

Penerapan Data Mining Klasifikasi C4.5 dalam Menentukan Tingkat Stres Mahasiswa Akhir

Anggi Trifani

STIKOM Tunas Bangsa, Pematangsiantar, Indonesia

Agus Perdana Windarto

STIKOM Tunas Bangsa, Pematangsiantar, Indonesia

Hendry Qurniawan

STIKOM Tunas Bangsa, Pematangsiantar, Indonesia

Korespondensi penulis: anggitrifani127@gmail.com

Abstract. *This study aims to classify the level of stress experienced by final students at STIKOM Tunas Bangsa using the C4.5 data mining technique. By knowing student stress, the campus can provide appropriate treatment and treatment of final students. Research data sources were obtained from several 6th and 8th semester classes as many as 110 students, through interviews and distribution of questionnaires. The attributes used as assessment parameters to determine the level of stress experienced by final students at STIKOM Tunas Bangsa, Pematang Siantar include: Interpersonal (C1), Intrapersonal (C2), Academic (C3) and Environment (C4). The testing process of this research uses the help of RapidMiner software to create a decision tree. From the results of C4.5 processing using the help of RapidMiner software, the Interpersonal attribute (C1) is the most influential attribute on the stress level of Final Students and the performance data shown on the suitability of the C4.5 method, the accuracy is 87,88%.*

Keywords: *Classification, Data Mining, C4.5, Stress, Final Student*

Abstrak. Penelitian ini bertujuan untuk melakukan Klasifikasi dalam menentukan tingkat Stres yang dialami oleh Mahasiswa Akhir di STIKOM Tunas Bangsa dengan menggunakan teknik data mining C4.5. Dengan mengetahui stres mahasiswa, pihak kampus dapat memberikan *treatment* dan perlakuan yang tepat terhadap mahasiswa akhir. Sumber data penelitian diperoleh dari beberapa kelas semester 6 dan 8 sebanyak 110 mahasiswa, melalui wawancara dan pembagian kuesioner. Atribut yang digunakan sebagai parameter penilaian untuk mengetahui tingkat stres yang dialami oleh mahasiswa akhir di STIKOM Tunas Bangsa, Pematang siantar antara lain: Interpersonal (C1), Intrapersonal (C2), Akademik (C3) dan Lingkungan (C4). Proses pengujian penelitian ini menggunakan bantuan *software RapidMiner* untuk membuat pohon keputusan. Dari hasil pengolahan C4.5 dengan menggunakan bantuan *software RapidMiner* atribut Interpersonal (C1) menjadi atribut yang paling berpengaruh terhadap tingkat stres Mahasiswa Akhir dan data *performance* yang ditunjukkan terhadap kesesuaian metode C4.5 akurasi adalah 87,88%.

Kata kunci: Klasifikasi, Data Mining, C4.5, Stres, Mahasiswa Akhir.

LATAR BELAKANG

Stres telah menjadi masalah yang nyata dalam kehidupan sehari-hari pada manusia sehingga menjadi bagian yang tidak terpisahkan (Sagita, Fairuz, & Aisyah, 2021). Stres merupakan salah satu gangguan psikologis yang dapat muncul ketika seorang individu tidak punya kesiapan untuk menghadapi tuntutan akademis yang seharusnya ia hadapi (Ambarwati, Pinilih, & Astuti, 2019). Stres adalah istilah yang umum digunakan untuk menjelaskan kondisi perasaan yang tidak stabil, yang disebabkan oleh kemarahan, frustrasi, kelelahan atau tekanan. Bahwa stres secara teoritis dapat dipandang sebagai upaya untuk menahan reaksi fisiologis ketika menghadapi kondisi atau bahaya yang menekan, yang disebut sebagai *stressor* (Tapidingan & Paseru, 2020). Menurut penelitian (Setiana & Cristin Wiyani, 2017) terdapat tiga kategori stres, yaitu stres rendah, stres sedang, dan stres berat. Penelitian yang dilakukan oleh (Merry & Mamahit, 2020) mengungkapkan terdapat dua pembagian stres, yakni *eustress* dan *distress*. *Eustress* adalah sebuah pengalaman stres yang kemunculannya di tandai ketika seorang individu berhasil menghadapi *stressor* dan pengalaman yang dirasakan berupa kebahagiaan atau kesenangan. Berbanding terbalik dengan *eustress*, *distress* adalah sebuah pengalaman stres yang tidak memunculkan kebahagiaan tetapi bersifat ancaman.

