

## FAKTOR-FAKTOR YANG BERPENGARUH UNTUK MENGESTIMASI WAKTU KELULUSAN MAHASISWA DENGAN PENDEKATAN ALGORITMA C4.5

Dewi Arianti Wulandari<sup>1\*</sup>, Herman Bedi<sup>2</sup>, Luqman<sup>3</sup>, Ahmad Arfan Mashuda<sup>4</sup>

<sup>1,2,3,4</sup>Teknik Informatika, Fakultas Telematika Energi, Institut Teknologi PLN, DKI Jakarta, Indonesia

Email: <sup>1\*</sup>dewiarianti@itpln.ac.id, <sup>2</sup>hermanbedi@itpln.ac.id, <sup>3</sup>luqman@itpln.ac.id, <sup>4</sup>arfan1831178@itpln.ac.id

(\*: corresponding author)

**Abstrak-** Faktor – faktor yang berpengaruh pada kelulusan mahasiswa untuk menyelesaikan studi di perguruan tinggi antara lain adalah nilai akademik, keaktifan dalam berorganisasi, asal sekolah, motivasi dan kesibukan dalam bekerja. Tujuan dari penelitian ini adalah untuk mengidentifikasi komponen yang mempengaruhi waktu kelulusan mahasiswa. Hasil yang didapat ternyata faktor Indeks Prestasi Kumulatif (IPK), asal sekolah menengah atas, aktif organisasi, aktif Unit Kegiatan Mahasiswa (UKM), sudah bekerja dan motivasi berpengaruh pada tingkat kelulusan mahasiswa baik yang tepat waktu maupun yang tidak. Dan berdasarkan pengujian menggunakan *confussion matrix* didapatkan hasil akurasi lebih dari 70% yang artinya dengan algoritma C4.5 dan model yang dihasilkan dapat mengestimasi dengan baik apakah mahasiswa dapat lulus tepat waktu atau tidak dengan faktor-faktor tersebut.

**Kata Kunci:** Algoritma C4.5, Data Mining, Prediksi, Kelulusan, Mahasiswa

**Abstract-** Factors that influence student graduation to complete studies at university include academic grades, activeness in organizations, school origin, motivation, and busyness in work. The purpose of this study was to identify the components that influence the graduation time of students. The results obtained turned out to be cumulative achievement index (GPA) factors, high school origin, active organizations, active student activity units (UKM), already working and motivation affect student graduation rates both on time and those who are not. Based on testing using a confusion matrix, accuracy results of more than 70% were obtained, which means that with the C4.5 algorithm, the resulting model can estimate well whether students can graduate on time or not with these factors.

**Keywords:** Algorithm C4.5, Data Mining, Prediction, Graduation, Student

### 1. PENDAHULUAN

Pada dunia yang kompetitif sekarang ini, sangat penting bagi sebuah lembaga untuk meramalkan kinerja siswa, mengklasifikasikan individu berdasarkan bakat mereka, dan berusaha untuk meningkatkan kinerja mereka dalam tes di masa depan. Siswa harus disarankan jauh sebelumnya untuk memusatkan upaya mereka di bidang tertentu untuk meningkatkan prestasi akademik mereka [1]. Di lembaga pendidikan tinggi, tingkat kelulusan adalah salah satu dari banyak aspek untuk menilai kesetaraan proses pembelajaran[2]. Keberhasilan dalam memprediksi prestasi akademik siswa sejak dini akan memudahkan guru dalam memberikan bantuan kepada siswa dengan prestasi akademik di bawah rata – rata kelas atau yang kesulitan mengikuti proses pembelajaran di kelas [3]. Putus sekolah siswa Brasil dari pendidikan tinggi adalah subjek yang telah *dieksplorasi* dengan baik, di mana tingginya tingkat siswa yang putus sekolah *diverifikasi*. Masalah yang timbul dari putus sekolah siswa masih belum memiliki solusi karena putus sekolah itu sendiri adalah masalah yang belum terpecahkan [4]. Siswa yang masuk dalam jumlah besar, tetapi siswa yang lulus tepat waktu di bawah standar yang ditentukan. Akibatnya, ada akumulasi siswa yang sangat besar di setiap periode kelulusan. Untuk mengatasi hal tersebut membutuhkan sistem berbasis *data mining* dalam memantau atau memanfaatkan perkembangan mahasiswa dalam memprediksi kelulusan menggunakan algoritma C4.5[5]. Memfasilitasi proses pembelajaran yang efektif adalah tujuan dari lembaga pendidikan tinggi. Meskipun ada peningkatan dalam kurikulum dan sumber daya, banyak siswa tidak dapat lulus tepat waktu. Sebagian besar, jumlah lulusan yang tepat waktu lebih rendah dari jumlah murid baru yang mendaftar ke universitas [6]. Ada 9 atribut yang digunakan untuk memprediksi faktor ketepatan waktu, dan metode yang dimaksud adalah algoritma C4.5 dan atribut hasil IPK adalah akar tertinggi yang memberikan pengaruh tertinggi[7].

