaskan	Lulus Memuaskan	Benar

Berdasarkan Tabel 2. didapatkan :

Jumlah data 12

Jumlah data dengan keterangan "Benar" 12

$$\text{Persentase Kecocokan} = \frac{\text{Jumlah data dengan keterangan "Benar"}}{\text{Jumlah Data}} \times 100\%$$

$$= \frac{12}{12} \times 100\% = 100\%$$

Pada tahap pengujian menggunakan data *training* sebanyak 12 yang dipilih secara acak, peneliti mendapatkan tingkat akurasi aplikasi dari hasil prediksi kelulusan sebesar 100%.

4 KESIMPULAN

Telah diperoleh suatu aplikasi berbasis *desktop* untuk memprediksi kelulusan mahasiswa menggunakan metode data *mining* dengan algoritma *decision tree* C5.0. Algoritma *decision tree* C5.0 di dalam perhitungannya, mengeliminasi perhitungan - perhitungan yang tidak diperlukan dan juga adanya pohon keputusan maka sampel diuji hanya berdasarkan kriteria atau kelas tertentu. Secara keseluruhan, tingkat akurasi aplikasi prediksi kelulusan dengan menggunakan algoritma C5.0 di atas 70% dengan menggunakan beberapa data sebagai data untuk pengujian, yaitu 12 data dari data *training*, 24 data testing (data alumni angkatan 2010 dan 2011).

DAFTAR PUSTAKA

- Darmayuda, K., 2010. *Pemrograman Aplikasi Database dengan Microsoft Visual Basic.Net 2008*, Bandung: Informatika.
- Fajar, Astuti. 2013. *Data Mining Edisi Pertama*, Yogyakarta : Penerbit Andi Jogiyanto.
2005. *Analisis & Disain*, Yogyakarta : Penerbit Andi Yogyakarta.
- Jogiyanto. 1999. *Analisis & Desain Sistem Informasi: Pendekatan Terstruktur Teori dan Praktek Aplikasi Bisnis*, Yogyakarta : Andi.
- Jogiyanto. 2010. *Analisis & Disain*, Yogyakarta : Penerbit Andi Yogyakarta.
- Kusrini; Emha, Taufik L. 2009. *Algoritma Data Mining*, Yogyakarta : Penerbit Andi.
- Larose, Daniel T. 2005. *Discovering Knowledge in Data: An Introduction to Data Mining*, John Willey & Sons, Inc.
- Mardani. 2011. *Data Mining Algoritma Decision Tree C5.0*. <http://mardani-g4ul.blogspot.co.id/2011/12/contoh-soal-algoritma-c45.html> Diakses pada tanggal 26 Januari 2016
- Pramudiono, I. 2007. *Pengantar Data Mining : Menambang Permata Pengetahuan di Gunung Data*. <http://www.ilmukomputer.org/wpcontent/uploads/2006/08/iko-datamining.zip> Diakses pada tanggal 26 januari 2016
- Suharli, Suryanto. 2005. *Membangun Aplikasi Berbasis Windows Dengan Visual Basic .Net*, Jakarta : PT Elex Media Komputindo.
- Universitas PGRI Ronggolawe Tuban. 2012. *Buku Pedoman Unirow Tuban*, Tuban : Universitas PGRI Ronggolawe.
- Whitten, Jeffrey L. ; Dittman, Kevin C. ; Bentley, Lonnie D. 2004. *Metode Design Dan Analisa Sistem Bibliografi ed.6*, Yogyakarta : Andi Offset.
- Yuswanto., Subari. (2007). *Pemrograman Database Visual Basic .Net*. Jakarta : Prestasi Pustakaraya.