

IMPLEMENTASI ALGORITMA C4.5 DALAM KLASIFIKASI KETEPATAN WAKTU KELULUSAN PADA DATA MAHASISWA PENERIMA BEASISWA BIDIKMISI (STUDI KASUS : UNIVERSITAS MARITIM RAJA ALI HAJI)

Taufik Ramadhan¹, Eka Suswaini², Alena Uperiati³

Taufikramadhan1901@gmail.com

Program studi Teknik Informatika Fakultas Teknik Universitas Maritim Raja Ali Haji

Abstract

Student graduation greatly impacts campus accreditation. The imbalance between the number of incoming and outgoing students causes an accumulation of the number of students, where a large number of students enter but the number who leave (graduate) on time is much smaller. Therefore, in this study, the C4.5 algorithm method was applied in order to classify the timeliness of student graduation and to determine the accuracy results of the system implemented. By using the existing data, then the data mining process is carried out, the results of the tests that have been carried out with the C4.5 algorithm are able to classify the punctuality of student graduation on time and not on time and produce classification results with an accuracy value of 95%. This proves that the decision tree algorithm C4.5 method can be implemented in the case of student graduation classification.

Keywords: Classification, Algorithm C4.5, Accuracy

I. Pendahuluan

Kelulusan mahasiswa sangat berpengaruh terhadap akreditasi kampus maka sangat diharapkan mahasiswa dapat lulus tepat pada waktunya (Martanto dkk., 2019). Berdasarkan data dari Universitas Maritim Raja Ali Haji, ditemukan ketidakseimbangan antara jumlah mahasiswa yang masuk dan keluar. Mahasiswa yang mendaftar sangat lah banyak namun yang bisa lulus tepat waktu tidak lah sebanding. Melihat kondisi seperti ini, ada masalah yang dihadapi oleh pihak Universitas Maritim Raja Ali Haji yakni tidak seimbang nya mahasiswa yang masuk dan mahasiswa yang lulus terlihat dari data mahasiswa bidikmisi tahun 2015 yang ada di Universitas Maritim Raja Ali Haji. Di tahun 2015 penerima beasiswa bidikmisi tidak seimbang antara yang mahasiswa penerima beasiswa dan mahasiswa yang lulus tepat Masalah ini akan berdampak panjang dalam masa studi mahasiswa yang menerima bidikmisi namun tidak lulus tepat pada waktunya, dan tentu menjadi perhatian serius bagi manajemen dan kepengurusan penerimaan beasiswa bidikmisi di Universitas Maritim Raja Ali Haji.

Berdasarkan masalah yang ada maka melakukan teknik penggalian data dengan prediksi bisa digunakan untuk mengetahui klasifikasi mahasiswa yang lulus tepat atau tidak tepat waktu. Teknik penggalian data yang akan dilakukan berdasarkan data pendidikan di perguruan tinggi (Han dkk., 2011). Sementara itu algoritma yang digunakan ialah Algoritma C4.5. Cara kerja algoritma ini adalah dengan mencari nilai tiap *entropy* yakni nilai paling besar pada atribut, kemudian mencari nilai *gain* yaitu nilai paling populer untuk pemilihan atribut, kemudian mencari nilai *split*

information yaitu nilai dengan informasi potensial dalam atribut, dan mencari nilai *gain ratio* yakni nilai tertinggi dari atribut untuk menjadi *root* (akar) yang kemudian akan di bentuk pohon keputusan data yang menjadi hasil dari proses klasifikasi pada algoritma C4.5.

Data yang digunakan dalam penelitian ini yakni data para mahasiswa yang menerima beasiswa Bidikmisi Universitas Maritim Raja Ali haji tahun 2015 sebanyak 222 data, dengan variable seperti data IPK (Indeks Prestasi Kumulatif) semester 7, data jumlah SKS (Satuan Kredit Semester) semester 7, data nilai mata kuliah Kerja Praktik/Magang/Kuliah Kerja Nyata (KKN/Magang/KP), dan data nilai MK (Mata Kuliah) Skripsi.