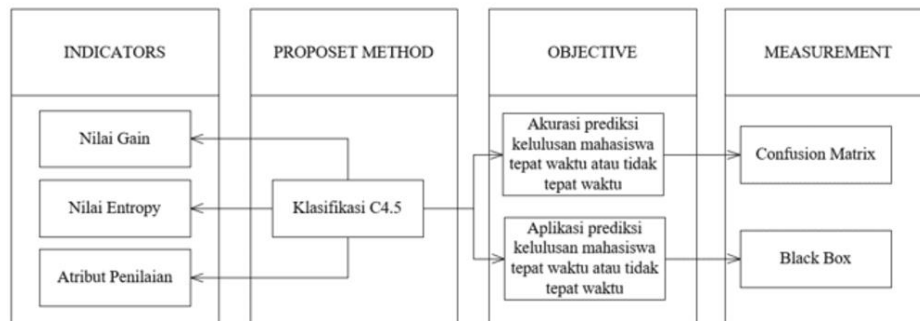
Memperkenalkan algoritma C4.5 ke dalam bidang ketenagakerjaan lulusan, mengatur dan menganalisis sejumlah besar data historis yang dikumpulkan, membangun pohon keputusan klasifikasi, menghasilkan set aturan, membangun model prediksi pekerjaan lulusan[8]. Metode pohon keputusan dan algoritma C4.5 sebagai pendukung, bersama dengan data mining, digunakan untuk memprediksi tingkat kelulusan [9]. Untuk mengetahui prediksi waktu lulus calon mahasiswa di akhir semester menggunakan algoritma C4.5[10]. Tinjauan literatur yang komprehensif untuk meramalkan kinerja siswa menggunakan teknik penambangan data khususnya menggunakan algoritma C4.5 direkomendasikan untuk meningkatkan prestasi siswa[11].

Hasil dari penelitian ini dapat mengestimasi memprediksi jumlah lulusan yang lulus tepat waktu berdasarkan variabel nilai Indeks Prestasi Kumulatif (IPK), asal sekolah menengah atas, aktif organisasi, aktif Unit Kegiatan Mahasiswa (UKM), sudah bekerja dan motivasi. Sehingga hasil dari penelitian ini dapat membantu institusi untuk membuat kebijakan terkait tentang kelulusan agar mahasiswanya dapat lulus tepat waktu.

## 2. METODE PENELITIAN

### 2.1 Kerangka Penelitian

Kerangka penelitian terdiri dari 4 tahapan seperti yang ditunjukkan pada gambar 1 yaitu, tahapan 1 penetapan indikator (*indicators*) yang terdiri dari nilai *gain*, nilai *entropy* dan atribut penilaian, tahapan 2 penentuan metode (*proposed method*) yang akan digunakan untuk penelitian yaitu algoritma C4.5, tahapan 3 menentukan tujuan penelitian (*objective*) yaitu untuk memperkirakan apakah mahasiswa akan lulus pada waktunya atau tidak dan yang terakhir adalah tahapan *measurement* yaitu pengukuran hasil dari penelitian pada tahapan ini menggunakan metode *confusion matrix* untuk pengujian akurasinya dan untuk pengujian aplikasinya menggunakan metode *black box testing*.



Gambar 1. Kerangka Penelitian

### 2.2 Algoritma C4.5

Pola penyelesaian prediksi dibuat dengan data mining, selanjutnya data dikelompokkan dengan cara diklasifikasikan untuk dapat memprediksi kelulusan tepat waktu dan terlambat dengan algoritma C4.5[12][13]. Algoritma C4.5 dapat diaplikasikan untuk memprediksi masa studi mahasiswa dengan baik [14][15].

Proses pengulangan metode *Decision Tree* ini berakhir ketika semua data dibagi rata, tidak ada lagi atribut yang dapat dibagi, dan tidak ada data dalam cabang yang kosong. Model algoritma C4.5 yang digunakan pada penelitian ini terdiri dari langkah-langkah menentukan variabel inputan yaitu Nim, IPK, Asal Sekolah, Aktif Organisasi, Aktif UKM, Bekerja dan Motivasi, kemudian dilakukan proses menghitung nilai *entropy* dan *gain* setiap atribut, selanjutnya membuat simpul akar pohon dan nilai *gain* tertinggi. Tahapan berikutnya mencari apakah ada atribut yang *ambigu* kalau ada dilanjutkan dengan menghitung kembali sampai dihasilkan atribut yang terpilih setelah itu membuat simpul akar pohon dari *gain* tertinggi. Sementara jika tidak maka proses selesai, tahapan ini dapat dilihat pada gambar 2.

Berikut adalah rumus yang digunakan untuk menghitung *gain* dan *entropy*.

Untuk menghitung *gain*, rumus berikut digunakan:

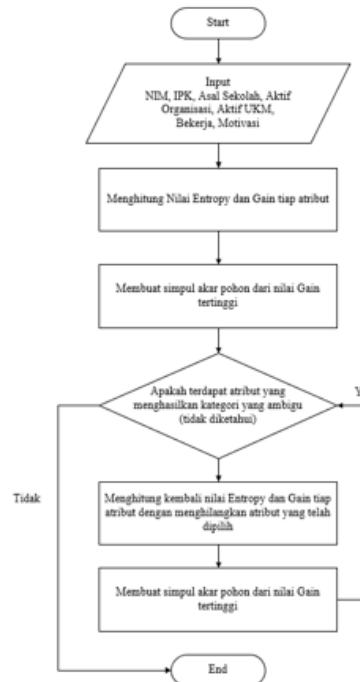
$$Gain(S, A) = Entropy(S) - \sum_{i=1}^n \frac{|S_i|}{|S|} * Entropy(S_i) \quad (1)$$