Stres yang dialami oleh tiap individu memiliki perbedaan dan keragaman. Dilihat dari kondisi lingkungan, terdapat beberapa jenis stres, yakni stres rumah tangga, stres kerja, dan stres akademik. Jenis stres yang banyak dialami oleh mahasiswa adalah stres akademik, tuntutan akademik dan tanggung jawab pada mahasiswa dapat menjadi bagian stres yang sudah biasa dialami oleh mahasiswa, sehingga stres akademik menjadi permasalahan yang paling umum dan cukup kompleks bagi mahasiswa (Sagita et al., 2021). Lingkungan kampus menjadi salah satu penyebab dari kemunculan stres sehingga stres dianggap sebagai sebuah hal yang umum dialami oleh mahasiswa, khususnya mahasiswa semester akhir. Tugas kuliah yang memiliki tingkat kesulitan tinggi seperti skripsi sering dianggap sebagai salah satu beban oleh mahasiswa. Stres yang dialami oleh mahasiswa tingkat akhir juga dapat di picu oleh adanya kesulitan dan hambatan yang muncul dari pihak eksternal seperti dosen ataupun lainnya ketika mengerjakan skripsi atau tugas akhir (Suhandiah, Ayuningtyas, & Sudarmaningtyas, 2021).

KAJIAN TEORITIS

1. Data Mining

Data mining adalah proses yang menggunakan teknik statistik, matematika, kecerdasan buatan dan *machine learning* untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai database besar (Irnanda, 2020). Data mining juga dikenal dengan istilah *pattern recognition*, merupakan suatu metode yang digunakan untuk pengolahan data guna menemukan pola yang tersembunyi dari data yang diolah (Oktavia, Hardinata, & Irawan, 2020). Data mining merupakan metode pengolahan data berskala besar, oleh karena itu data mining ini memiliki peranan penting dalam bidang industri, keuangan, cuaca, ilmu dan teknologi.

2. Klasifikasi

Klasifikasi adalah proses untuk menemukan model atau fungsi yang menjelaskan atau membedakan konsep atau kelas data, dengan tujuan untuk dapat memperkirakan kelas dari suatu objek yang labelnya tidak diketahui. Model itu sendiri bisa berupa aturan “jika maka”, *decision tree* ataupun formula matematis (Febriani & Sulistiani, 2021).

Ada dua jenis model klasifikasi, yaitu (Sadewo, Windarto, Damanik, & Hartama, 2018):

- a) Pemodelan Deskriptif (*descriptive modelling*), digunakan untuk membedakan objek-objek dalam kelas-kelas yang berbeda.
- b) Pemodelan Prediktif (*predictive modelling*), digunakan untuk memprediksi label kelas *record* yang tidak diketahui.

3. Algoritma C4.5

Algoritma ini adalah penerus dari algoritma sebelumnya, yaitu algoritma ID3 yang digunakan untuk menghasilkan sebuah rule berupa pohon keputusan. Algoritma ini dapat digunakan untuk klasifikasi data dan disebut sebagai *classifier statistic*. Algoritma C4.5 merupakan algoritma klasifikasi atau pengelompokan yang bersifat prediktif. Pernyataan klasifikasi pada pohon keputusan (*decision tree*) terdapat pada cabang-cabangnya dan kelas-kelas atau segmen-segmennya terdapat pada daun-daunnya. Algoritma C4.5 adalah salah satu keluarga pohon keputusan yang dapat menghasilkan sebuah pohon keputusan dan aturan atau rule untuk tujuan meningkatkan prediksi akurasi. Selain itu, model C4.5 mudah digunakan dan dipahami sebagai aturan yang diturunkan dan memiliki teknik interpretasi yang sangat mudah (Nurellisa & Fitriyah, 2020)

Secara umum, alur proses algoritma C4.5 untuk membangun pohon keputusan dalam data mining adalah sebagai berikut (Sadewo et al., 2018) :

- a) Pilih atribut sebagai simpul akar.
- b) Buat cabang untuk tiap nilai
- c) Bagi kasus dalam cabang
- d) Ulangi proses untuk setiap cabang sampai semua kasus pada cabang memiliki kelas yang sama. Pemilihan atribut sebagai simpul, baik simpul akar (*root*) atau simpul internal didasarkan pada nilai *gain* tertinggi dari atribut yang ada.