Setelah di lakukan analisis terhadap tingkat ketepatan waktu kelulusan mahasiswa maka diharapkan metode algoritma C4.5 ini dapat menjadi solusi untuk memprediksi kelulusan mahasiswa. Dimana dengan metode algoritma C4.5, Universitas Maritim Raja Ali Haji khususnya manajemen pengurus beasiswa bidikmisi dapat mengetahui informasi ketepatan waktu kelulusan mahasiswa penerima bidikmisi.

Pada penelitian ini juga akan mengimplementasikan sebuah sistem untuk klasifikasi data mahasiswa penerima beasiswa bidikmisi menggunakan metode C4.5 berbasis web, agar dapat mengklasifikasikan mahasiswa yang lulus tepat waktu atau tidak tepat waktu serta lebih mudah untuk mendapatkan informasi digital.

II. Metode Penelitian

2.1 Algoritma C4.5

Algoritma C 4.5 merupakan kelompok algoritma *Decision tree* yang mempunyai input berupa *training samples* dan *samples* (Nofriansyah., 2015)

a. Mencari Nilai *Entropy*

Mencari Nilai *entropy* menggunakan persamaan berikut :

$$Entropy(S) = \sum n_i = (-p_i) * \log_2(p_i)$$

Dimana :

- S = Himpunan Kasus
- n = Jumlah partisi atribut S
- pi = Proporsi dari Si ke S

b. Mencari Nilai *Gain*

Nilai *Gain* tertinggi digunakan untuk memilih nilai paling populer dari atribut data.

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

Dimana :

- S = Himpunan Kasus
- A = Atribut
- n = Jumlah Partisi atribut A
- |S| = Jumlah kasus pada S
- |S_i| = Jumlah kasus pada partisi i

c. *Split Information*

Split Information merupakan nilai yang dicari melalui pendekatan yang digunakan pada data dengan kata lain mencari nilai yang paling potensial dari data yang ada.

$$SplitInfo = - \sum_{n=1}^n \left(\frac{S_i}{S} \log_2 \frac{S_i}{S} \right)$$

Dimana :

- S = Ruang (*Sample*) yang digunakan untuk *training*
- A = Atribut
- Si = Jumlah *Sample* untuk atribut i

d. *Gain Ratio*

Gain Ratio adalah nilai yang dicari pada data dengan melihat nilai paling populer yang muncul kemudian nilai tertinggi akan menjadi *root* (akar) pada pembentukan pohon keputusan hasil klasifikasi.

$$Gain Ratio (A) = \frac{gain(A)}{Splitinfo(A)}$$

Dimana :

- A = Atribut
- *gain (A)* = *information gain* pada atribut a
- *Splitinfo (A)* = *split information* pada atribut a

- e. Nilai *Gain Ratio* tertinggi akan menjadi akar (*root*).
- f. Ulangi proses b sampai semua cabang memiliki kelas yang sama
- g. Maka akan terbentuklah pohon keputusan dari hasil klasifikasi yang digunakan pada data

2.2 Data Mining

Data Mining adalah metode cerdas yang di implementasikan dalam ekstraksi pada pola data untuk di analisis yang kemudian data bisa digunakan untuk meraih informasi atau pengetahuan (Han dkk., 2011).

Terdapat empat operasi umum data mining :

1. *Predictive and Classification Modelling*, ialah metode yang digunakan pada prediksi. Seperti prediksi cuaca atau prediksi curah hujan dengan membangun model pada satu nilai. Contoh metode yang banyak digunakan adalah metode *Linear regression*, metode *neural network*, metode *support vector Machine* dan lainnya.
2. *Association* (Asosiasi) ialah metode yang digunakan untuk analisis kasus yang terjadi pada pikiran dan perilaku manusia sehari-hari, dengan cara mempelajari hubungan antar data yang terjadi pada setiap kejadian yang tercipta. Contoh metode nya adalah metode *FP growth*, dan *Apriori*.
3. *Clustering* (Kluster) ialah data – data yang dikelompokkan dalam satu kelompok tertentu atau terpisah. Dengan menganalisis munculnya data pada setiap kejadian yang ada. Contoh implementasinya ialah metode *K-means*, *K-medoids*, *SOM*, *Fuzzy C-means* dan lainnya.
4. *Classification* (klasifikasi) ialah teknik untuk menemukan variable dependen yang ada pada setiap data yang nanti akan diberi nilai kemudian di klasifikasikan menurut kelas dan keadaan yang sama satu dengan yang lainnya. Contoh penerapan algoritmanya adalah pada algoritma *ID3* dan *KNN* (*K Nearest Neighbors*).

