

## ANALISIS PREDIKSI TINGKAT KEBERHASILAN STUDI DI PERGURUAN TINGGI BERDASARKAN PRESTASI DENGAN MENGGUNAKAN METODE ITERATIVE DICHOTOMIZER 3(ID3) (STUDI KASUS: FTI UKSW)

**Andi Kristanto**

Fakultas Teknologi Informasi, Program Studi Teknik  
Informatika Universitas Kristen Satya Wacana  
Email: andi.kris28@gmail.com

**Eko Sedyono**

Fakultas Teknologi Informasi  
Universitas Kristen Satya Wacana  
Email: ekosed1@yahoo.com

### ABSTRAK

Prestasi akademis merupakan hasil dari proses pembelajaran yang dilakukan selama semester- semester yang telah berlangsung, yang diukur dari nilai dan indeks prestasi kumulatif. Dimana FTI-UKSW memiliki jumlah mahasiswa yang banyak sehingga perlu untuk melihat prestasi mahasiswanya. Dengan melakukan prediksi keberhasilan studi dapat dilihat tingkat mahasiswa yang berprestasi dan faktor-faktor yang dapat menjadi indikator dalam penentuan keberhasilan studi tersebut. Data yang digunakan dalam penelitian ini menggunakan data yang diterima dari Biro Administrasi dan *Registrasi Akademik (BARA)*, yaitu data admisi/rapot SLTA, dan data mahasiswa aktif yaitu mahasiswa yang berstatus kuliah. Dengan menggunakan *spreadsheet* sebagai alat bantu untuk melakukan pengolahan data yang menerapkan rumusan metode ID3. Dengan menggunakan metode ID3 dapat membantu membuat suatu keputusan yang berbentuk diagram pohon keputusan yang melalui indikator-indikator, dan persentase dari masing-masing indikator. Hasil dari penelitian analisis ini nantinya dapat digunakan untuk mengetahui dari ketiga indikator, mana yang menjadi komponen utama dalam menentukan apakah mahasiswa tersebut berprestasi.

**Kata kunci:** prestasi; keberhasilan studi; data admisi; data mahasiswa aktif; ID3; *spreadsheet*.

### ABSTRACT

*Academic achievement is the result of the learning process undertaken during the semesters that have taken place, which is measured by the cumulative grades and grades. Where the FTI-SWCU has a large number of students so it is necessary to see student achievements. By predicting the success of the study can be seen the level of students who excel and the factors that can be indicators in determining the success of the study. The data used in this study uses data received from the Administrative Bureau and Academic Registration (BARA), namely data admissions / high school report cards, and active student data, namely students with college status. By using spreadsheets as a tool to do data processing that applies the ID3 method formula. Using the ID3 method can help make a decision in the form of a decision tree through indicators, and the percentage of each indicator. The results of this analysis study can later be used to find out from the three indicators, which are the main components in determining whether the student is performing.*

**Keywords:** *achievement; the success of the study; data procedures; data of active students; spreadsheet; ID3.*

### 1. PENDAHULUAN

Dalam proses pembelajaran, mahasiswa merupakan komponen yang terpenting didalam maupun diluar kelas. Dimana dengan melihat nilai indeks prestasi mereka, dapat digunakan untuk penentuan prestasi [1]. Dengan demikian faktor utama dalam menentukan prestasi dapat dilihat dari bermacam-macam faktor, baik dari nilai IPK, kegiatan yang diikuti selama masa studi, keaktifan, dll.

Indeks Prestasi merupakan hasil pembelajaran mahasiswa dalam tiap-tiap semester, untuk melihat hasil dalam pembelajaran [2]. Sehingga dalam proses pembelajaran, IP penting dalam membantu proses pembelajaran dan penentuan prestasi.

Keberhasilan studi dalam masa-masa perkuliahan dapat dilihat dari perkembangan prestasi pada tiap mahasiswa, baik dengan melihat hasil studi dari tiap-tiap semester yang sudah ditempuh, baik dari *Indeks Prestasi Kumulatif* (IPK), dan keaktifan mahasiswa dalam proses pembelajaran dikelas.

Pada saat ini muncul permasalahan yang berkaitan dengan keberhasilan studi khususnya di Fakultas Teknologi Informasi, Universitas Kristen Satya Wacana. Dengan banyaknya lowongan pekerjaan pada perusahaan-perusahaan di-indonesia sendiri, membuat banyak mahasiswa yang berusaha serta berlomba-lomba untuk mendapatkan pekerjaan pada perusahaan tersebut, yang sesuai dengan bidang minat atau keahlian mereka. Persaingan dalam mendapatkan tempat pekerjaan sendiri membuat penelitian ini dikembangkan, dengan tujuan untuk mengetahui persentase keberhasilan studi mahasiswa dari indikator-indikator yang digunakan dalam penelitian di Fakultas Teknologi Infomasi.

Dalam penelitian yang berjudul “Sistem Pendukung Keputusan Prediksi Kecepatan Studi Mahasiswa Menggunakan Metode ID3” yang diteliti oleh Ronny Ardi Giovani, Paulus Mudjihartono dan Pranowo. Pada penelitian ini membahas tentang pembuatan sistem pendukung keputusan yang digunakan untuk menganalisis dan mengolah data dalam menentukan kecepatan tingkat studi mahasiswa dalam menempuh proses perkuliahan, dan dimana sistem ini mampu digunakan untuk menampung semua data yang telah ada dan dapat digunakan dalam proses penentuan kecepatan studi mahasiswa [2].

