

# Penerapan Data Mining Algoritma C4.5 dalam Menentukan Rekam Jejak Kinerja Dosen STT Harapan Medan

Rismayanti

<sup>1</sup>Program studi teknik informatika,  
Universitas Harapan Medan  
risma.stth@gmail.com

Fera Damayanti

<sup>2</sup>Program studi teknik informatika,  
Universitas Harapan Medan  
feradamayantii@gmail.com

Khairunnisa

<sup>3</sup>Program studi teknik informatika,  
Universitas Harapan Medan  
khairunnisajv2@gmail.com

*Abstract*— Kinerja kerja dosen merupakan suatu hal yang sangat penting dalam upaya lembaga perguruan tinggi untuk mencapai tujuannya. Oleh karena itu perguruan tinggi perlu melakukan evaluasi dan penilaian yang menggambarkan kinerja dosen sebagai prestasi kerja yang dihasilkan oleh dosen tersebut sesuai dengan peranannya. Evaluasi dan penilaian kinerja dosen yang dilakukan di STT Harapan masih belum maksimal sehingga tingkat kesadaran setiap dosen untuk meningkatkan kinerjanya juga tidak maksimal. Penelitian ini bertujuan menerapkan *data mining* algoritma C4.5 dalam menentukan rekam jejak kinerja dosen berdasarkan 3 atribut *input* (publikasi, pengabdian dan pengajaran) dan 1 atribut *output* yang bernilai kinerja “kurang” “cukup” dan “tinggi”. Hasil penelitian berupa *decision tree* beserta *rules* yang memberikan informasi hasil evaluasi kinerja dosen STT Harapan Medan dalam melaksanakan tri darma perguruan tinggi.

*Kata Kunci*—*Data Mining; Algoritma C4.5; Decision Tree*

## I. PENDAHULUAN

Kinerja dosen pada suatu perguruan tinggi merupakan perilaku nyata yang ditampilkan setiap dosen sebagai prestasi kerja yang dihasilkan oleh dosen tersebut sesuai dengan peranannya. Penilaian kinerja dosen merupakan suatu proses di mana lembaga melakukan penilaian terhadap kinerja dosen dan mengevaluasi hasil pekerjaan dosen. Ada berbagai cara dalam melakukan evaluasi dan penilaian kinerja dosen, dan di STT Harapan evaluasi dan penilaian kinerja dosen yang dilakukan adalah dengan cara pengisian angket oleh mahasiswa yang menilai kinerja dosen dalam hal pengajaran. Hal tersebut tentunya masih belum maksimal dalam melakukan evaluasi dan penilaian terhadap kinerja dosen, akibat nya muncul permasalahan di mana

dosen di STT Harapan tidak memperoleh umpan balik yang maksimal dari lembaga atas hasil kerja mereka sehingga tingkat kesadaran setiap dosen untuk meningkatkan kinerjanya juga tidak maksimal baik itu dalam hal kinerja tri darma perguruan tinggi maupun prestasi hasil kerja lainnya.

*Data Mining* Algoritma C4.5 cocok digunakan untuk membentuk *decision tree* yang berguna untuk mengeksplorasi data dan menemukan hubungan yang tersembunyi antara sejumlah variabel input dengan satu variabel target. Penerapan *Data Mining* Algoritma C4.5 dalam menggali informasi terhadap data-data yang berkaitan dengan kinerja dosen di STT Harapan, yang meliputi data diri dosen, kinerja dosen di bidang tri darma perguruan tinggi akan berpotensi menghasilkan informasi baru yang berguna bagi STT Harapan dalam melakukan

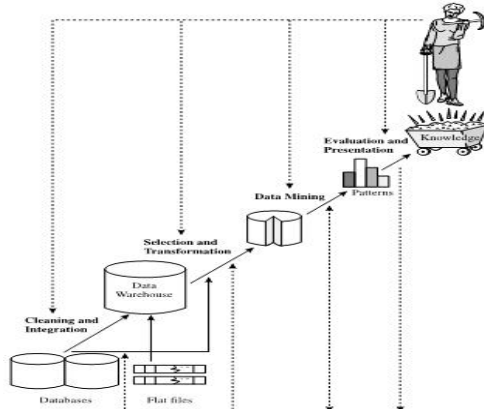
evaluasi dan penilaian terhadap kinerja dosen. Selain itu dosen diharapkan dapat lebih termotivasi dalam meningkatkan kinerja kerja sehingga dapat mempermudah dalam mengajukan usulan kenaikan jabatan fungsional maupun serdos.