2.3 Klasifikasi

Klasifikasi adalah konsep dalam ekstraksi data menjadi model yang dapat memberikan gambaran suatu pengetahuan dan informasi melalui nilai dari variable dependen setiap nilai yang ada (Han dkk., 2011).

2.4 Beasiswa Bidikmisi

Beasiswa Bidikmisi merupakan beasiswa yang disalurkan untuk calon mahasiswa atau mahasiswi yang sudah menjadi mahasiswa atau mahasiswi yang kurang mampu namun memiliki prestasi akademik yang baik. Beasiswa ini diterima selama seorang mahasiswa dan menyelesaikan program studinya selama waktu yang telah ditetapkan oleh pemerintah dan kampus, juga selama mahasiswa tersebut mengikuti aturan yang ada dan berlaku (Universitas Papua, 2018).

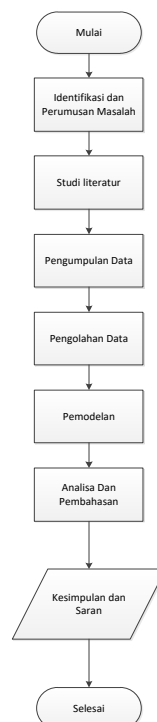
2.5 Tingkat Akurasi

Mencari akurasi dari kasus klasifikasi adalah membagi jumlah tepat dari setiap hasil dari data yang diklasifikasikan dengan jumlah keseluruhan data yang telah diklasifikasikan selanjutnya akan dikalikan dengan 100% untuk mendapatkan hasil akurasi yang maksimal (Widianingsih., 2019). Dibawah ini ialah rumus mencari akurasi pada klasifikasi:

$$Akurasi = \frac{Jumlah\ Tepat}{Jumlah\ Total\ Data} \times 100\%$$

2.6 Flowchart Prosedur Penelitian

Adapun Flowchart Prosedur Penelitian yang dilakukan pada Penerapan Algoritma C4.5 seperti terlihat pada Gambar 1.

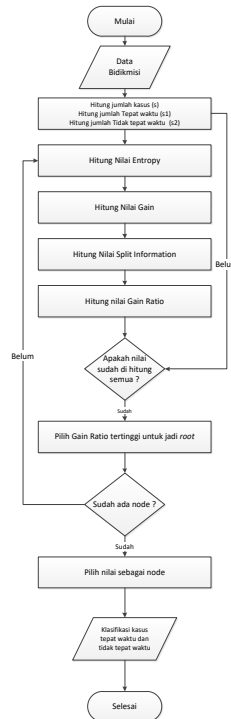


Gambar 1. Flowchart Prosedur Penelitian

Pada gambar 1 merupakan metode penelitian dengan awal mengidentifikasi dan merumuskan masalah mengenai data bidikmisi tahun 2015, kemudian mencari studi literatur sebagai bahan

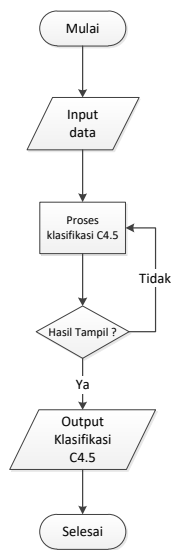
kajian untuk penelitian yang diangkat, kemudian melakukan koleksi kebutuhan data-data yang mendukung pada proses penelitian, lalu mengolahnya, kemudian membuat model perangkat lunak yang digunakan untuk implementasi data, lalu melakukan analisa hasil penelitian dan menyampaikan pembahasan di dokumen, kemudian mengambil kesimpulan dari hasil pembahasan dan memberikan saran atas penelitian yang dilakukan.

Adapun Flowchart metode yang dilakukan pada penerapan Algoritma C4.5 seperti terlihat pada Gambar 2.