Ada beberapa tahap dalam membuat sebuah pohon keputusan dengan algoritma C4.5, yaitu (Asmaul Husnah Nasrullah, 2018)

- a) Menyiapkan data *training*. Data *training* biasanya diambil dari data histori yang pernah terjadi sebelumnya dan sudah dikelompokkan ke dalam kelas – kelas tertentu.
- b) Menentukan akar dari pohon. Akar akan diambil dari atribut yang terpilih, dengan cara menghitung nilai *gain* dari masing-masing atribut, nilai *gain* yang paling tinggi yang akan menjadi akar pertama. Sebelum menghitung nilai *gain* dari atribut, hitung dahulu nilai *entropy* menggunakan rumus 2.1 :

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

Di mana :

S : himpunan kasus

n : jumlah partisi S

p_i : proporsi S_i terhadap S

Setelah menghitung *entropy* dari setiap kasus selanjutnya menghitung nilai *gain* untuk pemisah objek dengan rumus 2.2.

$$Gain(S, A) = Entropy(S) - \sum_{i=1}^n \frac{|S_i|}{|S|} * Entropy(S_i) \quad (2)$$

Di mana :

S : himpunan kasus

A : Atribut

N : Jumlah partisi atribut A

$|S|$: Jumlah kasus pada partisi ke-1

$|S_i|$: Jumlah kasus dalam S

- c) Ulangi langkah ke-2 hingga semua record terpartisi.
- d) Proses partisi pohon keputusan akan berhenti saat:
 - 1. Semua record dalam simpul N mendapat kelas yang sama.
 - 2. Tidak ada atribut di dalam record yang dipartisi lagi.
 - 3. Tidak ada record di dalam cabang yang kosong.

METODE PENELITIAN

1. Lokasi dan Waktu Penelitian

Penelitian ini dilakukan di Sekolah Tinggi Ilmu Komputer (STIKOM) Tunas Bangsa, Pematangsiantar. Pengambilan data penelitian ini dimulai dari 27 Juni 2022 sampai 04 Juli 2022 dengan membagikan kuisisioner kepada mahasiswa akhir tersebut.

2. Diagram Rancangan Penelitian

Diperlukan beberapa langkah dalam penelitian ini untuk mencapai tujuan penelitian, penjelasan langkah-langkah pada rancangan penelitian ini dapat dilihat pada diagram berikut.



Gambar 1. Diagram Rancangan Penelitian

Berdasarkan rancangan penelitian pada Gambar 1, maka dapat diuraikan tahapan-tahapannya sebagai berikut:

- a) **Analisa Masalah**
Menganalisa suatu masalah mengenai tingkat stres mahasiswa akhir yang ada di STIKOM Tunas Bangsa, Pematang Siantar.
- b) **Teknik Pengumpulan Data**
Mengumpulkan data yang ingin diperoleh dari hasil dengan menyebarkan kuesioner kepada Mahasiswa akhir di Kampus STIKOM Tunas Bangsa, Pematang Siantar.

c) Menentukan Teknik Data Mining

Setelah permasalahan sudah dianalisa dan data sudah dikumpulkan, selanjutnya peneliti mengukur stres mahasiswa akhir.

3. Prosedur Pengumpulan Data

Metode pengumpulan data yang digunakan dalam penelitian ini terdapat beberapa metode sebagai berikut :

a) Penelitian Kepustakaan (*Library Research*)

Penelitian ini memanfaatkan perpustakaan untuk mencari buku-buku yang berkaitan dengan penelitian ini sebagai referensi untuk pengumpulan data.

b) Penelitian Lapangan (*Field Work Research*)

Penelitian yang dilakukan secara langsung di lapangan untuk mendapatkan data yang diinginkan dengan beberapa prosedur yaitu;

1. Wawancara

Yaitu proses pengumpulan data atau informasi melalui tatap muka antara pihak penanya (*interviewer*) dengan pihak yang ditanya atau penjawab (*interviewee*). Hal ini peneliti melakukan tanya jawab dengan mahasiswa akhir yang ada di Kampus STIKOM Tunas Bangsa, Pematangsiantar.

2. Kuisisioner

Kuisisioner adalah angket atau daftar pertanyaan yang tertulis secara tertutup atau responden hanya dapat memberikan jawaban dan secara terbuka atau responden dapat menuliskan pendapatnya berdasarkan kebutuhan data yang ingin ditanyakan. Untuk pengambilan jumlah sampel minimal. Setelah itu, peneliti membagikan kuisisioner pada mahasiswa akhir di Kampus STIKOM Tunas Bangsa, Pematangsiantar sebanyak 100 responden agar mendapatkan data yang akurat.

3. Studi Pustaka

Pada penelitian juga menggunakan teknik pengumpulan data studi pustaka, dilakukan dengan mempelajari teori-teori atau literatur dari buku-buku, jurnal, referensi-referensi yang berhubungan dengan penelitian ini untuk melengkapi data.

4. Observasi

Penulis melakukan pengamatan secara langsung terkait fenomena stres pada mahasiswa akhir di Kampus STIKOM Tunas Bangsa, Pematangsiantar.