Algoritma ID3 (*Iterative Dichotomizer 3*) adalah algoritma yang digunakan untuk menghasilkan suatu keputusan yang berbentuk sebuah pohon keputusan. Dengan menggunakan algoritma ID3 dibentuklah sebuah pohon keputusan untuk menentukan persentase keberhasilan studi mahasiswa di Fakultas Teknologi Informasi, Universitas Kristen Satya Wacana dengan menggunakan data *mining* yang berisi hasil dari nilai raport SMA dan hasil dari proses pembelajaran mahasiswa dan mahasiswi selama masa studi atau perkuliahan.

## 2. METODOLOGI PENELITIAN

Dalam metodologi penelitian merupakan hal terpenting dalam melakukan proses penelitian, karena dengan adanya metodologi penelitian tersebut dapat membantu melihat proses atau langkah-langkah yang akan dilakukan dalam melakukan pekerjaan dalam penelitian [4].

ID3 adalah metode yang digunakan untuk mengklasifikasikan dengan membentuk pohon keputusan yang digunakan untuk menentukan atau memprediksi pada suatu penelitian. Pada metode ini menggunakan konsep dari *entropy informasi*. Secara ringkas, langkah kerja Algoritma ID3 dapat digambarkan sebagai berikut:

Penghitungan *Information Gain* dari setiap *atribut* dengan menggunakan :

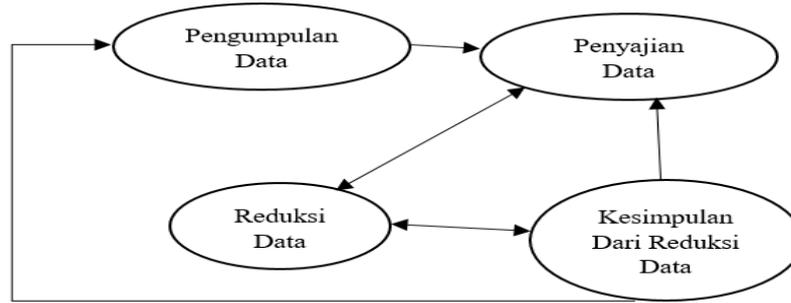
$$Gain(S, A) = Entropy(S) - \sum_{v \in nilai(A)} \frac{|S_v|}{|S|} Entropy(S_v) \quad (1)$$

$$Entropy(S) = -P_+ \log_2 P_+ - P_- \log_2 P_- \quad (2)$$

Dimana dengan menggunakan konsep *entropy informasi* ini dapat membantu dalam proses perhitungan atau pengolahan data yang telah dibuat ke dalam kategori masing-masing.

Pada penelitian ini menggunakan software RapidMiner yang nantinya akan diterapkan dengan menggunakan module metode ID3 yang digunakan untuk membuat sebuah pohon keputusan yang berguna dalam membantu menentukan apakah mahasiswa tersebut proses bagaimana dalam menentukan prestasi berdasarkan faktor-faktor atau indikator-indikator permasalahan dari penelitian pada kasus ini. Algoritma pada metode ini menggunakan konsep dari entropi informasi.

Adapun pada penelitian ini data yang diperoleh adalah data *kuantitatif*. Data *kuantitatif* adalah data yang didapat berbentuk dari tempat penelitian baik dengan cara wawancara, dengan angket atau data *excel* yang berupa angka. Adapun tahapan atau langkah – langkah ditunjukkan pada gambar 1.



**Gambar 1. Tahapan Dalam Penelitian**

Dimana pada gambar 1 dapat dijelaskan sebagai berikut:

- Tahap pengumpulan data merupakan tahap dimana dalam penelitian ini melakukan pengumpulan data dengan cara meminta data ke tempat penelitian, dimana pada penelitian ini tempat yang dituju yaitu Bagian *Administrasi Registrasi Akademik* (BARA), Bagian Nilai dan Ijazah. Dimana pada bagian nilai dan ijazah memberikan data yang berupa data mahasiswa yang masih dalam proses studi, serta data mahasiswa yang sudah lulus.
- Tahap reduksi data merupakan proses pemilahan data dan melakukan pengambilan untuk digunakan sample dalam penelitian.
- Tahap penyajian data merupakan tahapan penyajian setelah proses pemilahan dan pengambilan data sample dan data uji telah selesai.
- Tahap penarikan kesimpulan merupakan tahap akhir yang dilakukan dalam proses menganalisa data, yang dimana data yang sudah diolah dapat disimpulkan sesuai dengan penelitian.

Pada tahapan awal penganalisan keberhasilan studi ini sendiri dapat dilakukan dengan melakukan pengambilan data dengan cara permintaan data langsung ke Bagian *Administrasi Registrasi Akademik* (BARA), bagian nilai dan ijazah. Adapun data yang diperoleh untuk penganalisan ini merupakan data yang berupa angka yang sudah diklasifikasikan, baik nilai raport setelah lulus SMA atau SMK, dan juga hasil nilai keseluruhan selama studi atau yang dikenal dengan nama *Indeks Prestasi Kumulatif* (IPK). Sehingga data yang digunakan dalam penganalisan ini adalah data nilai raport dan data IPK pada tiap-tiap mahasiswa.