Rumus berikut dapat digunakan untuk menghitung nilai *entropy*:

$$Entropy(S) = \sum_{i=1}^n - p_i * \log_2 p_i \quad (2)$$

*Confusion Matrix* merupakan salah satu metode untuk menguji tingkat keakuratan metode *data mining*. Akurasi merupakan presentase dari total data yang diidentifikasi dan dinilai. Rumusnya adalah sebagai berikut:

$$Akurasi = \frac{(TP+TN)}{(TP+TN+FP+FN)} * 100\% \quad (3)$$



**Gambar 2.** Model Algoritma C4.5

### 3. HASIL DAN PEMBAHASAN

Data yang digunakan dalam penelitian ini adalah data mahasiswa IT-PLN angkatan 2016-2017 sebagai data *training* dan data mahasiswa angkatan 2019 sebagai data *testing*. Adapun data kelulusan mahasiswa tahun 2016-2017 sebanyak 664 data yang terdiri dari 6 atribut yang akan dijadikan sebagai variabel perhitungan. Deskripsi dari atribut yang terdapat pada data set yang digunakan sebagai data *training* seperti yang tercantum dalam tabel 1.

**Tabel 1.** Atribut-Atribut Variabel

Atribut	Proses	Keterangan	Skala Penilaian
IPK	IPK	Nilai Indeks Prestasi Kumulatif Semester 4	A= IPK lebih besar sama dengan 3.5 B= IPK lebih besar dari 3&IPK kurang dari 3.5 C=IPK lebih besar sama dengan 2.5 & IPK kurang dari 3 D=IPK kurang dari 2.5
Asal Sekolah	Asal Sekolah	Penilaian terhadap asal sekolah	- Negeri - Swasta
Aktif Organisasi	Aktif Organisasi	Penilaian terhadap aktif atau tidak dalam organisasi	- Aktif - Tidak Aktif
Aktif Kegiatan Mahasiswa Bekerja	Aktif UKM Status Bekerja	Penilaian terhadap aktif atau tidak dalam Kegiatan Mahasiswa Penilaian terhadap status bekerja atau tidak saat menjadi mahasiswa	- Aktif - Tidak Aktif - Bekerja - Tidak Bekerja
Motivasi	Motivasi dalam menjalankan kuliah	Penilaian terhadap motivasi mahasiswa dalam menjalankan kuliah	- Kuat - Cukup - Lemah
Kategori	Label	Keputusan	- Lulusan Tepat Waktu (LTW) - Lulusan Tidak Tepat Waktu (LTTW)

Berdasarkan data yang diperoleh, kemudian dilakukan proses klasifikasi dengan data *training* yang tercantum dalam tabel 2.

**Tabel 2.** Data Hasil Klasifikasi

NIM	IPK	ASAL SEKOLAH	AKTIF ORGANISASI	AKTIF UKM	BEKERJA	MOTIVASI	KET (Tepat Waktu: TW; Tidak Tepat Waktu: TTW)
20161	2.96	Negeri	Aktif	Aktif	Bekerja	Kuat	TW
20162	3.37	Negeri	Aktif	Aktif	Bekerja	Lemah	TTW
20164	3.14	Negeri	Aktif	Tidak Aktif	Bekerja	Cukup	TTW
20166	3.46	Swasta	Aktif	Aktif	Bekerja	Lemah	TTW
20168	3.42	Swasta	Aktif	Tidak Aktif	Tidak Bekerja	Kuat	TW
20169	3.23	Negeri	Tidak Aktif	Aktif	Tidak Bekerja	Kuat	TW
201610	3.15	Swasta	Aktif	Aktif	Tidak Bekerja	Lemah	TTW
.....	...	.....	.....	.....	.....	.....	.....
X <sub>n</sub>	X <sub>n</sub>	X <sub>n</sub>	X <sub>n</sub>	X <sub>n</sub>	X <sub>n</sub>	X <sub>n</sub>	X <sub>n</sub>

Setelah diklasifikasikan selanjutnya data masuk ke tahap *transformasi* berdasarkan skala penilaian dari masing-masing atribut yang sudah ditentukan pada tahap data *understanding*, yang terlihat pada tabel 3.

**Tabel 3.** Data Hasil Transformasi

NIM	IPK	ASAL SEKOLAH	AKTIF ORGANISASI	AKTIF UKM	BEKERJA	MOTIVASI	KET (Tepat Waktu: TW; Tidak Tepat Waktu: TTW)
20161	C	Negeri	Aktif	Aktif	Bekerja	Kuat	TW
20162	B	Negeri	Aktif	Aktif	Bekerja	Lemah	TTW
20164	B	Negeri	Aktif	Tidak Aktif	Bekerja	Cukup	TTW
20166	B	Swasta	Aktif	Aktif	Bekerja	Lemah	TTW
20168	B	Swasta	Aktif	Tidak Aktif	Tidak Bekerja	Kuat	TW
20169	B	Negeri	Tidak Aktif	Aktif	Tidak Bekerja	Kuat	TW
201610	B	Swasta	Aktif	Aktif	Tidak Bekerja	Lemah	TTW
.....	...	.....	.....	.....	.....	.....	.....
X <sub>n</sub>	X <sub>n</sub>	X <sub>n</sub>	X <sub>n</sub>	X <sub>n</sub>	X <sub>n</sub>	X <sub>n</sub>	X <sub>n</sub>

Selanjutnya dicari nilai *entropy* dari masing-masing atribut, seperti yang ditunjukkan dalam tabel 4.