4. Analisis Data

Pada penelitian ini, 100 sampel di ambil dari mahasiswa akhir di kampus STIKOM Tunas Bangsa, Pematangsiantar. Data yang digunakan dari hasil kuesioner kemudian diolah dengan melakukan perhitungan menggunakan algoritma C4.5, setelah itu dilakukan pengujian data menggunakan aplikasi *Rapidminer*. Berikut variabel yang akan digunakan dalam penelitian tingkat stres mahasiswa akhir, yaitu sebagai berikut :

a) Interpersonal, yaitu *stressor* yang dihasilkan dari hubungan dengan orang lain, misalnya konflik dengan orang tua, teman atau pacar. Variabel ini akan dikelompokkan menjadi 4 kategori yaitu : Sangat Setuju, Setuju, Tidak Setuju, Sangat Tidak Setuju.

b) Intrapersonal, yaitu *stressor* yang berasal dari dalam diri individu itu sendiri, misalnya kesulitan keuangan, perubahan kebiasaan makan, tidur dan kesehatan menurun. Variabel ini akan dikelompokkan menjadi 4 kategori yaitu : Sangat Setuju, Setuju, Tidak Setuju, Sangat Tidak Setuju.

**KUESIONER PENELITIAN
KLASIFIKASI TINGKAT STRES MAHASISWA AKHIR**

Nama :
 Jenis Kelamin :
 Jurusan :
 Semester :

Keterangan
 Berilah tanda centang (√) pada tabel dibawah untuk setiap pernyataan yang sesuai menurut pendapat anda.
 Sangat Setuju (SS), Setuju (S), Tidak Setuju (TS), Sangat Tidak Setuju (STS).

No.	Keterangan	SS	S	TS	STS
Interpersonal					
1.	Saya merasa tertekan oleh tuntutan orang tua yang menyuruh saya segera wisuda				
2.	Teman Saya selalu mengajak bermain game online sehingga saya tidak fokus mengerjakan tugas dan skripsi				
Intrapersonal					
1.	Saya mengalami kesulitan tidur dan sering begadang				
2.	Saya mengalami kesulitan keuangan dan keuangan terbatas				
Akademik					
1.	Saya kesulitan dalam memahami materi yang disampaikan oleh Dosen				
2.	Saya merasa kesulitan ketika harus mencari literatur dan referensi				
Lingkungan					
1.	Saya tidak bisa bermain dan istirahat dengan leluasa				
2.	Saya merasa tidak nyaman dirumah karena berisik membuat tidak fokus				

Dari pernyataan diatas, apakah anda merasa stres sebagai mahasiswa akhir ?

Tidak Stres
 Sangat Stres

**Mohon diisi dengan jujur dan dengan penuh rasa tanggung jawab.*

Gambar 2. Form Data Kuisisioner

Berikut ini merupakan data penelitian yang sudah diolah.

Tabel 1. Data penelitian yang sudah diolah

Responden	C1	C2	C3	C4	Hasil
R1	S	SS	S	SS	Sangat Stres
R2	S	S	S	TS	Sangat Stres
R3	S	S	SS	S	Sangat Stres
R4	TS	S	S	TS	Tidak Stres
R5	S	S	SS	TS	Sangat Stres
R6	TS	S	S	S	Sangat Stres
R7	TS	SS	S	TS	Tidak Stres
R8	S	SS	SS	S	Sangat Stres
R9	S	S	S	S	Sangat Stres
R10	TS	S	S	S	Sangat Stres
...
R110	S	S	SS	SS	Sangat Stres

Berikut ini merupakan penjelasan untuk mendapatkan hasil pada tabel tersebut:

a) Mengubah hasil skala linkert kuesioner ke dalam angka di Ms.excel

Dimana :

SS : 4

S : 3
 TS : 2
 STS : 1

- b) Jumlahkan 2 pernyataan dalam 1 variabel.
- c) Selanjutnya hasil dari penjumlahan 2 pernyataan dalam 1 variabel dibagi 2 pernyataan, kemudian didapatkan hasil rata-rata.
- d) Dari hasil rata-rata yang telah diperoleh, maka diubah kembali kedalam skala linkert.

HASIL DAN PEMBAHASAN

1. Proses Perhitungan Algoritma C4.5

Perhitungan Algoritma C4.5 untuk menentukan faktor dominan penyebab Stres dapat diuraikan sebagai berikut:

- Langkah 1: Menghitung jumlah kasus, jumlah kasus untuk keputusan Sangat Stres, jumlah kasus untuk keputusan Tidak Stres.
- Langkah 2: Menghitung *entropy* dari semua kasus dan kasus yang dibagi berdasarkan kelas atribut dengan persamaan (1) selanjutnya dilakukan perhitungan *gain* untuk masing-masing atribut dengan persamaan (2).