Setelah proses tahapan pengumpulan data selesai maka dilanjutkan dengan proses pengelompokan data yang dimana data tersebut dikelompokkan berdasarkan keperluan dalam penganalisan tersebut.

Dengan menggunakan metode *iterative dichotomizer* (ID3) ini, data yang sudah dikelompokkan dapat mulai diolah dengan menggunakan perhitungan *information gain* dan *entropy* serta menggunakan *spreadsheet* untuk mengolah data dengan menerapkan perhitungan *information gain* dan *entropy*. Adapun setelah proses pada tahap perhitungan selesai dapat di eksekusi dengan menggunakan *software RapidMiner* dengan menggunakan *module* ID3 yang dapat menghasilkan sebuah keputusan dalam menentukan keberhasilan studi ini, dengan berbentuk sebuah pohon keputusan.

### 3. HASIL DAN PEMBAHASAN

Dalam hasil dan pembahasan pada penelitian ini, data mahasiswa yang diambil dari BARA, Bag. Nilai dan Ijazah adalah data admisi, data mahasiswa aktif, dan data lulusan yang akan diolah untuk membantu penelitian selanjutnya. Dengan menggunakan *spreadsheet* dan dengan menggunakan metode *Iterative Dichotomizer 3* atau *ID3*. Data yang sudah diklasifikasikan dapat dibentuk ke dalam tabel *dataset*. Adapun sebagian sampel data nilai raport yang diterima dari BARA sebagai berikut:

**Tabel 1. Tabel data sampel nilai raport**

Nilai Raport							
Ing 1	Ing 2	Ing 4	Ing 5	Mat 1	Mat 2	Mat 4	Mat 5
84	81	84	78	84	82	71	68
58	59	70	73	61	65	48	70
63	63	62	65	62	63	68	72
86	89	84	70	67	72	70	64
63	65	69	66	64	62	77	65
85	90	86	90	72	74	62	68
75	72	74	70	53	56	58	66

<i>Nilai Raport</i>							
<i>Ing 1</i>	<i>Ing 2</i>	<i>Ing 4</i>	<i>Ing 5</i>	<i>Mat 1</i>	<i>Mat 2</i>	<i>Mat 4</i>	<i>Mat 5</i>
63	66	63	69	66	64	59	62
74	81	71	82	65	71	73	71
70	80	70	73	70	70	75	85
73	76	78	78	65	65	70	75
73	81	76	76	62	54	73	75
70	71	73	73	65	65	65	67
71	74	69	65	67	62	62	65
70	70	65	65	60	72	61	61
76	65	71	70	67	71	67	71
60	60	60	67	50	74	69	75
60	60	60	70	80	74	67	73

Pada tabel 1 merupakan tabel data sampel nilai raport yang diterima dari bagian BARA, dimana nantinya data yang akan digunakan berjumlah 100 data. Adapun data mahasiswa Lulus yang diterima dari bagian nilai dan ijazah sebagai berikut:

**Tabel 2. Tabel sampel data lulusan**

<i>Tgl.Lulus</i>	<i>Pria/Wanita</i>	<i>SKS</i>	<i>AK</i>	<i>IPK</i>
7-Apr-17	Pria	145	427	2.94
7-Apr-17	Pria	146	462	3.16
7-Apr-17	Pria	144	451.5	3.13
7-Apr-17	Pria	144	450.5	3.12
7-Apr-17	Pria	145	501.5	3.45
7-Apr-17	Pria	146	478.5	3.27
7-Apr-17	Pria	148	542	3.66
7-Apr-17	Pria	144	459.0	3.18
7-Apr-17	Pria	146	568	3.89
7-Apr-17	Wanita	146	571.5	3.91
7-Apr-17	Wanita	155	533	3.43
7-Apr-17	Wanita	145	554	3.82
7-Apr-17	Pria	144	529	3.67
7-Apr-17	Pria	144	524	3.63
7-Apr-17	Pria	149	560.5	3.76
7-Apr-17	Pria	146	520	3.56
7-Apr-17	Pria	144	475.5	3.30
7-Apr-17	Pria	146	484	3.31
8 Mei 2017	Pria	144	469	3.25
8 Mei 2017	Pria	146	483	3.30
8 Mei 2017	Pria	146	432.5	2.96
8 Mei 2017	Pria	147	497.5	3.38
8 Mei 2017	Pria	147	540	3.67
8 Mei 2017	Wanita	147	499	3.39
8 Mei 2017	Pria	146	449.5	3.07
8 Mei 2017	Pria	147	495	3.36
8 Mei 2017	Pria	147	516.5	3.51
8 Mei 2017	Pria	147	522	3.55
8 Mei 2017	Pria	146	472.5	3.23
8 Mei 2017	Pria	144	521.5	3.62
8 Mei 2017	Wanita	147	550.5	3.74
8 Mei 2017	Pria	146	537.5	3.68
5 Juni 2017	Pria	144	491	3.40

Tabel 2 merupakan sampel data yang diterima dari bagian nilai dan ijazah, serta data yang diberikan dari tahun 2010 – 2017.