**Tabel 4.** Hasil Nilai Entropy Dari Masing-Masing Atribut

	Jumlah Data	Tepat Waktu	Tidak Tepat Waktu	Entropy
Total	664	368	298	$(-\frac{368}{664} * \log_2(\frac{368}{664})) + (-\frac{296}{664} * \log_2(\frac{396}{664}))$
<b>IPK</b>				
A	175	114	61	$(-\frac{115}{175} * \log_2(\frac{115}{175})) + (-\frac{61}{175} * \log_2(\frac{61}{175}))$
B	375	225	150	$(-\frac{225}{375} * \log_2(\frac{225}{375})) + (-\frac{150}{375} * \log_2(\frac{150}{375}))$
C	52	29	23	$(-\frac{29}{52} * \log_2(\frac{29}{52})) + (-\frac{23}{52} * \log_2(\frac{23}{52}))$
D	62	0	62	$(-\frac{0}{62} * \log_2(\frac{0}{62})) + (-\frac{62}{62} * \log_2(\frac{62}{62}))$
<b>Asal Sekolah</b>				
Negeri	519	287	232	$(-\frac{287}{519} * \log_2(\frac{297}{519})) + (-\frac{232}{519} * \log_2(\frac{232}{519}))$
swasta	145	81	64	$(-\frac{81}{145} * \log_2(\frac{81}{145})) + (-\frac{64}{146} * \log_2(\frac{64}{145}))$
<b>Aktif Organisasi</b>				
Aktif	366	184	182	$(-\frac{184}{366} * \log_2(\frac{184}{366})) + (-\frac{182}{366} * \log_2(\frac{182}{366}))$

	Jumlah Data	Tepat Waktu	Tidak Tepat Waktu	Entropy	
Tidak Aktif	298	184	114	$(-\frac{184}{298} * \log_2(\frac{184}{298})) + (-\frac{114}{298} * \log_2(\frac{114}{298}))$	0.96
Aktif UKM					
Aktif	378	184	194	$(-\frac{184}{378} * \log_2(\frac{184}{378})) + (-\frac{194}{378} * \log_2(\frac{194}{378}))$	1.00
Tidak Aktif	286	184	102	$(-\frac{184}{286} * \log_2(\frac{184}{286})) + (-\frac{102}{286} * \log_2(\frac{102}{286}))$	0.94
Bekerja					
Bekerja	331	149	182	$(-\frac{149}{331} * \log_2(\frac{149}{331})) + (-\frac{182}{331} * \log_2(\frac{182}{331}))$	0.99
Tidak Aktif	333	219	114	$(-\frac{219}{333} * \log_2(\frac{219}{333})) + (-\frac{114}{333} * \log_2(\frac{114}{333}))$	0.93
Motivasi					
Lemah	160	0	160	$(-\frac{0}{160} * \log_2(\frac{0}{160})) + (-\frac{160}{160} * \log_2(\frac{160}{160}))$	0.00
Cukup	175	104	71	$(-\frac{104}{175} * \log_2(\frac{104}{175})) + (-\frac{71}{175} * \log_2(\frac{71}{175}))$	0.97
Kuat	329	264	65	$(-\frac{264}{329} * \log_2(\frac{264}{329})) + (-\frac{65}{329} * \log_2(\frac{65}{329}))$	0.72

Setelah didapat nilai *entropy* setiap atribut, maka tindakan berikutnya adalah mencari nilai *gain* dari masing-masing atribut yang tercantum pada tabel 5.

**Tabel 5.** Nilai *Gain* Dari Masing-Masing Atribut

Atribut	Perhitungan	Nilai Gain
IPK	$= 0.99 - ((\frac{175}{664} \times 0.93) + (\frac{375}{664} \times 0.97) + (\frac{52}{664} \times 0.99) + (\frac{62}{664} \times 0.00))$	0.1197496
Asal Sekolah	$= 0.99 - ((\frac{519}{664} \times 0.99) + (\frac{145}{664} \times 0.99))$	0.0000158
Aktif Organisasi	$= 0.99 - ((\frac{366}{664} \times 1.00) + (\frac{298}{664} \times 0.96))$	0.0095447
Aktif Unit Kegiatan Mahasiswa	$= 0.99 - ((\frac{378}{664} \times 1.00) + (\frac{286}{664} \times 0.94))$	0.0176922
Bekerja	$= 0.99 - ((\frac{331}{664} \times 0.99) + (\frac{333}{664} \times 0.93))$	0.0316705
Motivasi	$= 0.99 - ((\frac{160}{664} \times 0.00) + (\frac{175}{664} \times 0.97) + (\frac{329}{664} \times 0.72))$	0.3794691