Berikut ini hasil perhitungan nilai *entropy* dan *gain* yang diuraikan pada tabel 2. berikut ini :

Tabel 2. Hasil Perhitungan Node 1

Node 1		Jlh Kasus	Sangat Stres	Tidak Stres	Entropy	Information Gain
Total		110	63	47	0,984684157	
C1						0,328180116
	SS	13	13	0	0	
	S	50	38	12	0,795040279	
	TS	35	12	23	0,927526588	
	STS	12	0	12	0	
C2						0,144499828
	SS	34	25	9	0,833764907	
	S	61	37	24	0,966985296	
	TS	13	1	12	0,391243564	
	STS	2	0	2	0	
C3						0,165554296
	SS	20	19	1	0,286396957	
	S	69	39	30	0,987692509	
	TS	20	5	15	0,811278124	
	STS	1	0	1	0	
C4						0,323788395
	SS	18	17	1	0,309543429	
	S	52	39	13	0,811278124	
	TS	34	7	27	0,733537929	
	STS	6	0	6	0	

- Langkah 3: Dari hasil perhitungan pada Tabel 4.1. diperoleh nilai atribut tertinggi adalah C1= Setuju, dengan *gain* sebesar 0,328180116. Maka atribut C1

dipilih sebagai *node* akar. Nilai kelas atribut Sangat Setuju adalah kosong, sehingga tidak perlu dilakukan perhitungan pada kelas atribut tersebut. Untuk kelas atribut Tidak Setuju dan Sangat Tidak Setuju belum diperoleh hasil antara keputusan Sangat Stres dan Tidak Stres, maka perlu dilakukan perhitungan lebih lanjut. Berikut ini hasil perhitungan dari kelas atribut C1= Setuju, yang ditunjukkan pada tabel 3 berikut.

Tabel 3. Hasil Perhitungan Node 1.1

<i>Node 1.1</i>		Jumlah Kasus	Sangat Stres	Tidak Stres	<i>Entropy</i>	<i>Information Gain</i>
C1-Setuju		50	38	12	0,795040279	
C2						0,196918444
	SS	12	11	1	0,41381685	
	S	34	27	7	0,733537929	
	TS	4	0	4	0	
	STS	0	0	2	0	
C3						0,237673766
	SS	9	9	0	0	
	S	37	29	8	0,753197991	
	TS	4	0	3	0	
	STS	0	0	1	0	
C4						0,269513404
	SS	10	10	0	0	
	S	28	24	4	0,591672779	
	TS	10	4	6	0,970950594	
	STS	2	0	2	0	

Dari hasil perhitungan pada Tabel 3, atribut yang menjadi *node* cabang dari C1 = Setuju adalah C4 dengan nilai *gain* tertinggi sebesar 0,269513404. Nilai kelas atribut Sangat Setuju dan Sangat Tidak Setuju adalah kosong, sehingga tidak perlu dilakukan perhitungan pada kelas atribut tersebut. Sedangkan nilai kelas atribut Setuju dan Tidak Setuju belum diperoleh hasilnya, maka akan dilakukan perhitungan lanjut. Berikut ini hasil perhitungan C1 = Setuju dan C4 = Setuju yang ditunjukkan pada Tabel 4 berikut ini:

Tabel 4. Hasil Perhitungan Node 1.1.1

<i>Node 1.1.1</i>		Jumlah Kasus	Sangat Stres	Tidak Stres	<i>Entropy</i>	<i>Information Gain</i>
C1-Setuju, C4-Setuju		28	24	4	0,591672779	
C2						0,256675926
	SS	6	6	0	0	
	S	20	18	2	0,468995594	
	TS	2	0	2	0	
	STS	0	0	0	0	
C3						0,121075746
	SS	9	9	0	0	
	S	24	22	2	0,41381685	
	TS	4	3	1	0,811278124	
	STS	0	0	1	0	

Dari hasil perhitungan pada Tabel 4, yang menjadi *node* cabang dari atribut C1 = Setuju dan C4 = Setuju adalah C2 dengan nilai *gain* tertinggi yaitu sebesar 0,256675926. Nilai

kelas Sangat Setuju, Tidak Setuju dan Sangat Tidak Setuju adalah kosong, sehingga tidak perlu dilakukan perhitungan pada kelas tersebut. Berikut ini hasil perhitungan C1 = Setuju, C4 = Setuju, C2 = Setuju dan C3 = Setuju yang ditunjukkan pada Tabel 5 berikut ini:

Tabel 5. Hasil Perhitungan Node 1.1.2

<i>Node 1.1.1</i>		Jumlah Kasus	Sangat Stres	Tidak Stres	<i>Entropy</i>	<i>Information Gain</i>
C1-Setuju, C4-Setuju, C2-Setuju		20	18	2	0,468995594	
C3						0,468995594
	SS	1	1	0	0	
	S	17	17	0	0	
	TS	2	0	2	0	
	STS	0	0	0	0	

Dari hasil perhitungan pada Tabel 5, hasil C1= Setuju, C4= Setuju, C2= Setuju dan C3 adalah kosong tidak perlu dilakukan perhitungan lebih lanjut. Dengan demikian *node* 1.1.2. adalah *node* cabang terakhir yang terbentuk.