## II. TINJAUAN PUSTAKA

### A. Data Mining

Secara sederhana *Data Mining* adalah penambangan atau penemuan informasi baru dengan mencari pola atau aturan tertentu dari sejumlah data yang sangat besar. *Data Mining* merupakan proses iteratif dan interaktif untuk menemukan pola atau model baru yang dapat digeneralisasi untuk masa yang akan datang, bermanfaat dan dapat dimengerti dalam suatu *database* yang sangat besar (*massive database*) [4]. *Data mining* disisi lain adalah kegiatan meliputi pengumpulan, pemakaian data historis untuk menemukan keteraturan, pola atau hubungan dalam set data berukuran besar [9].

Sebagai suatu rangkaian proses, *Data Mining* dapat dibagi menjadi beberapa tahap. Tahap-tahap tersebut bersifat interaktif di mana pemakai terlibat langsung atau dengan perantara *knowledge base* [10], terlihat pada gambar di bawah ini.



Gambar 1. Tahapan dalam *Data Mining*

### B. Pengelompokan Data Mining

*Data Mining* dibagi menjadi beberapa kelompok berdasarkan tugas yang dapat dilakukan yaitu: [1]

#### 1. Deskripsi

Menggambarkan pola dan kecenderungan yang terdapat dalam data, deskripsi dari pola dan kecenderungan sering memberikan penjelasan untuk suatu pola atau kecenderungan.

#### 2. Estimasi

Estimasi hampir sama dengan klasifikasi, kecuali variabel target estimasi lebih ke arah numerik dari pada ke arah kategori.

#### 3. Prediksi

Dalam prediksi, nilai dari hasil akan ada di masa mendatang.

#### 4. Klasifikasi

Dalam klasifikasi, terdapat target variabel kategori.

#### 5. Pengklusteran

Merupakan pengelompokan *record*, pengamatan, atau memperhatikan dan membentuk kelas objek-objek yang memiliki kemiripan.

#### 6. Asosiasi

Menemukan atribut yang muncul dalam satu waktu.

### C. Algoritma C4.5

Algoritma C4.5 memiliki kelebihan utama yaitu dapat menghasilkan model berupa *tree* atau aturan yang mudah diinterpretasikan, memiliki tingkat akurasi yang dapat diterima, dapat menangani atribut bertipe diskrit dan numerik [5]. Algoritma C4.5 merupakan algoritma yang digunakan untuk membentuk *decision tree* berdasarkan kriteria-kriteria pembentuk *decision* [7]. Menurut [2] Rumus algoritma C4.5 terbagi menjadi 2 rumus, yang pertama rumus untuk mencari nilai *Gain* yaitu:

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

Keterangan:

S = himpunan kasus

A = fitur

n = jumlah partisi atribut A

$|S_i|$  = proporsi  $S_i$  terhadap S

|S| = jumlah kasus dalam S

Dan rumus yang kedua adalah untuk mencari nilai *Entropy*, yaitu:

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

Keterangan:

S = himpunan kasus

n = jumlah partisi S

$P_i$  = proporsi  $S_i$  terhadap S

### D. Decision Tree

*Decision tree* adalah struktur *flowchart* yang mempunyai *tree* (pohon), di mana setiap simpul internal menandakan suatu tes atribut, setiap cabang



“tinggi”. Penentuan nilai kinerja dilihat berdasarkan *score* publikasi, *score* pengabdian dan *score* pengajaran.

*Score* publikasi diperoleh dari jumlah banyaknya publikasi dosen dikalikan dengan nilai angka kredit publikasi yang terdapat pada pedoman penilaian angka kredit kenaikan pangkat/jabatan akademik dosen. Kemudian *score* pengabdian diperoleh dari jumlah pengabdian masyarakat yang telah dilakukan dosen STT Harapan Medan. Dan untuk *score* pengajaran, *score* diperoleh dari penjumlahan nilai rata-rata *point* pengajaran A sd H pada masing-masing dosen. Berdasarkan hal tersebut maka dapat dilakukan transformasi pada data kinerja 85 dosen di STT Harapan Medan sebagai berikut pada tabel 3.