**Tabel 3. Tabel perhitungan manual dataset mahasiswa dari A-CV**

<i>Nama</i>	<i>Rapot</i>	<i>IPK</i>	<i>Status</i>	<i>Prestasi</i>
A	Bagus	Sedang	Lulus	Ya
B	Cukup	Sedang	Lulus	Ya
C	Cukup	Tinggi	Lulus	Ya
D	Bagus	Tinggi	Lulus	Ya
E	Cukup	Tinggi	Lulus	Ya
F	Bagus	Tinggi	Lulus	Ya
G	Cukup	Tinggi	Lulus	Ya
H	Cukup	Tinggi	Lulus	Ya
I	Bagus	Tinggi	Lulus	Ya
J	Bagus	Tinggi	Lulus	Ya
...	...	...	...	...
CT	Bagus	Tinggi	Lulus	Ya
CU	Bagus	Tinggi	Belum	Ya
CV	Kurang	Sedang	Belum	Tidak

Keterangan :

	Data Latih
	Data Uji

Penjelasan pada masing-masing tabel, pada tabel berwarna kuning adalah merupakan data latih, dimana data latih memiliki total data berjumlah 97 data. Tabel berwarna hijau adalah data uji yang memiliki total data uji berjumlah 3 data yang akan digunakan dalam menguji data latih. Adapun langkah pengerjaan atau langkah perhitungan sebagai berikut:

### 3.1 Langkah Pertama

Dengan menggunakan rumus *Entropy* yang dapat dijabarkan atau dirumuskan sebagai berikut:

$$Entropy (S) = -(Ya/Semua \text{ kasus}) * \log_2(Ya/Semua \text{ kasus}) + -(tidak/Semua \text{ kasus}) * \log_2(tidak/Semua \text{ kasus}).$$

**Tabel 4. Tabel menentukan *entropy* untuk keseluruhan kasus**

Jumlah kasus	97
Jumlah Kasus dengan jawaban Ya	84
Jumlah Kasus dengan jawaban Tidak	13

Dari rumus *entropy* diatas dapat dijadikan untuk melakukan perhitungan sebagai berikut:  
 $(-(84/97) * \log_2(84/97) + -(13/97) * \log_2(13/97))$   
 $= 0.568362534$

Dimana pada hasil 0.568362534 akan masuk ke dalam nilai *entropy* pada baris Total, mencari nilai *Entropy* dari tiap masing-masing atribut yaitu: Nilai Rapot, Nilai IPK, Status dapat diterapkan pada tabel 5.

**Tabel 5. Tabel perhitungan pencarian nilai *Entropy* dan *Gain***

	<i>Jumlah</i>	<i>Ya</i>	<i>Tidak</i>	<i>Entropy</i>	<i>Gain</i>
Total	97	84	13	0.568363	
Nilai Rapot					-0.00428
Bagus	65	57	8	0.538141	
Cukup	28	24	4	0.591673	
Kurang	4	2	2	1	
Nilai IPK					0.423172
Tinggi	62	62	0	0	

	<i>Jumlah</i>	<i>Ya</i>	<i>Tidak</i>	<i>Entropy</i>	<i>Gain</i>
	Sedang	21	20	1	0.276195
	Rendah	14	12	2	0.591673
Status					0.090383
	Lulus	48	47	1	0.146094
	Belum	49	37	12	0.803091

Dengan rumus *entropy* digunakan untuk menghitung total nilai masing-masing *entropy* pada IPK, Rapot dan Status, sehingga dapat dijadikan untuk melakukan perhitungan sebagai berikut:

- a. Rapot: Bagus  
Jumlah data yang yang bernilai bagus secara keseluruhan berjumlah: 65.  
Jumlah data yang bernilai Ya: 57.  
Jumlah data yang bernilai Tidak: 8.  
 $(-(57/65) * \log_2(57/65) + -(8/65) * \log_2(8/65))$   
 $= 0.538141$
- b. Rapot: Cukup  
Jumlah data yang yang bernilai cukup secara keseluruhan berjumlah: 28.  
Jumlah data yang bernilai Ya: 24.  
Jumlah data yang bernilai Tidak: 4.  
 $(-(24/28) * \log_2(24/28) + -(4/28) * \log_2(4/28))$   
 $= 0.591673$
- c. Rapot: Kurang  
Jumlah data yang yang bernilai kurang secara keseluruhan berjumlah: 4.  
Jumlah data yang bernilai Ya: 2.  
Jumlah data yang bernilai Tidak: 2.  
 $(-(2/4) * \log_2(2/4) + -(2/4) * \log_2(2/4))$   
 $= 1$
- d. IPK: Tinggi  
Jumlah data yang yang bernilai tinggi secara keseluruhan berjumlah: 62.  
Jumlah data yang bernilai Ya: 62.  
Jumlah data yang bernilai Tidak: 0.  
 $(-(62/62) * \log_2(62/62) + -(0/62) * \log_2(0/62))$   
 $= 0$
- e. IPK: Sedang  
Jumlah data yang yang bernilai sedang secara keseluruhan berjumlah: 21.  
Jumlah data yang bernilai Ya: 20.  
Jumlah data yang bernilai Tidak: 1.  
 $(-(20/21) * \log_2(20/21) + -(1/21) * \log_2(1/21))$   
 $= 0.276195$
- f. IPK: Rendah  
Jumlah data yang yang bernilai rendah secara keseluruhan berjumlah: 14.  
Jumlah data yang bernilai Ya: 12.  
Jumlah data yang bernilai Tidak: 2.  
 $(-(12/14) * \log_2(12/14) + -(2/14) * \log_2(2/14))$   
 $= 0.591673$
- g. Status: Lulus  
Jumlah data yang yang bernilai lulus secara keseluruhan berjumlah: 48.  
Jumlah data yang bernilai Ya: 47.  
Jumlah data yang bernilai Tidak: 1.  
 $(-(47/48) * \log_2(47/48) + -(1/48) * \log_2(1/48))$   
 $= 0.146094$
- h. Status: Belum  
Jumlah data yang yang bernilai belum secara keseluruhan berjumlah: 49.  
Jumlah data yang bernilai Ya: 37.  
Jumlah data yang bernilai Tidak: 12.  
 $(-(37/49) * \log_2(37/49) + -(12/49) * \log_2(12/49))$   
 $= 0.803091$