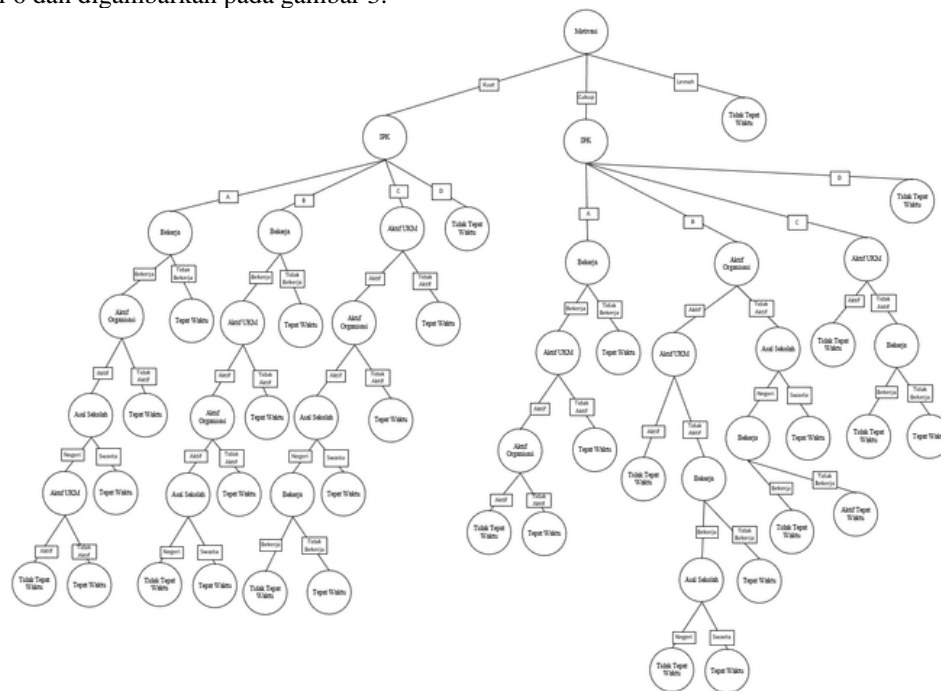
Pada tahapan terakhir menentukan *node* sebagai dasar dalam membuat model pohon keputusannya seperti yang disajikan pada tabel 6.

**Tabel 6.** Hasil Perhitungan Untuk Mencari *Node*

Node	Atribut					
	IPK	Asal Sekolah	Aktif Organisasi	Aktif UKM	Bekerja	Motivasi
1	0.1197496	0.0000158	0.0095447	0.0176922	0.0316705	0.3794691
1.1	0.2667320	0.0028229	0.0189690	0.0426429	0.0504623	-
1.2	0.1446527	0.0017942	0.0549157	0.0643267	0.0528878	-
1.1.1	-	0.0432382	0.1241411	0.1402609	0.0053631	-
1.1.2	-	0.0366561	0.1229057	0.1165491	0.1507829	-
1.1.3	-	0.0307622	0.0648999	0.0725758	0.0991791	-
1.2.1	-	0.1379254	0.3112781	0.5487949	0.3112781	-
1.2.2	-	0.0162843	0.0715792	0.0450532	0.0113299	-
1.2.3	-	0.0002792	0.0577201	0.1899162	0.4163430	-
1.1.1.1	-	0.0851236	0.2340681	-	0.0003791	-
1.1.2.1	-	0.1078610	0.2488649	0.2590615	-	-
1.1.3.1	-	0.0571049	0.1999446	0.1187148	-	-
1.2.1.1	-	0.0729056	0.1709506	-	0.0316705	-

Node	Atribut					
	IPK	Asal Sekolah	Aktif Organisasi	Aktif UKM	Bekerja	Motivasi
1.2.2.1	-	0.1185794	-	0.0044832	0.0316705	-
1.2.2.2	-	0.0000578	-	0.1879772	0.0089162	-
1.2.3.1	-	0.0105230	0.1388147	0.3810602	-	-
1.1.1.1.1	-	0.2488424	-	-	0.0069003	-
1.1.2.1.1	-	0.3159853	0.4468034	-	-	-
1.1.3.1.1	-	0.1864743	-	0.1431085	-	-
1.2.2.1.1	-	-	-	0.0272064	0.0316705	-
1.2.2.2.1	-	0.0094968	-	-	0.1350424	-
1.2.3.1.1	-	0.0153588	0.3210238	-	-	-
1.1.1.1.1.1	-	-	-	-	0.0058021	-
1.1.2.1.1.1	-	0.8571484	-	-	-	-
1.1.3.1.1.1	-	-	-	0.2128966	-	-
1.2.2.2.1.1	-	0.0467848	-	-	-	-

Langkah selanjutnya adalah membuat pohon keputusan berdasarkan hasil pencarian *node* yang terdapat pada tabel 6 dan digambarkan pada gambar 3.



Gambar 3. Pohon Keputusan

Gambar 3 menjelaskan tentang pohon keputusan yang dihasilkan dari proses klasifikasi sehingga menghasilkan sebanyak 32 *rules* yang akan dijadikan dasar keputusan kelulusan mahasiswa.

### 3.1. Pengujian dan Pembahasan

Pada tahapan ini dilakukan pengujian, dan hasilnya dapat dilihat pada tabel 7, 8 dan 9.

Tabel 7. Data Uji

NIM	IPK	Asal Sekolah	Aktif Organisasi	Aktif UKM	Bekerja	Motivasi	Ket
20171435	B	Negeri	Aktif	Aktif	Bekerja	Kuat	TW
20171436	B	Swasta	Tidak Aktif	Aktif	Bekerja	Lemah	TTW
20171437	A	Negeri	Tidak Aktif	Aktif	Bekerja	Cukup	TTW
20171438	B	Negeri	Aktif	Tidak Aktif	Bekerja	Cukup	TW
20171439	D	Negeri	Tidak Aktif	Aktif	Bekerja	Kuat	TTW



NIM	IPK	Asal Sekolah	Aktif Organisasi	Aktif UKM	Bekerja	Motivasi	Ket
20171440	B	Negeri	Tidak Aktif	Aktif	Tidak Bekerja	Kuat	TW
Xn	Xn	Xn	Xn	Xn	Xn	Xn	Xn

Pengujian akurasi pada aplikasi dengan data *testing* yang digunakan pada penelitian ini adalah berjumlah 10% dari data *training* yaitu 64 data, data *training* harus lebih banyak dari data *testing* karena hasil analisa data *training* yang akan menjadi *rule* untuk diuji ke data *testing*, dimana data tersebut merupakan bagian dari data *training* untuk menguji akurasi yang telah dibuat pada tahap *modelling phase* dalam pembuatan keputusan dengan Algoritma C4.5. Setelah data *testing* sudah diprediksi langkah selanjutnya adalah membuat tabel *Confusion Matrix*.