Tabel 3. Transformasi Data Kinerja Dosen STT Harapan Medan

No	Inisial Dosen	Score Publikasi	Score Pengabdian	Score Pengajaran	Kinerja
1	AKH	tidak ada	Kurang	baik	kurang
2	PUT	cukup	Cukup	baik	cukup
3	HSF	cukup	Cukup	baik	cukup
4	BON	cukup	Cukup	baik	cukup
5	DOD	tinggi	Tinggi	sangat baik	tinggi
6	ARZ	tidak ada	Kurang	baik	kurang
7	ARI	tinggi	Cukup	baik	tinggi
8	IFS	tinggi	Cukup	baik	tinggi
9	RKO	cukup	tidak ada	baik	kurang
10	HER	kurang	Kurang	baik	kurang
11	RIN	kurang	Kurang	baik	kurang
12	DVH	tidak ada	Kurang	baik	kurang
13	YUS	kurang	Kurang	baik	kurang
14	YYN	tinggi	Cukup	baik	tinggi
15	NNA	tinggi	Cukup	baik	tinggi
16	JND	cukup	tidak ada	baik	kurang
17	ABM	kurang	Kurang	baik	kurang
18	FZI	tidak ada	Kurang	baik	kurang
19	SPN	cukup	tidak ada	baik	kurang
20	SRT	kurang	Kurang	baik	kurang
.	.	.	.	.	.
.	.	.	.	.	.
.	.	.	.	.	.
.	.	.	.	.	.
.	.	.	.	.	.
81	YET	tidak ada	Kurang	baik	kurang
82	YUN	kurang	Kurang	baik	kurang
83	SFL	tidak ada	Kurang	baik	kurang
84	ARS	cukup	Cukup	baik	cukup
85	LIA	kurang	Kurang	baik	kurang

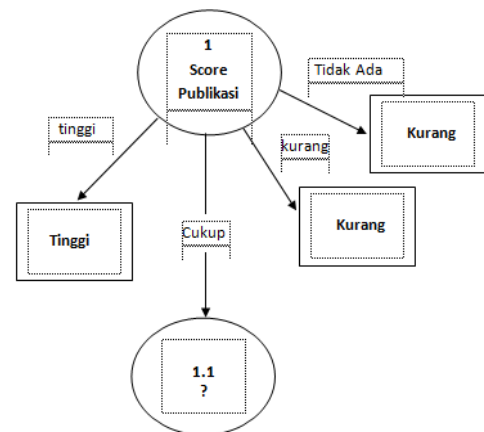
E. Merancang *Decision Tree* dengan Algoritma C4.5

Berdasarkan tabel 3, maka dilakukan proses perhitungan *Gain* dan *Entropy* untuk menentukan akar (*root*) dari pohon keputusan dalam membantu menentukan rekam jejak kinerja dosen. Di bawah ini adalah hasil perhitungan *Gain* dan *Entropy* node 1, pada tabel 4.

Tabel 4. Hasil Perhitungan *Gain* dan *Entropy* pada Node 1.

Node	Atribut	Kelas	Jumlah Kasus (S)	Kurang	Cukup	Tinggi	Entropy	Gain	
1	Total		85	55	17	13	1,285069		
	Score							1,025785	
	Publikasi	Tidak Ada		15	15	0	0	0	
		kurang		33	33	0	0	0	
		Cukup		24	8	16	0	0,918296	
		tinggi		13	0	0	13	0	
	Score							0,755575	
	Pengabdian	Tidak Ada		5	5	0	0	0	
		kurang		44	44	0	0	0	
		Cukup		29	3	17	9	1,314151	
	Score	tinggi		7	3	0	4	0,985228	
	Pengajaran								0,100243
		Cukup		2	2	0	0	0	
		baik		81	54	16	11	1,243336	
	sangat baik		2	0	0	2	0		

Dari hasil tabel 4, dapat diketahui atribut dengan *Gain* tertinggi adalah *score* publikasi, yaitu 1,025785. Dengan demikian *score* publikasi dapat dijadikan sebagai *node* akar. Dari hasil perhitungan nilai *Gain* dan *Entropy* pada *node* 1, maka dapat digambarkan *decision tree* sementara terlihat pada gambar 1.