Pada gambar 2 merupakan *flowchart* algoritma C4.5 dengan awal memasukkan data Bidikmisi 2015, kemudian menghitung jumlah atribut yang ada, kemudian menghitung jumlah kasus yang tepat waktu dan tidak tepat waktu, kemudian menghitung nilai *entropy* masing masing atribut, kemudian menghitung nilai *gain*, dilanjutkan mencari nilai *split information*, kemudian menemukan nilai *gain ratio* dari semua nilai atribut data, lalu masuk ke kondisi jika belum terhitung semua atribut kasus, maka kembali menghitung ulang *entropy* dari awal, jika sudah maka pilih nilai *gain ratio* tertinggi untuk menjadi *root*, kemudian masuk ke kondisi jika belum ada *node* atau cabang maka hitung ulang dari mencari nilai *entropy* atribut selanjutnya yang sudah dinormalisasi, jika sudah maka pilih cabang dengan nilai *gain ratio* tertinggi menjadi *node* atau cabang, kemudian klasifikasikan kasus tepat atau tidak tepat waktu, dan kemudian selesai .

Adapun Flowchart metode yang dilakukan pada penerapan Algoritma C4.5 seperti terlihat pada Gambar 3.



Gambar 3. Flowchart Sistem Klasifikasi Ketepatan Waktu Kelulusan

Pada gambar 3 merupakan bagan alur dari algoritma sistem dengan awal memasukkan/ input data mahasiswa, kemudian melakukan proses algoritma C4.5 dengan meng-input jawaban tiap pilihan dari pertanyaan-pertanyaan yang akan di proses untuk mencari nilai *entropy*, *gain*, *split information* dan *gain ratio*. Kemudian hasil akan otomatis tampil pada layar dan klasifikasi pun telah dilakukan berupa dengan ditandai *output* atau keluaran Tepat waktu atau Tidak Tepat Waktu. Kemudian proses sistem klasifikasi ketepatan waktu kelulusan pun selesai.

III. Hasil dan Pembahasan

3.1 Hasil Penelitian

Data pada pada sistem klasifikasi kali ini ialah data kelulusan mahasiswa ini sebanyak 222 data yaitu nilai IPK semester 7, jumlah SKS semester 7, nilai KP/Magang/KKN, dan nilai Mk Skripsi. Berikut data yang akan digunakan terlihat pada tabel 1.

3.2 Pengujian dan Analisa SKKWK

Proses pengujian akan dilakukan pada data aktual yang berjumlah 222.

Table 1. Data Aktual

Data Bidikmisi				
n	Nilai IPK sem 7	Jumlah SKS sem 7	Nilai KP/Magang/KKN	Nilai Skripsi
1	3.27	144	4	4
2	3.19	135	4	0
3	3.27	144	4	0
4	3.47	138	4	0
5	2.93	129	4	0
6	3.00	120	4	0
7	3.06	129	4	0

8	3.45	138	4	0
9	3.22	144	4	0
10	3.43	144	4	3,5
11	3.18	135	4	0
12	3.42	138	4	0
13	3.23	144	4	3,5
14	3.09	144	4	3,5
15	3.11	138	4	0
16	3.05	132	4	0
17	3.42	144	4	0
18	3.21	129	4	0
19	3.19	129	4	0
20	3.4	144	4	4
...
222	3.36	148	4	0

3.3 Data pilihan

Tabel berikut ini merupakan tabel yang digunakan untuk acuan dalam memberikan pilihan dari tiap- tiap soal yang akan di jadikan data bobot jawaban. Tabel ini menyediakan empat pertanyaan dengan delapan jawaban di masing-masing pertanyaan.