Setelah proses perhitungan nilai *entropy* pada masing-masing indikator selesai, maka dilanjutkan untuk proses perhitungan nilai *gain*. Dengan menggunakan rumus *gain* maka akan dapat disimpulkan dari masing-masing indikator mana yang memiliki nilai *gain* paling besar. Adapun perhitungan dapat dijelaskan sebagai berikut:

Rapot:

$(\text{Total nilai } entropy) - ((\text{jumlah data indikator raport bagus} \div \text{jumlah data latih}) * \text{jumlah nilai } entropy \text{ bagus}) - ((\text{jumlah data indikator raport cukup} \div \text{jumlah data latih}) * \text{jumlah nilai } entropy \text{ cukup}) - ((\text{jumlah data indikator raport kurang} \div \text{jumlah data latih}) * \text{jumlah nilai } entropy \text{ kurang}).$

Dapat diterapkan dengan cara perhitungan diatas menjadi:

$$(0.568363) - ((65 \div 97) * 0.538141) - ((28 \div 97) * 0.591673) - ((4 \div 97) * 1) = -0.00428$$

IPK:

$$(0.568363) - ((62 \div 97) * 0) - ((21 \div 97) * 0.276195) - ((14 \div 97) * 0.591673) = 0.423172$$

Status:

$$(0.568363) - ((48 \div 97) * 0.146094) - ((49 \div 97) * 0.803091) = 0.090383$$

Dimana pada perhitungan *gain* diatas, dapat disimpulkan bahwa nilai *gain* tertinggi adalah IPK, dan IPK menjadi *Root Node* yang berjumlah 0.423172.

**Tabel 6. Sampel yang sudah difilterisasi. Dimana IPK merupakan *Root Node***

<i>IPK</i>	<i>Rapot</i>	<i>Status</i>	<i>Prestasi</i>
Sedang	Bagus	Lulus	Ya
Sedang	Bagus	Belum	Ya
Sedang	Cukup	Lulus	Ya
Sedang	Kurang	Belum	Tidak
Sedang	Kurang	Belum	Ya
Rendah	Kurang	Belum	Tidak
Rendah	Bagus	Lulus	Ya
Rendah	Cukup	Belum	Tidak
Rendah	Cukup	Lulus	Ya

Pada tabel 6 merupakan data sampel yang dipilih secara acak dan difilterisasi. Yang dimana IPK menjadi *Root Node*.

### 3.2 Langkah ke Dua

Menentukan nilai *entropy* dengan jumlah data 5 yang diambil dari tabel 5, pada IPK yang bernilai sedang:

$$(-2/5) * \log_2(2/5) + (-3/5) * \log_2(3/5) = 0.970951$$

Dari perhitungan diatas maka nilai *entropy* dapat digunakan dalam pengolahan data nilai *entropy* dan *gain*.

**Tabel 7. Node internal di fitur atau atribut IPK yang bernilai cukup**

	<i>Jumlah</i>	<i>Ya</i>	<i>Tidak</i>	<i>Entropy</i>	<i>Gain</i>
Total	5	2	3	0.970951	
Nilai Rapot					0.570951
Bagus	2	0	2	0	
Cukup	1	1	0	0	
Kurang	2	1	1	1	
Status					0.419973
Lulus	2	2	0	0	
Belum	3	2	1	0.918296	

- a. Rapot: Bagus  
Jumlah data yang yang bernilai bagus secara keseluruhan berjumlah:2.  
Jumlah data yang bernilai Ya: 0.  
Jumlah data yang bernilai Tidak: 2.  
$$(-(0/2) * \log_2(0/2) + -(2/2) * \log_2(2/2))$$
$$= 0$$
- b. Rapot: Cukup  
Jumlah data yang yang bernilai cukup secara keseluruhan berjumlah:1.  
Jumlah data yang bernilai Ya:1.  
Jumlah data yang bernilai Tidak: 0.  
$$(-(1/1) * \log_2(1/1) + -(0/1) * \log_2(0/1))$$
$$= 0$$
- c. Rapot: Kurang  
Jumlah data yang yang bernilai kurang secara keseluruhan berjumlah: 2.  
Jumlah data yang bernilai Ya: 1.  
Jumlah data yang bernilai Tidak: 1.  
$$(-(1/2) * \log_2(1/2) + -(1/2) * \log_2(1/2))$$
$$= 1$$
- d. Status: Lulus  
Jumlah data yang yang bernilai lulus secara keseluruhan berjumlah:2.  
Jumlah data yang bernilai Ya: 2.  
Jumlah data yang bernilai Tidak: 0.  
$$(-(2/2) * \log_2(2/2) + -(0/2) * \log_2(0/2))$$
$$= 0$$
- e. Status: Belum  
Jumlah data yang yang bernilai belum secara keseluruhan berjumlah: 3.  
Jumlah data yang bernilai Ya: 2.  
Jumlah data yang bernilai Tidak: 1.  
$$(-(2/3) * \log_2(2/3) + -(1/3) * \log_2(1/3))$$
$$= 0.918296$$