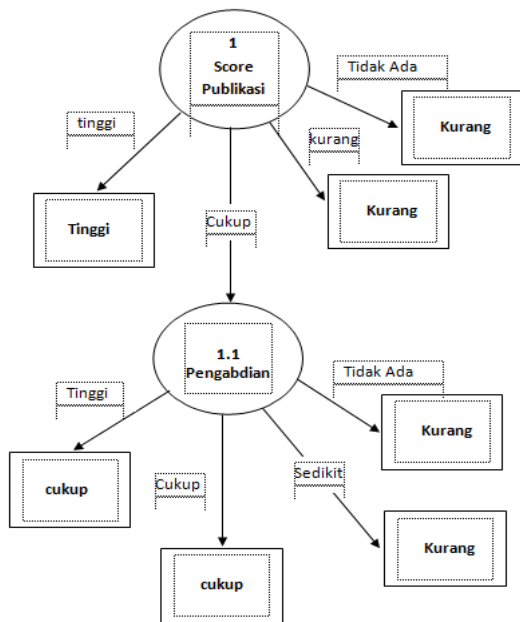
Gambar 1. *Decision Tree* Node 1 Penentuan Rekam Jejak Kinerja Dosen

Ada 4 nilai atribut dari “*score* publikasi” yaitu tidak ada, kurang, cukup dan tinggi. Berdasarkan data kinerja 85 dosen, maka *score* publikasi yang bernilai “tidak ada” dan “kurang” sudah mengklasifikasikan nilai menjadi 1 keputusan yaitu kinerja “kurang”. Kemudian untuk *score* publikasi bernilai “tinggi” juga sudah mengklasifikasikan nilai menjadi 1 keputusan yaitu kinerja “tinggi”. Sedangkan “*score* publikasi” yang bernilai “cukup” belum mengklasifikasikan nilai menjadi 1 keputusan, sehingga perlu dilakukan proses perhitungan *Gain* dan *Entropy* lebih lanjut, terlihat pada tabel 5.

Tabel 5. Hasil Perhitungan *Gain* dan *Entropy* pada *Node* 1.1

Node	Atribut	Kelas	Jumlah Kasus (S)	Kurang	Cukup	Tinggi	Entropy	Gain	
1	Total	Publikasi-							
		Cukup	24	8	16	0	0,918296		
	Score	Pengabdian	Tidak Ada	5	5	0	0	0	0,918296
			kurang	3	3	0	0	0	
			Cukup	16	0	16	0	0	
			tinggi	0	0	0	0	0	
	Pengajaran								0
			Cukup	0	0	0	0	0	
			baik	24	8	16	0	0,918296	
			sangat baik	0	0	0	0	0	

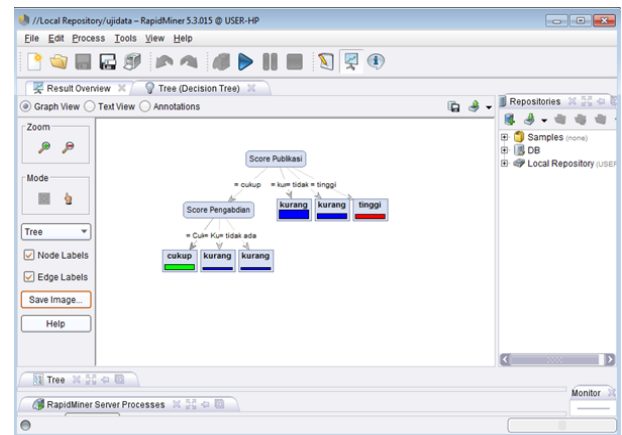
Berdasarkan tabel 5 dapat diketahui bahwa atribut dengan *Gain* tertinggi adalah *score* pengabdian yaitu 0,918296. Dengan demikian, *score* pengabdian dapat menjadi *node* cabang dari *score* publikasi bernilai “cukup”. Untuk atribut *score* pengabdian memiliki 4 nilai atribut yaitu tidak ada, kurang, cukup dan tinggi. Ke-4 nilai atribut tersebut sudah mengklasifikasikan kasus menjadi satu keputusan, sehingga dapat digambarkan *decision tree* akhir dalam menentukan rekam jejak kinerja dosen STT Harapan Medan, terlihat pada gambar 2.



Gambar 2. *Decision Tree* Node 1.1 Penentuan Rekam Jejak Kinerja Dosen STT Harapan Medan

## VI. IMPLEMENTASI

Implementasi dilakukan menggunakan salah satu *software Data Mining* yaitu *Rapid Miner 5.3.015*. Semua atribut *input* dan atribut *output* disimpan dalam format *xlsx*, kemudian di *import* ke *software Rapid Miner 5.3.015* dan menghasilkan *decision tree* sebagai berikut, pada gambar 3.