Tabel 2. Data Pilihan

no	a	b	c	D	e	f	g	h
Pilihan a	< Rentang IPK 2,85	Rentang IPK 2,86 - 3,04	Rentang IPK 3,05-3,09	Rentang IPK 3,10- 3,18	Rentang IPK 3,15-3,19	Rentang IPK 3,20- 3,24	Rentang IPK 3,25- 3,49	Rentan g IPK 3,50 - 4,00
Pilihan b	< Rentang SKS 100	Rentang SKS 100 - 105	Rentang SKS 106-110	Rentang SKS 111- 115	Rentang SKS 116 - 120	Rentang SKS 121- 129	Rentang SKS 130 - 139	Rentan g SKS 141- 144
Pilihan c	Bobot Skripsi E = 0	Bobot Skripsi D = 1.00	Bobot Skripsi C = 2.00	Bobot Skripsi B- = 2.50	Bobot Skripsi B = 3.00	Bobot Skripsi B+ = 3.25	Bobot Skripsi A- = 3.50	Bobot Skripsi A = 4.00
Pilihan d	bobot KKN/magang /Kp E = 0	bobot KKN/ma gang/Kp D = 1.00	bobot KKN/magang /Kp C = 2.00	bobot KKN/ma gang/Kp B- = 2.50	bobot KKN/magang/ Kp B = 3.00	bobot KKN/mag ang/Kp B+ = 3.25	bobot KKN/mag ang/Kp A- = 3.50	bobot KKN/ magan g/Kp A = 4.00

3.4 Tabel Transformasi data aktual ke jawaban tiap pilihan

Table berikut ini merupakan transformasi rincian jawaban dari masing – masing soal yang telah di pilih. Tabel ini berisikan data mahasiswa dan bobot masing-masing jawaban A, B, C, D, E, F, G dan H tiap user.

Tabel 3. Data transformasi data aktual ke jawaban

n	Jawaban A	Jawaban B	Jawaban C	Jawaban D	Jawaban E	Jawaban F	Jawaban G	Jawaban H
1	0	0	0	1	0	0	0	3
2	0	0	0	0	0	1	0	3
3	0	0	0	1	0	0	0	3
4	0	0	0	0	0	1	0	3
5	1	1	0	0	1	0	0	1
6	1	0	2	0	0	0	0	1
7	1	0	1	0	1	0	0	1
8	0	0	0	1	0	0	0	3
9	0	0	0	0	0	1	0	3

10	0	0	0	0	0	0	2	2
11	0	0	0	0	0	0	2	2
12	0	0	0	0	0	0	2	2
13	0	0	0	0	0	0	2	2
14	0	0	0	1	0	0	0	3
15	0	0	0	0	0	1	0	3
16	0	0	0	0	0	0	2	2
17	0	0	0	0	0	0	2	2
18	0	0	0	1	0	0	0	3
19	0	0	0	0	0	1	0	3
20	0	0	0	0	0	0	1	3
...
222	0	0	0	0	0	1	0	3

3.5 Pembahasan

Pada penelitian implementasi algoritma C4.5 dalam menentukan klasifikasi ketepatan waktu kelulusan kali ini yang dilakukan untuk klasifikasi kelulusan mahasiswa bidikmisi angkatan tahun 2015 Universitas Maritim Raja Ali Haji. Penelitian menggunakan 222 data ini akan di hasilkan untuk 2 hasil klasifikasi yaitu “tepat waktu dan tidak tepat waktu”.

3.6 Pengujian Algoritma C4.5

Pengujian ini bertujuan untuk mengetahui hasil klasifikasi dari parameter pertama yaitu Bobot Nilai IPK Semester 7.

Tabel 4. Hasil Uji Parameter Pertama

Data Bidikmisi			
n	IPK sem 7	Kelas Asli	Kelas hasil
1	3.27	Tepat waktu	Tepat waktu
2	3.19	Tepat waktu	Tepat waktu
3	3.27	Tepat waktu	Tepat waktu
4	3.47	Tepat waktu	Tepat waktu
5	2.93	Tidak Tepat waktu	Tidak Tepat waktu
6	3.00	Tidak Tepat waktu	Tidak Tepat waktu
7	3.06	Tidak Tepat waktu	Tidak Tepat waktu
8	3.45	Tepat waktu	Tepat waktu
9	3.22	Tepat waktu	Tepat waktu
10	3.43	Tepat waktu	Tepat waktu
11	3.18	Tepat waktu	Tepat waktu
12	3.42	Tepat waktu	Tepat waktu
13	3.23	Tepat waktu	Tepat waktu
14	3.09	Tepat waktu	Tepat waktu
15	3.11	Tepat waktu	Tepat waktu
16	3.05	Tepat waktu	Tepat waktu
17	3.42	Tepat waktu	Tepat waktu
18	3.21	Tepat waktu	Tepat waktu