Setelah proses perhitungan nilai *entropy* pada masing-masing indikator selesai, maka dilanjutkan untuk proses perhitungan nilai *gain*. Dengan menggunakan rumus *gain* maka akan dapat disimpulkan dari masing-masing indikator mana yang memiliki nilai *gain* paling besar. Adapun perhitungan dapat dijelaskan sebagai berikut:

- a. Rapot:  
$$(0.970951) - ((2 \div 5) * 0) - ((1 \div 5) * 0) - ((2 \div 5) * 1) = 0.570951$$
- b. Status:  
$$(0.970951) - ((2 \div 5) * 0) - ((3 \div 5) * 0.918296) = 0.419973$$

Dimana pada perhitungan gain diatas, dapat disimpulkan bahwa nilai *gain* tertinggi adalah Rapot, dan Rapot menjadi akar pencabangan dari IPK yang bernilai sedang yang berjumlah 0.570951.

### 3.3 Langkah ke Tiga

Untuk langkah ke tiga sendiri, *Node* 3 di fitur IPK dengan cabang Tinggi bernilai nol karena sudah dipastikan semua data node cabang tinggi masuk di kelas yang sama dan menjadi *node* terminal atau daun.

### 3.4 Langkah ke Empat

Menentukan nilai *entropy* dengan jumlah data 4 yang diambil dari tabel 5, pada IPK yang bernilai rendah.

$$(-(2/4) * \log_2(2/4) + -(2/4) * \log_2(2/4))$$
$$= 1$$

**Tabel 8. Node internal di cabang IPK yang bernilai rendah**

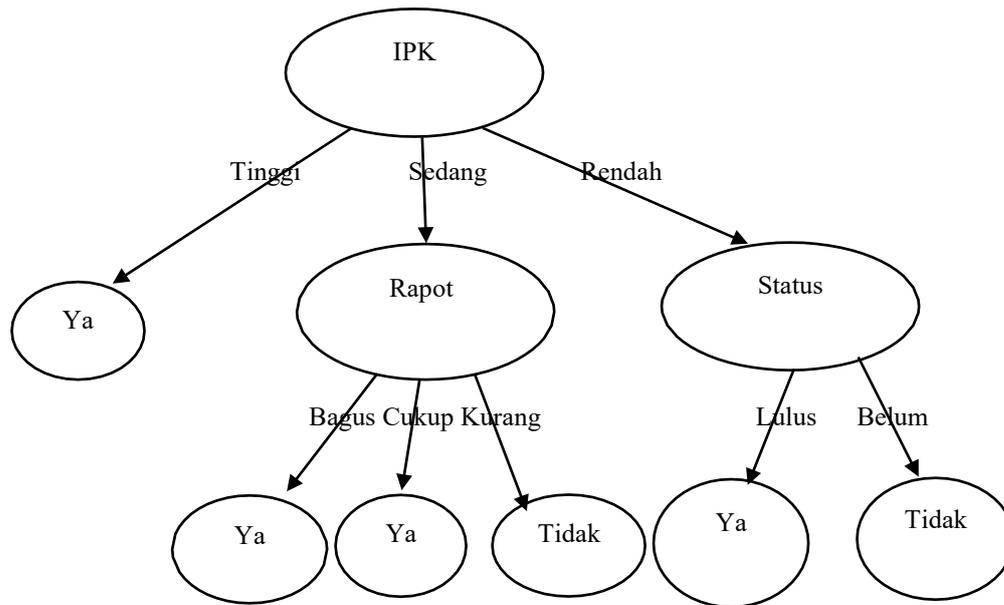
	<i>Jumlah</i>	<i>Ya</i>	<i>Tidak</i>	<i>Entropy</i>	<i>Gain</i>
Total	4	2	2	1	
Nilai Rapot					0.5
Bagus	1	1	0	0	
Cukup	2	1	1	1	
Kurang	1	0	1	0	
Status					1
Lulus	2	2	0	0	
Belum	2	0	2	0	

- a. Rapot: Bagus  
Jumlah data yang yang bernilai bagus secara keseluruhan berjumlah:1.  
Jumlah data yang bernilai Ya: 1.  
Jumlah data yang bernilai Tidak: 0.  
 $(-(1/1) * \log_2(1/1) + -(0/1) * \log_2(0/1))$   
 $= 0$
- b. Rapot: Cukup  
Jumlah data yang yang bernilai cukup secara keseluruhan berjumlah:2.  
Jumlah data yang bernilai Ya:1.  
Jumlah data yang bernilai Tidak: 1.  
 $(-(1/2) * \log_2(1/2) + -(1/2) * \log_2(1/2))$   
 $= 1$
- c. Rapot: Kurang  
Jumlah data yang yang bernilai kurang secara keseluruhan berjumlah: 1.  
Jumlah data yang bernilai Ya: 0.  
Jumlah data yang bernilai Tidak: 1.  
 $(-(0/1) * \log_2(0/1) + -(1/1) * \log_2(1/1))$   
 $= 0$
- d. Status: Lulus  
Jumlah data yang yang bernilai lulus secara keseluruhan berjumlah:2.  
Jumlah data yang bernilai Ya: 2.  
Jumlah data yang bernilai Tidak: 0.  
 $(-(2/2) * \log_2(2/2) + -(0/2) * \log_2(0/2))$   
 $= 0$
- e. Status: Belum  
Jumlah data yang yang bernilai belum secara keseluruhan berjumlah: 2.  
Jumlah data yang bernilai Ya: 2.  
Jumlah data yang bernilai Tidak: 0.  
 $(-(2/2) * \log_2(2/2) + -(0/2) * \log_2(0/2))$   
 $= 0$