Tabel 6. Hasil Perhitungan Node 1.2.

<i>Node 1.2</i>		Jumlah Kasus	Sangat Stres	Tidak Stres	<i>Entropy</i>	<i>Information Gain</i>
C1-Tidak Setuju		35	12	23	0,927526588	
C2						0,090834802
	SS	10	5	5	1	
	S	21	7	14	0,918295834	
	TS	4	0	4	0	
	STS	0	0	0	0	
C3						0,161463953
	SS	4	3	1	0,811278124	
	S	25	9	16	0,942683189	
	TS	6	0	6	0	
	STS	0	0	0	0	
C4						0,358715347
	SS	0	0	0	0	
	S	16	11	5	0,896038233	
	TS	18	1	17	0,309543429	
	STS	1	0	1	0	

Dari hasil perhitungan pada Tabel 6, atribut yang menjadi *node* cabang C1= Tidak Setuju adalah C4 dengan nilai *Gain* tertinggi sebesar 0,358715347. Nilai kelas Setuju dan Tidak Setuju diperoleh satu keputusan sehingga masih perlu dilakukan perhitungan lebih lanjut. Berikut ini hasil perhitungan C1= Tidak Setuju dan C4 = Setuju yang ditunjukkan pada Tabel 7 berikut ini:

Tabel 7. Hasil Perhitungan Node 1.2.1.

<i>Node 1.2.1</i>		Jumlah Kasus	Sangat Stres	Tidak Stres	<i>Entropy</i>	<i>Information Gain</i>
C1-Tdk Setuju, C4-Tdk Setuju		18	1	17	0,309543429	
C2						0,109007847
	SS	5	1	4	0,721928095	
	S	10	0	10	0	
	TS	3	0	3	0	
	STS	0	0	0	0	
C3						0,026978633
	SS	1	0	1	0	
	S	13	1	12	0,391243564	
	TS	4	0	4	0	
	STS	0	0	0	0	

Dari hasil perhitungan pada Tabel 7, yang menjadi *node* cabang dari atribut C1= Tidak Setuju dan C4= Tidak Setuju adalah C2 dengan nilai *gain* tertinggi yaitu sebesar 0,109007847. Nilai kelas Setuju, Tidak Setuju dan Sangat Tidak Setuju adalah kosong, sehingga tidak perlu dilakukan perhitungan pada kelas atribut tersebut. Untuk nilai kelas atribut Sangat setuju belum diperoleh satu keputusan sehingga masih perlu dilakukan perhitungan lebih lanjut. Berikut hasil perhitungan C1= Tidak Setuju, C4= Tidak Setuju dan C2= Sangat Setuju yang ditunjukkan pada Tabel 8 berikut ini:

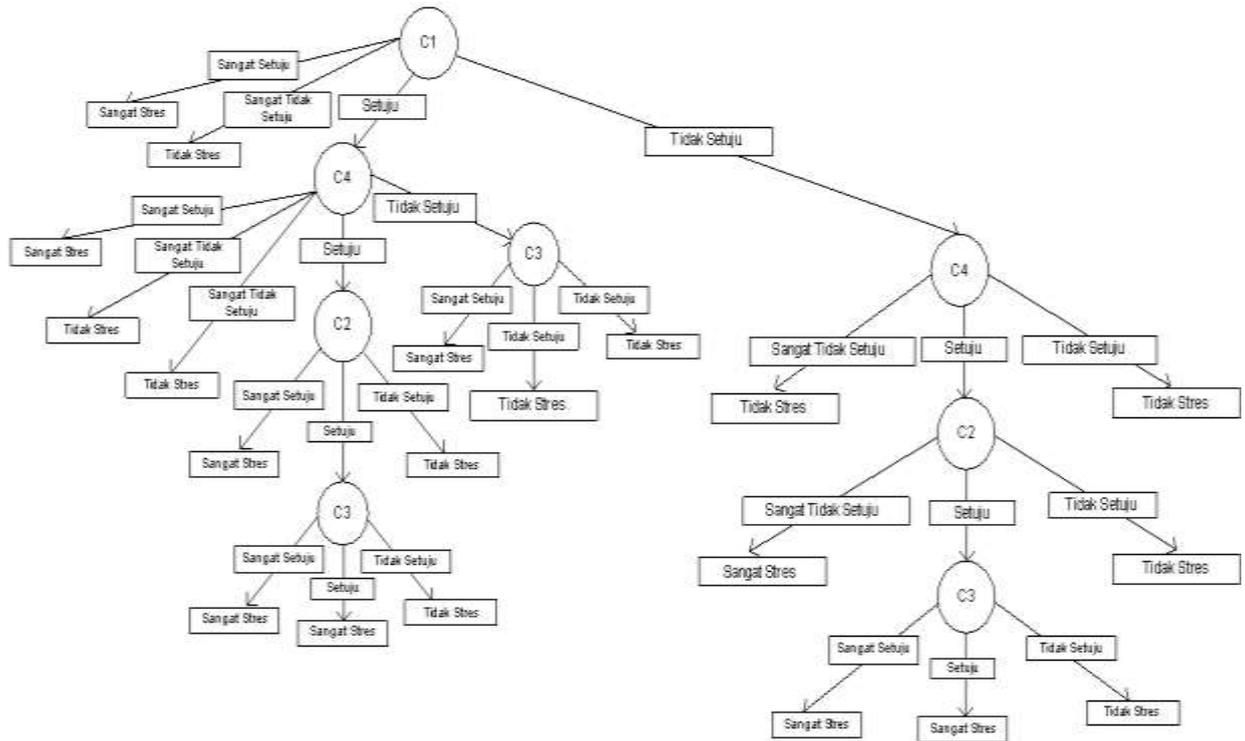
Tabel 8. Hasil Perhitungan Node 1.2.1.1

<i>Node 1.2.1.1</i>		Jumlah Kasus	Sangat Stres	Tidak Stres	<i>Entropy</i>	<i>Information Gain</i>
C1-Tdk Setuju, C4-Tdk Setuju, C2-Sangat Setuju		5	1	4	0,721928095	
C3						0
	SS	0	0	0	0	
	S	5	1	4	0,721928095	
	TS	0	0	0	0	
	STS	0	0	0	0	

Dari hasil perhitungan pada Tabel 8. diatas hasil dari C1= Tidak Setuju, C4 = Tidak Setuju dan C2= Sangat Setuju tidak perlu dilakukan perhitungan lebih lanjut. Dengan demikian *node* 1.2.1.1 adalah *node* cabang terakhir yang terbentuk.

2. Pohon Keputusan

Pohon keputusan pada Gambar 4.4. merupakan hasil akhir yang terbentuk dari hasil perhitungan *node* 1.1.1.1 dengan atribut Akademik (C3) sebagai *node* cabang, *node* 1.2.1.1 adalah atribut Akademik (C3) sebagai *node* cabang. Dari hasil tersebut diketahui semua kasus sudah masuk kedalam kelas. Dengan demikian dapat digambarkan hasil pohon keputusan hasil perhitungan menggunakan Algoritma C4.5 pada Gambar 3 berikut.



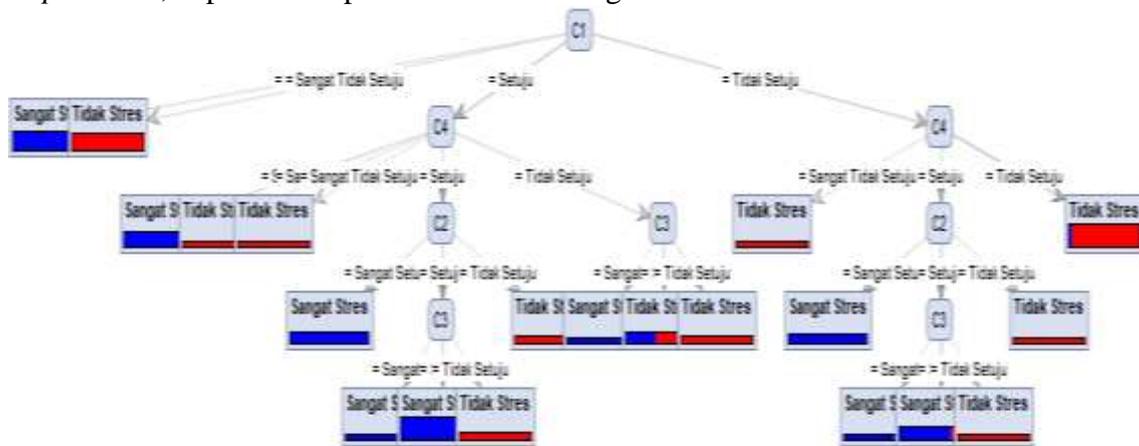
Gambar 3. Hasil Akhir Pohon Keputusan

Dari perhitungan diatas terdapat 20 *rules* yang dapat dijadikan sebagai pedoman dalam menentukan faktor dominan penyebab Stres Mahasiswa akhir. Adapun atau *rule* yang terbentuk berdasarkan pohon keputusan pada Gambar 3 yaitu 9 *rule* keputusan Sangat Stres dan 11 (Sebelas) *rule* keputusan Tidak Stres dijelaskan melalui teks narasi sebagai berikut :