19	3.19	Tepat waktu	Tepat waktu
20	3.41	Tepat waktu	Tepat waktu
...
222	3.36	Tepat waktu	Tepat waktu

Pengujian ini bertujuan untuk mengetahui hasil klasifikasi dari parameter pertama yaitu Bobot Nilai SKS semester 7.

Tabel 5. Hasil Uji Parameter Kedua

Data Bidikmisi			
n	SKS sem 7	Kelas Asli	Kelas hasil
1	144	Tepat waktu	Tepat waktu
2	135	Tepat waktu	Tepat waktu
3	144	Tepat waktu	Tepat waktu
4	138	Tepat waktu	Tepat waktu
5	129	Tidak Tepat waktu	Tidak Tepat waktu
6	120	Tidak Tepat waktu	Tidak Tepat waktu
7	129	Tidak Tepat waktu	Tidak Tepat waktu
8	138	Tepat waktu	Tepat waktu
9	144	Tepat waktu	Tepat waktu
10	144	Tepat waktu	Tepat waktu
11	135	Tepat waktu	Tepat waktu
12	138	Tepat waktu	Tepat waktu
13	144	Tepat waktu	Tepat waktu
14	144	Tepat waktu	Tepat waktu
15	138	Tepat waktu	Tepat waktu
16	132	Tepat waktu	Tepat waktu
17	144	Tepat waktu	Tepat waktu
18	129	Tepat waktu	Tepat waktu
19	129	Tepat waktu	Tepat waktu
20	144	Tepat waktu	Tepat waktu
...
222	148	Tepat waktu	Tepat waktu

Pengujian ini bertujuan untuk mengetahui hasil klasifikasi dari parameter pertama yaitu Bobot Nilai KP/Magang/KKN.

Tabel 6. Hasil Uji Parameter Ketiga

Data Bidikmisi			
n	Nilai KP/Magang/KKN	Kelas Asli	Kelas hasil
1	4	Tepat waktu	Tepat waktu

2	4	Tepat waktu	Tepat waktu
3	4	Tepat waktu	Tepat waktu
4	4	Tepat waktu	Tepat waktu
5	4	Tidak Tepat waktu	Tidak Tepat waktu
6	4	Tidak Tepat waktu	Tidak Tepat waktu
7	4	Tidak Tepat waktu	Tidak Tepat waktu
8	4	Tepat waktu	Tepat waktu
9	4	Tepat waktu	Tepat waktu
10	4	Tepat waktu	Tepat waktu
11	4	Tepat waktu	Tepat waktu
12	4	Tepat waktu	Tepat waktu
13	4	Tepat waktu	Tepat waktu
14	4	Tepat waktu	Tepat waktu
15	4	Tepat waktu	Tepat waktu
16	4	Tepat waktu	Tepat waktu
17	4	Tepat waktu	Tepat waktu
18	4	Tepat waktu	Tepat waktu
19	4	Tepat waktu	Tepat waktu
20	4	Tepat waktu	Tepat waktu
...
222	4	Tepat waktu	Tepat waktu

Pengujian ini bertujuan untuk mengetahui hasil klasifikasi dari parameter pertama yaitu Bobot Nilai MK Skripsi.

Tabel 7. Hasil Uji Parameter Keempat

Data Bidikmisi			
N	Nilai MK Skripsi	Kelas Asli	Kelas hasil
1	4	Tepat waktu	Tepat waktu
2	0	Tepat waktu	Tepat waktu
3	0	Tepat waktu	Tepat waktu
4	0	Tepat waktu	Tepat waktu
5	0	Tidak Tepat waktu	Tidak Tepat waktu
6	0	Tidak Tepat waktu	Tidak Tepat waktu
7	0	Tidak Tepat waktu	Tidak Tepat waktu
8	0	Tepat waktu	Tepat waktu
9	0	Tepat waktu	Tepat waktu
10	3,5	Tepat waktu	Tepat waktu
11	0	Tepat waktu	Tepat waktu
12	0	Tepat waktu	Tepat waktu