Setelah proses perhitungan nilai *entropy* pada masing-masing indikator selesai, maka dilanjutkan untuk proses perhitungan nilai *gain*. Dengan menggunakan rumus *gain* maka akan dapat disimpulkan dari masing-masing indikator mana yang memiliki nilai *gain* paling besar. Adapun perhitungan dapat dijelaskan sebagai berikut:

- a. Rapot:  
 $(1) - ((1 \div 4) * 0) - ((2 \div 4) * 1) - ((1 \div 4) * 0) = 0.5$
- b. Status:  
 $(1) - ((2 \div 4) * 0) - ((2 \div 4) * 0) = 1$

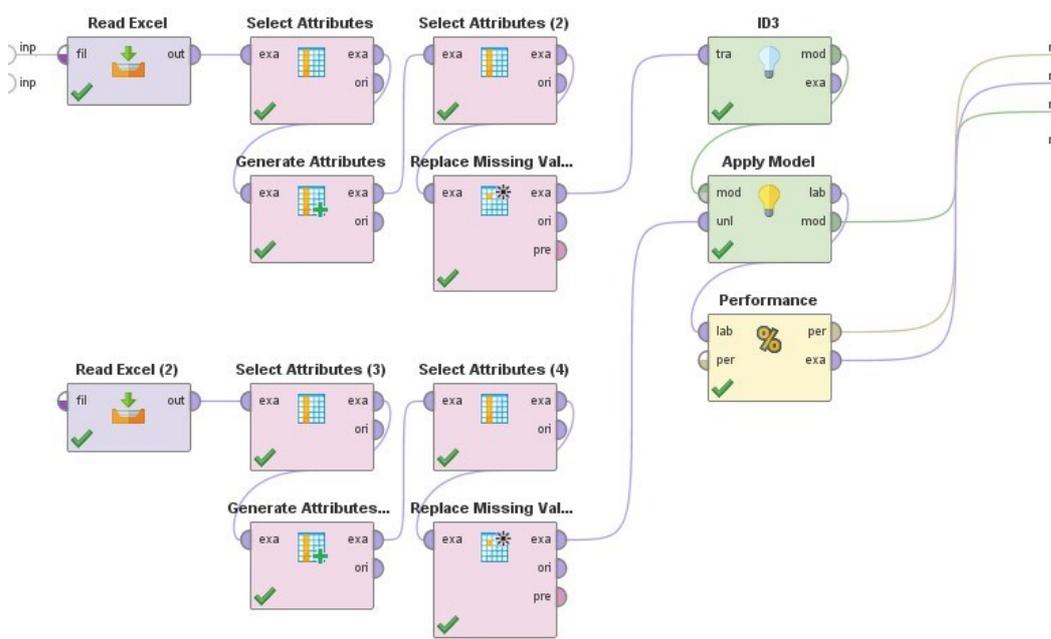
Dimana pada perhitungan gain diatas, dapat disimpulkan bahwa nilai *gain* tertinggi selanjutnya adalah status, dan status menjadi akar pencabangan dari IPK yang bernilai rendah yang berjumlah 1. Pada perhitungan diatas dapat disimpulkan bahwa *Root Node* pada penelitian diatas adalah Nilai IPK, dan dengan sejumlah perhitungan diatas maka akan dibentuk pohon keputusan dengan *Root Node* dari Nilai IPK yang ditunjukkan pada gambar 2.



**Gambar 2. Hasil Dari Perhitungan Manual**

Dimana *Root Nodenya* adalah Nilai IPK jika Nilai IPK Tinggi maka Ya. Jika Sedang Maka akan dilihat ke bagian Nilai Rapot, dimana jika bagus maka Ya, jika cukup maka Ya jika kurang maka Tidak. Jika bernilai Rendah maka akan dilihat ke bagian status dimana jika dia Lulus maka Ya, jika belum maka Tidak.

Adapun dalam penganalisisan ini melibatkan pengujian dengan menggunakan *software RapidMiner*, dimana dengan menggunakan *software RapidMiner* ini maka akan diuji apakah hasil dari pengujian data secara manual dan dengan menggunakan *software* sama atau tidaknya, maka dapat dilihat pada alur proses tahapan perancangan sebagai berikut:



**Gambar 3. Alur Tahapan Proses Perancangan Perhitungan Dengan RapidMiner**

Dimana *Read Excel 1* adalah data latih, sedangkan *Read Excel 2* adalah data uji. Sedangkan fungsi *performance* menghitung jumlah persentase dari keakuratan perhitungan data tersebut dengan menggunakan metode ID3. Hasil dari model proses tersebut ditunjukkan pada gambar 4.