**Tabel 8.** Hasil pengujian antara data aktual dengan hasil pengujian

Data	Data Aktual	Hasil Pengujian
Data Testing 1	LTTW	LTTW
Data Testing 2	LTTW	LTTW
Data Testing 3	LTTW	LTW
Data Testing 4	LTTW	LTW
Data Testing 5	LTTW	LTTW
Data Testing 6	LTW	LTW
Data Testing 7	LTTW	LTTW
Data Testing 8	LTW	LTTW
Data Testing 9	LTTW	LTTW
Data Testing 10	LTTW	LTTW
Data Testing 11	LTW	LTTW
Data Testing 12	LTTW	LTW
Data Testing 13	LTTW	LTTW
Data Testing 14	LTW	LTW
Data Testing 15	LTTW	LTTW
Data Testing 16	LTTW	LTTW
Data Testing 17	LTW	LTW
Data Testing 18	LTW	LTW
Data Testing 19	LTTW	LTW
Data Testing 20	LTTW	LTTW
Data Testing 21	LTW	LTW
Data Testing 22	LTTW	LTTW
Data Testing 23	LTTW	LTTW
Data Testing 24	LTW	LTW
Data Testing 25	LTTW	LTTW
Data Testing 26	LTW	LTW
Data Testing 27	LTTW	LTTW
Data Testing 28	LTW	LTW
Data Testing 29	LTTW	LTW
Data Testing 30	LTTW	LTTW
Data Testing 31	LTTW	LTTW
Data Testing 32	LTW	LTW
Data Testing 33	LTW	LTW
Data Testing 34	LTW	LTW
Data Testing 35	LTTW	LTW
Data Testing 36	LTTW	LTTW
Data Testing 37	LTW	LTW
Data Testing 38	LTW	LTW
Data Testing 39	LTW	LTW
Data Testing 40	LTW	LTW
Data Testing 41	LTW	LTW
Data Testing 42	LTW	LTW
Data Testing 43	LTTW	LTTW
Data Testing 44	LTTW	LTTW

Data	Data Aktual	Hasil Pengujian
Data Testing 45	LTW	LTW
Data Testing 46	LTW	LTW
Data Testing 47	LTW	LTW
Data Testing 48	LTTW	LTTW
Data Testing 49	LTTW	LTTW
Data Testing 50	LTTW	LTTW
Data Testing 51	LTW	LTTW
Data Testing 52	LTTW	LTTW
Data Testing 53	LTW	LTW
Data Testing 54	LTW	LTW
Data Testing 55	LTW	LTW
Data Testing 56	LTW	LTTW
Data Testing 57	LTTW	LTTW
Data Testing 58	LTW	LTW
Data Testing 59	LTW	LTTW
Data Testing 60	LTW	LTTW
Data Testing 61	LTW	LTW
Data Testing 62	LTW	LTW
Data Testing 63	LTTW	LTTW
Data Testing 64	LTTW	LTTW

Pada tabel 8, terdapat 12 data yang berbeda antara data aktual dengan data hasil pengujian. Hal ini dapat dibuktikan pada hasil pengujian menggunakan confusion matrix seperti yang terlihat pada tabel 9.

**Tabel 9.** Hasil confusion matrix

<i>Actual</i>	<i>Predicted Class</i>	
	Tepat Waktu	Tidak Tepat Waktu
Tepat Waktu	28	6
Tidak Tepat Waktu	5	25

$$\text{Akurasi} = \left( \frac{TP+TN}{TP+TN+FP+FN} \right) * 100\%$$

$$\text{Akurasi} = \left( \frac{28+25}{28+25+5+6} \right) * 100 \% = 82.81 \%$$

Pada tahap *modelling* dihasilkan *rule* sebanyak 32 *rule* yang akan digunakan sebagai dasar keputusan untuk prediksi kelulusan mahasiswa yang tercantum dalam tabel 10.

**Tabel 10.** Hasil Rule

No	Rule
1	Apabila Motivasi Lemah maka Lulusan Tidak Tepat Waktu
2	Apabila Motivasi Cukup dan IPK D maka Lulusan Tidak Tepat Waktu
3	Apabila Motivasi Cukup, IPK C dan Aktif UKM maka Lulusan Tidak Tepat Waktu
4	Apabila Motivasi Cukup, IPK C, Aktif UKM dan Tidak Bekerja maka Lulusan Tepat Waktu
5	Apabila Motivasi Cukup, IPK C, Aktif UKM dan Bekerja maka Lulusan Tidak Tepat Waktu
6	Apabila Motivasi Cukup, IPK B, Tidak Aktif Organisasi dan Asal Sekolah Swasta maka Lulusan Tepat Waktu
7	Apabila Motivasi Cukup, IPK B, Tidak Aktif Organisasi, Asal Sekolah Negeri dan Tidak Bekerja maka Lulusan Tepat Waktu
8	Apabila Motivasi Cukup, IPK B, Tidak Aktif Organisasi, Asal Sekolah Negeri dan Bekerja maka Lulusan Tidak Tepat Waktu
9	Apabila Motivasi Cukup, IPK B, Aktif Organisasi, dan Aktif UKM maka Lulusan Tidak Tepat Waktu
10	Apabila Motivasi Cukup, IPK B, Aktif Organisasi, Tidak Aktif UKM, dan Tidak Bekerja maka Lulusan Tepat Waktu