13	3,5	Tepat waktu	Tepat waktu
14	3,5	Tepat waktu	Tepat waktu
15	0	Tepat waktu	Tepat waktu
16	0	Tepat waktu	Tepat waktu
17	0	Tepat waktu	Tepat waktu
18	0	Tepat waktu	Tepat waktu
19	0	Tepat waktu	Tepat waktu
20	4	Tepat waktu	Tepat waktu
...
222	0	Tepat waktu	Tepat waktu

3.7 Tingkat Akurasi Algoritma C4.5

Pada langkah ini akan dilakukan pencarian nilai akurasi yang terlihat dari data yang ada. Dalam pencarian nilai akurasi kali ini membagi sebanyak nilai dari jumlah tepat dengan total data keseluruhan kemudian dikalikan dengan 100% untuk mendapatkan hasil akhir dari akurasi sistem.

Tabel 8. Akurasi Data uji

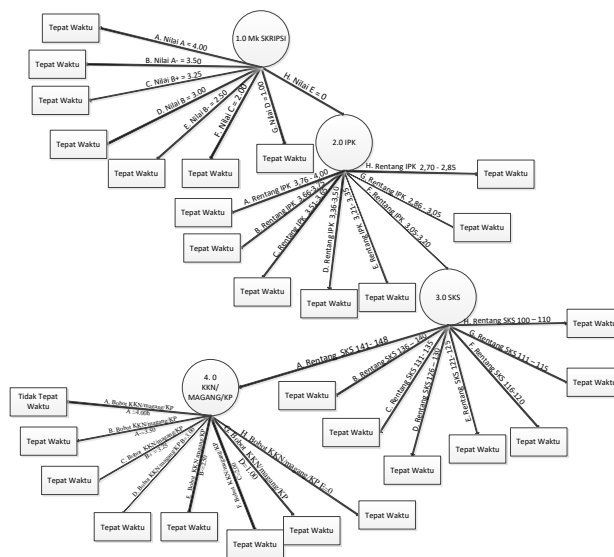
	Keputusan Tepat (Tepat Waktu)	Keputusan Tidak Tepat (Tidak Tepat Waktu)
Tepat Waktu	11	1
Tidak Tepat Waktu	0	8

$$\text{Akurasi | Data Testing} = \frac{(11+8)}{20} \times 100\% = 95\%$$

Maka berdasarkan pada hasil data testing proses mining maka didapatkan tingkat akurasi klasifikasi algoritma C4.5 proses mining sebesar 95%.

3.8 Pohon Keputusan

Berdasarkan tabel – tabel hasil uji diatas maka didapatlah Pohon keputusan (*Decission tree*) sebagaimana pada gambar berikut.



Gambar 4. Pohon keputusan

Pada gambar 4 diatas menggambarkan hasil klasifikasi dari pencarian nilai *entropy*, *gain*, *split information* dan *gain rasio* tertinggi dari masing – masing atribut yang ada pada data aktual mahasiswa. Maka di dapatkanlah pohon keputusan untuk klasifikasi ketepatan waktu kelulusan mahasiswa dengan hasil :

1. Nilai MK Skripsi E = 0 = Tidak tepat waktu (*root*)
2. Rentang Nilai IPK 3 05 – 3,20 = Tepat waktu
3. Rentang SKS < 100 = Tidak Tepat Waktu
4. Nilai KKN/Magang/KP = 0 = Tidak Tepat Waktu

IV. Kesimpulan

Kesimpulan yang dapat diambil dari penelitian ini adalah :