Gambar 3. *Decision Tree* Penentuan Rekam Jejak Kinerja Dosen STT Harapan Medan

Berdasarkan pengujian yang telah dilakukan, maka rules yang diperoleh dari hasil analisis penentuan rekam jejak kinerja dosen STT Harapan Medan menggunakan algoritma *C4.5* adalah sebagai berikut:

1. IF *Score* Publikasi “Kurang” THEN Kinerja “Kurang”
2. IF *Score* Publikasi “Tidak Ada” THEN Kinerja “Kurang”
3. IF *Score* Publikasi “Tinggi” THEN Kinerja “Tinggi”
4. IF *Score* Publikasi “Cukup” And *Score* Pengabdian “Cukup” THEN Kinerja “Cukup”
5. IF *Score* Publikasi “Cukup” And *Score* Pengabdian “Kurang” THEN Kinerja “Kurang”
6. IF *Score* Publikasi “Cukup” And *Score* Pengabdian “Tidak Ada” THEN Kinerja “Kurang”

## V. KESIMPULAN

Setelah melakukan analisis, perancangan, implemmentasi beserta pengujian dalam menerapkan algoritma *C4.5* untuk menentukan rekam jejak kinerja dosen STT Harapan, maka diperoleh kesimpulan bahwa:

1. Hasil *decision tree* penentuan rekam jejak kinerja dosen STTH Medan menghasilkan keputusan bahwa dosen yang berkinerja tinggi adalah dosen yang memiliki *score* publikasi tinggi.
2. Dosen yang berkinerja “cukup” adalah dosen yang memiliki *score* publikasi dan *score* pengabdian bernilai cukup.
3. Dosen yang berkinerja “kurang” adalah dosen yang memiliki *score* publikasi dan *score* pengabdian “Tidak Ada” dan “Kurang”.

#### REFERENSI

- [1] Gunadi, Goldie, Indra Sensue, “Penerapan Data Mining Market Basket Analysis Terhadap Data Penjualan Produk Buku dengan Menggunakan Algoritma Aprior dan Frequent Pattern Growth(Fb:Growth): Studi Kasus Percetakan Gramedia”, Jurnal Telematika M.Kom Vol. 4, No.1, 2012.
- [2] Syahputra, Guntur, “ Penerapan Algoritma C4.5 dalam Analisa Kelayakan Penerima Bonus Tahunan Pegawai(Studi Kasus: PT. Multi Pratama Nauli Medan)”, Jurnal Mantik Penusa, Vol. 16, No. 2, Desember 2014.
- [3] Ina, Wenefrida Tulit, “Klasifikasi Data Rekam Medis Berdasarkan Kode Penyakit International Menggunakan Algoritma C4.5”, Jurnal Media Elektro, Vol. 1, No. 3, 2013.
- [4] Luvia, Yuni Sara, Dedi Hartama n dkk, “ Penerapan Algoritma C4.5 untuk Klasifikasi Predikat Keberhasilan Mahasiswa di Amik Tunas Bangsa”, JURASIK(Jurnal Riset Sistem Informasi & Teknik Informatika), Vol. 1, No. 1, Juli 2016.
- [5] Rismayanti, “Implementasi Algoritma C4.5 untuk Menentukan Penerima Beasiswa di STT Harapan Medan”, Jurnal Media Infotama, Vol. 12, No. 2, September 2016.
- [6] Rismayanti, “Decision Tree Penentuan Masa Studi Mahasiswa Prodi Teknik Informatika(Studi Kasus: Fakultas Teknik dan Komputer Univ. Harapan Medan)”, QUERY: Jurnal Sistem Informasi, Vol. 2, No. 1, April 2018.
- [7] Wajhillah, Rusda, Ita Yulianti, “Penerapan Algoritma C4.5 untuk Prediksi Penggunaan Jenis Kontrasepsi Berbasis Web”, KLIK(Kumpulan Jurnal Ilmu Komputer)”, Vol. 4, No. 2, September 2017.
- [8] Sunjana, “Aplikasi Data Mining Data Mahasiswa dengan Metode Klasifikasi Decision Tree “, SNATI Yogyakarta, 2010.
- [9] Swastina, Liliana, “Penerapan Algoritma C4.5 untuk Penentuan Jurusan Mahasiswa”, Jurnal Gema Aktualita, Vol. 2, No. 1, 2013.
- [10] Yunus, Mahmud, Harry Soekotjo n dkk, “ SPK Pemilihan Calon Pendorong Darah Potensial dengan Algoritma C4.5 dan Fuzzy Tahani”, Jurnal EECCIS, 2014, Vol.8 No1.