Gambar 4. Pohon Keputusan Keberhasilan Studi

Gambar 4 merupakan hasil dari proses atau tahapan-tahapan dalam model proses yang dihitung dengan menggunakan *software RapidMiner* tersebut, yang dimana *Root Node* adalah IPK. Dimana Nilai IPK yang bernilai Ya berjumlah *Ratio of total*: 63,92%. Sedangkan Nilai IPK yang terbagi ke nilai Rapor dengan bernilai Bagus berjumlah *Ratio of total*: 15,46%, sedangkan yang bernilai Cukup berjumlah *Ratio of total*: 5,15%, yang bernilai Kurang berjumlah *Ratio of total*: 1,03%. Sedangkan nilai IPK yang terbagi ke Status Lulus berjumlah *Ratio of total*: 2,06, sedangkan Status Belum Lulus berjumlah *Ratio of total*: 12,37%.

#### 4. KESIMPULAN

Penerapan algoritma *Iterative Dichotomizer 3 (ID3)* dimana proses awal dimulai dari proses dimana data-data mahasiswa yang telah didapat dipilih dan dirubah ke dalam format kolom penganalisisan data sesuai dengan penelitian yang dilakukan. Dari perhitungan diatas maka dapat disimpulkan yang menjadi faktor utama dalam penelitian ini adalah Nilai IPK. Dimana Nilai IPK memiliki nilai Gain yang lebih tinggi dalam perhitungan awal yaitu berjumlah 0,423172. Dimana nilai IPK merupakan akar dari pohon keputusan yang mencabangkan ke Nilai Rapor dan Status. Dengan masing-masing persentase dari perhitungan yang terdiri dari:

- Nilai IPK yang bernilai Ya berjumlah *Ratio Of Total*: 63,92%.
- Nilai IPK yang terbagi ke nilai Rapor dengan bernilai Bagus berjumlah *Ratio of total*: 15,46%.
- Bernilai Cukup berjumlah *Ratio of total*: 5,15%.
- Bernilai Kurang berjumlah *Ratio of total*: 1,03%.
- IPK yang terbagi ke Status dan dinyatakan Lulus berjumlah *Ratio of total*: 2,06.
- IPK yang terbagi ke Status dan Belum Lulus berjumlah *Ratio of total*: 12,37%.

Dimana prediksi yang dilakukan dalam penelitian ini mampu memberikan informasi terkait dengan tingkat persentase dalam masing-masing indikator yang berguna untuk mendorong mahasiswa agar lebih dapat membantu dalam proses prestasi.

#### DAFTAR PUSTAKA

- Nugroho, A. K., and Iskandar, D. (2015). Algoritma Iterative Dichotomizer 3 (ID3) Pengambilan Keputusan. *Dinamika Rekayasa*, 11(2), 44-48.
- Giovani, Ronny Ardi., et al. (2011) Sistem Pendukung Keputusan Prediksi Kecepatan Studi Mahasiswa Menggunakan Metode ID3. *Jurnal Buana Informatika*, 2.2.
- Tyasti, Avia Enggar., et al. (2015). Algoritma Iterative Dichotomiser 3 (Id3) Untuk Mengidentifikasi Data Rekam Medis (Studi Kasus Penyakit Diabetes Mellitus Di Balai Kesehatan Kementerian Perindustrian, Jakarta). *Jurnal Gaussian*, 4.2: 237-246.
- Bagus, A. W., Et Al. (2012). Perancangan Dan Implementasi Sistem Pendukung Keputusan Untuk Jalan Menggunakan Metode Id3 (Studi Kasus Bappeda Kota Salatiga).
- Danny, Himawan. (2011). Aplikasi Data Mining Menggunakan Algoritma Id3 Untuk Mengklasifikasi Kelulusan Mahasiswa Pada Universitas Dian Nuswantoro Semarang. *Skripsi, Fakultas Ilmu Komputer*.
- Obbie, Kristanto. (2014). Penerapan Algoritma Klasifikasi Data Mining Id3 Untuk Menentukan Penjurusan Siswa Sman6 Semarang. *Skripsi, Fakultas Ilmu Komputer*.
- MOCHLIAT, Maria Magdalena. (2017). *Perancangan dan Implementasi Sistem Pengambilan Keputusan Mendaki Gunung Menggunakan Algoritma ID3: Studi Kasus Balai Taman Nasional Gunung Merbabu*. Thesis. Program Studi Teknik Informatika FTI-UKSW.

- [8] Adiana, Sedyono, & Dwi, 2014. Pendekatan Metode Pohon Keputusan Menggunakan Algoritma ID3 Untuk Sistem Informasi Pengukuran Kinerja PNS. *Jurnal Sistem Informasi Bisnis*, 4.2: 75-86.
- [9] Arista, N., Bettiza, M., & Kurniawan, H. (2014). Penerapan Algoritma ID3 dan C4. 5 dalam Menemukan Hubungan Data Awal Masuk Mahasiswa dengan Prestasi Akademik. *Universitas Maritim Raja Ali Aji Tanjungpinang*, 1-10.
- [10] Agustina, D. M. (2016). Analisis Perbandingan Algoritma ID3 Dan C4. 5 Untuk Klasifikasi Penerima Hibah Pemasangan Air Minum Pada PDAM Kabupaten Kendal. *Journal of Applied Intelligent System*, 1(3), 234-244.