1. Metode algoritma C4.5 dapat digunakan untuk melakukan klasifikasi ketepatan waktu kelulusan mahasiswa.
2. Pada sistem ini pengklasifikasian digunakan dengan mengolah data jawaban dari data nilai tiap-tiap variabel nilai mahasiswa.
3. Pada pengujian klasifikasi menggunakan algoritma C4.5 setelah mencari seluruh nilai *entropy* mendapatkan nilai *entropy* terbesar yaitu pada parameter Nilai IPK semester 7 pada rentang 3,50 – 3,20 dengan nilai 0,918295834, lalu mendapatkan nilai *gain* populer pada parameter nilai Skripsi semester 7 E = 0 dengan nilai 0,068394234, mendapatkan nilai *split information* pada parameter nilai IPK Semester 7, jumlah SKS semester 7, Nilai KKN/Magang/KP semester 7 dan Nilai Skripsi semester 7 yaitu 0, dan nilai *gain rasio* pada parameter nilai IPK Semester 7, jumlah SKS semester 7, Nilai KKN/Magang/KP semester 7 dan Nilai Skripsi semester 7 yaitu 0, maka dihasilkanlah pohon keputusan dari pengujian data tersebut.
4. Pada pengujian yang dilakukan dengan metode algoritma C4.5 didapatkan tingkat hasil akurasi sebesar 95%.

V. Daftar Pustaka

- Budiyantara, A., & Irwansyah., 2019, Prediksi Mahasiswa Lulus Tepat Waktu Menggunakan Algoritma Decision Tree (C 4.5), *Infotec*, Jakarta. C 4.5, *Jurnal Media Infotama*, 130 - 138.
- Darmawan, E., 2018, C4.5 Algorithm Application for Prediction of Self Candidate New Students In Higher Education, *JOIN (Jurnal Online Informatika)*, 22 - 28
- Han, J., Kamber, M., Pei, J., 2011, Data Mining : Concepts And Techniques, 3rd ed. p cm, ISBN 978-0-12-381479-1, *Morgan Kaufman Publisher*, USA.
- Haryati, S., Sudarsono, A., Suryana, E., 2015, Implementasi Data Mining Untuk Memprediksi Masa Studi Mahasiswa Menggunakan Algoritma
- Jaman, J. H., 2013, Prediksi Kelulusan Mahasiswa Dengan Metode Algoritma C 4.5, *SYNTAX* , 39 - 43.
- Martanto., Irfan, A., Mulyawan., 2019, Prediksi Tingkat Kelulusan Mahasiswa Menggunakan Machine Learning dengan Teknik Deep Learning, *Jurnal Informatika: Jurnal Pengembangan IT Poltek Tegal*, 10.30591/jpit.v4i2-2.1877, 191-194
- Nofriansyah, D., & Nurcahyo, G. W., 2015, Algoritma Data Mining dan Pengujian, ISBN 878-602-2018-666-0, *Deepublisher*.
- Rohman, A., & Rufiyanto, A., 2019, Implementasi Data Mining dengan Algoritma Decision Tree (C 4.5) Untuk Prediksi Kelulusan Mahasiswa, *Proceeding SYNTAK* , 134 - 139.
- Universitas Papua., 2018, Tentang Beasiswa Bidikmisi, <http://unipa.ac.id/?p=204611>, accessed 29 January 2021
- Widianingsih, S., 2019, Perbandingan Metode Data Mining Untuk Prediksi Nilai Dan Waktu Kelulusan Mahasiswa Prodi Teknik Informatika Dengan Algoritma C4.5, Naïve Bayes, Knn, Dan Svm, *Jurnal Tekno Insentif*, Halaman 16-25.
- Yang, F., 2018, Decision Tree Algorithm Based University Graduate Employment Trend Prediction, *Informatika* 43, 573 - 580.

VI. Ucapan Terimakasih

Penulis mengucapkan terima kasih yang sebesar-besarnya dan penghargaan yang mulia kepada :

1. Allah SWT, Kedua orang tua dan keluarga yang telah memberikan kekuatan dan pertolongan demi kelancaran pengerjaan tugas akhir dan jurnal ini
2. Ibu Eka Suswaini S.T., M.T. dan Ibu Alena Uperiati, S.T., M.Cs. sebagai Pembimbing 1 dan 2, bapak Tekad Matulatan, S. Sos., S. Kom., M.Inf.Tech sebagai Penasehat Akademik saya
3. Seluruh tim penguji, teman, sahabat, dan rekan rekan yang juga sudah memberikan support.