No	Rule
11	Apabila Motivasi Cukup, IPK B, Aktif Organisasi, Tidak Aktif UKM, Bekerja dan Asal Sekolah Swasta maka Lulusan Tepat Waktu
12	Apabila Motivasi Cukup, IPK B, Aktif Organisasi, Tidak Aktif UKM, Bekerja dan Asal Sekolah Negeri maka Lulusan Tidak Tepat Waktu
13	Apabila Motivasi Cukup, IPK A dan Tidak Bekerja maka Lulusan Tepat Waktu
14	Apabila Motivasi Cukup, IPK A, Bekerja, dan Tidak Aktif UKM maka Lulusan Tepat Waktu
15	Apabila Motivasi Cukup, IPK A, Bekerja, Aktif UKM dan Tidak Aktif Organisasi maka Lulusan Tepat Waktu
16	Apabila Motivasi Cukup, IPK A, Bekerja, Aktif UKM dan Aktif Organisasi maka Lulusan Tidak Tepat Waktu
17	Apabila Motivasi Kuat dan IPK D maka Lulusan Tidak Tepat Waktu
18	Apabila Motivasi Kuat, IPK C, dan Tidak Aktif UKM maka Lulusan Tepat Waktu
19	Apabila Motivasi Kuat, IPK C, Aktif UKM dan Tidak Aktif Organisasi maka Lulusan Tepat Waktu
20	Apabila Motivasi Kuat, IPK C, Aktif UKM, Aktif Organisasi dan Asal Sekolah Swasta maka Lulusan Tepat Waktu
21	Apabila Motivasi Kuat, IPK C, Aktif UKM, Aktif Organisasi, Asal Sekolah Negeri dan Tidak Bekerja maka Lulusan Tepat Waktu
22	Apabila Motivasi Kuat, IPK C, Aktif UKM, Aktif Organisasi, Asal Sekolah Negeri dan Bekerja maka Lulusan Tidak Tepat Waktu
23	Apabila Motivasi Kuat, IPK B dan Tidak Bekerja maka Lulusan Tepat Waktu
24	Apabila Motivasi Kuat, IPK B, Bekerja dan Tidak Aktif UKM maka Lulusan Tepat Waktu
25	Apabila Motivasi Kuat, IPK B, Bekerja, Aktif UKM dan Tidak Aktif Organisasi maka Lulusan Tepat Waktu
26	Apabila Motivasi Kuat, IPK B, Bekerja, Aktif UKM, Aktif Organisasi dan Asal Sekolah Swasta maka Lulusan Tepat Waktu
27	Apabila Motivasi Kuat, IPK B, Bekerja, Aktif UKM, Aktif Organisasi dan Asal Sekolah Negeri maka Lulusan Tidak Tepat Waktu
28	Apabila Motivasi Kuat, IPK A dan Tidak Bekerja maka Lulusan Tepat Waktu
29	Apabila Motivasi Kuat, IPK A, Bekerja dan Tidak Aktif Organisasi maka Lulusan Tepat Waktu
30	Apabila Motivasi Kuat, IPK A, Bekerja, Aktif Organisasi dan Asal Sekolah Swasta maka Lulusan Tepat Waktu
31	Apabila Motivasi Kuat, IPK A, Bekerja, Aktif Organisasi, Asal Sekolah Negeri dan Tidak Aktif UKM maka Lulusan Tepat Waktu
32	Apabila Motivasi Kuat, IPK A, Bekerja, Aktif Organisasi, Asal Sekolah Negeri dan Aktif UKM maka Lulusan Tidak Tepat Waktu

#### 4. KESIMPULAN

Hasil penelitian ini adalah untuk mengidentifikasi komponen yang mempengaruhi tingkat kelulusan mahasiswa. Berdasarkan hasil perhitungan yang dilakukan melalui algoritma C4.5 ternyata faktor Indeks Prestasi Kumulatif (IPK), asal sekolah menengah atas, aktif organisasi, aktif Unit Kegiatan Mahasiswa (UKM), sudah bekerja dan motivasi berpengaruh pada tingkat kelulusan mahasiswa baik yang tepat waktu maupun yang tidak. Berdasarkan pengujian menggunakan *confusion matrix* didapatkan hasil akurasi lebih dari 70% yang artinya dengan algoritma C4.5, model yang dikembangkan dapat memprediksi dengan cermat apakah mahasiswa akan lulus pada waktunya dengan faktor-faktor tersebut. Dimana pada tahap *modelling* dihasilkan *rule* sebanyak 32 *rule* yang akan digunakan sebagai dasar keputusan untuk prediksi kelulusan mahasiswa dan hasil akurasi pengujian dari proses prediksi kelulusan yang mencakup 64 data *testing* dan 600 data *training*, dimana data *testing* diambil dari 10% dari data *training* didapatkan hasil akurasi sebesar 82,81% atau lebih besar dari 70% dengan pengujian menggunakan *confusion Matrix* ini sehingga algoritma C4.5 dikatakan dapat memprediksi dengan baik apakah mahasiswa akan lulus tepat waktu atau tidak.

## UCAPAN TERIMA KASIH

Kami ucapkan kepada Lembaga Penelitian dan Pengabdian Pada Masyarakat (LPPM) Institut Teknologi PLN yang telah banyak memberikan dukungan baik materil maupun moril sehingga penelitian ini dapat diselesaikan dengan baik.

## DAFTAR PUSTAKA

- [1] H. Pallathadka, A. Wenda, E. Ramirez-Asís, M. Asís-López, J. Flores-Albornoz, and K. Phasinam, "Classification and prediction of student performance data using various machine learning algorithms," *Materials Today: Proceedings*, vol. 80. pp. 3782–3785, 2023. doi: 10.1016/j.matpr.2021.07.382.
- [2] A. Mailana, A. A. Putra, S. Hidayat, and A. Wibowo, "Comparison of C4.5 Algorithm and Support Vector Machine in Predicting the Student Graduation Timeliness," *Jurnal Online Informatika*, vol. 6, no. 1. p. 11, 2021. doi: 10.15575/join.v6i1.608.
- [3] U. Pujiyanto, W. Agung Prasetyo, and A. Rakhmat Taufani, "Students Academic Performance Prediction with k-Nearest Neighbor and C4.5 on SMOTE-balanced data," *2020 3rd International Seminar on Research of Information Technology and Intelligent Systems, ISRITI 2020*. pp. 348–353, 2020. doi: 10.1109/ISRITI51436.2020.9315439.
- [4] A. M. Mariano, A. B. De Magalhães Lelis Ferreira, M. R. Santos, M. L. Castilho, and A. C. F. L. C. Bastos, "Decision trees for predicting dropout in Engineering Course students in Brazil," *Procedia Computer Science*, vol. 214, no. C. pp. 1113–1120, 2022. doi: 10.1016/j.procs.2022.11.285.
- [5] D. Kurniawan, A. Anggrawan, and H. Hairani, "Graduation Prediction System On Students Using C4.5 Algorithm," *MATRIK : Jurnal Manajemen, Teknik Informatika dan Rekayasa Komputer*, vol. 19, no. 2. pp. 358–365, 2020. doi: 10.30812/matrik.v19i2.685.
- [6] H. Yuliansyah, R. A. P. Imaniati, A. Wirasto, and M. Wibowo, "Predicting Students Graduate on Time Using C4.5 Algorithm," *Journal of Information Systems Engineering and Business Intelligence*, vol. 7, no. 1. p. 67, 2021. doi: 10.20473/jisebi.7.1.67-73.
- [7] Y. T. Samuel, J. J. Hutapea, and B. Jonathan, "Predicting the timeliness of student graduation using decision tree c4.5 algorithm in universitas advent Indonesia," *Proceedings of 2019 International Conference on Information and Communication Technology and Systems, ICTS 2019*. pp. 276–280, 2019. doi: 10.1109/ICTS.2019.8850948.
- [8] F. Zhou, L. Xue, Z. Yan, and Y. Wen, "Research on college graduates employment prediction model based on C4.5 algorithm," *Journal of Physics: Conference Series*, vol. 1453, no. 1. 2020. doi: 10.1088/1742-6596/1453/1/012033.
- [9] F. Wiza and B. Febriadi, "Classification Analysis Using C4.5 Algorithm To Predict The Level of Graduation of Nurul Falah Pekanbaru High School Students," *IJISTECH (International Journal Of Information System & Technology)*, vol. 2, no. 2. p. 43, 2019. doi: 10.30645/ijistech.v2i2.21.
- [10] A. Ramadhanu, S. Defit, and S. W. Kareem, "Hybrid Data Mining with the Combination of K-Means Algorithm and C4.5 to Predict Student Achievement," *International Journal of Artificial Intelligence Research*, vol. 5, no. 2. 2021. doi: 10.29099/ijair.v6i1.225.
- [11] L. I. P. Aji and A. Sunyoto, "An Implementation of C4.5 Classification Algorithm to Analyze student's Performance," *2020 3rd International Conference on Information and Communications Technology, ICOIACT 2020*. pp. 126–130, 2020. doi: 10.1109/ICOIACT50329.2020.9332088.
- [12] N. A. Prahastiwi, R. Andreswari, and R. Fauzi, "Students Graduation Prediction Based on Academic Data Record Using the Decision Tree Algorithm C4.5 Method," *JURTEKSI (Jurnal Teknologi dan Sistem Informasi)*, vol. 8, no. 3. pp. 295–304, 2022. doi: 10.33330/jurteks.v8i3.1680.
- [13] A. Bulololo and E. Bu'ulolo, "Implementasi Algoritma Data Mining J48 Untuk Klasifikasi Mahasiswa Yang Layak Mendapat Beasiswa PPA," *Building of Informatics, Technology and Science (BITS)*, vol. 4, no. 1. 2022. doi: 10.47065/bits.v4i1.1641.
- [14] Y. A. Gerhana, I. Fallah, W. B. Zulfikar, D. S. Maylawati, and M. A. Ramdhani, "Comparison of naive Bayes classifier and C4.5 algorithms in predicting student study period," *Journal of Physics: Conference Series*, vol. 1280, no. 2. 2019. doi: 10.1088/1742-6596/1280/2/022022.
- [15] X. Wang, C. Zhou, and X. Xu, "Application of C4.5 decision tree for scholarship evaluations," *Procedia Computer Science*, vol. 151. pp. 179–184, 2019. doi: 10.1016/j.procs.2019.04.027.