1. Jika Interpersonal (C1) = Sangat Setuju, maka hasilnya Sangat Stres {Sangat Stres=13, Tidak Stres=0}
2. Jika Interpersonal (C1) = Sangat Tidak Setuju, maka hasilnya Tidak Stres {Sangat Stres=0, Tidak Stres=12}
3. Jika Interpersonal (C1) = Setuju, Lingkungan (C4) = Sangat Setuju, maka hasilnya Sangat Stres {Sangat Stres=10, Tidak Stres=0}
4. Jika Interpersonal (C1) = Setuju, Lingkungan (C4) = Sangat Tidak Setuju, maka hasilnya Tidak Stres {Sangat Stres=0, Tidak Stres=1}
5. Jika Interpersonal (C1) = Setuju, Lingkungan (C4) = Sangat Tidak Setuju, maka hasilnya Tidak Stres {Sangat Stres=0, Tidak Stres=1}
6. Jika Lingkungan (C4) = Setuju, Intrapersonal (C2) = Sangat Setuju, maka hasilnya Sangat Stres {Sangat Stres=6, Tidak Stres=0}
7. Jika Intrapersonal (C2) = Setuju, Akademik (C3) = Sangat Setuju, maka hasilnya Sangat Stres {Sangat Stres=1, Tidak Stres=0}
8. Jika Intrapersonal (C2) = Setuju, Akademik (C3) = Setuju, maka hasilnya Sangat Stres {Sangat Stres=17, Tidak Stres=0}
9. Jika Intrapersonal (C2) = Setuju, Akademik (C3) = Tidak Setuju, maka hasilnya Tidak Stres {Sangat Stres=0, Tidak Stres=2}
10. Jika Intrapersonal (C2) = Tidak Setuju, maka hasilnya Tidak Stres {Sangat Stres=0, Tidak Stres=2}

3. Pengujian

Hasil pengolahan data dengan model pohon keputusan sesuai dengan *Software Rapidminer*, dapat dilihat pada Gambar 4 sebagai berikut.



Gambar 4. Pohon Keputusan dengan Rapidminer

Berdasarkan Gambar 4 dapat dijelaskan bahwa Interpersonal adalah faktor dominan yang menyebabkan Stres pada Mahasiswa Akhir di STIKOM Tunas Bangsa Pematang Siantar. Penyederhanaan dari grafik diatas dapat dilihat pada *Rule Decision Tree* Gambar 5 berikut.

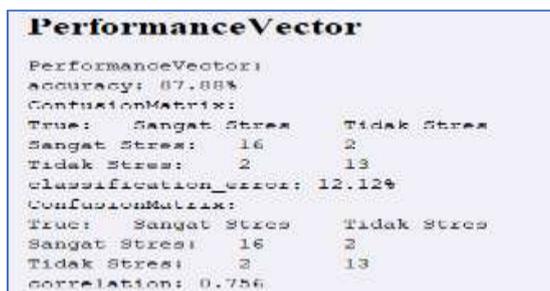
```

C1 = Sangat Setuju: Sangat Stres {Sangat Stres=13, Tidak Stres=0}
C1 = Sangat Tidak Setuju: Tidak Stres {Sangat Stres=0, Tidak Stres=12}
C1 = Setuju
| C4 = Sangat Setuju: Sangat Stres {Sangat Stres=10, Tidak Stres=0}
| C4 = Sangat Tidak Setuju: Tidak Stres {Sangat Stres=0, Tidak Stres=1}
| C4 = Sangat Tidak Setuju: Tidak Stres {Sangat Stres=0, Tidak Stres=1}
| C4 = Setuju
| | C2 = Sangat Setuju: Sangat Stres {Sangat Stres=6, Tidak Stres=0}
| | C2 = Setuju
| | | C3 = Sangat Setuju: Sangat Stres {Sangat Stres=1, Tidak Stres=0}
| | | C3 = Setuju: Sangat Stres {Sangat Stres=17, Tidak Stres=0}
| | | C3 = Tidak Setuju: Tidak Stres {Sangat Stres=0, Tidak Stres=2}
| | C2 = Tidak Setuju: Tidak Stres {Sangat Stres=0, Tidak Stres=2}
| C4 = Tidak Setuju
| | C3 = Sangat Setuju: Sangat Stres {Sangat Stres=1, Tidak Stres=0}
| | C3 = Setuju: Tidak Stres {Sangat Stres=3, Tidak Stres=4}
| | C3 = Tidak Setuju: Tidak Stres {Sangat Stres=0, Tidak Stres=2}
C1 = Tidak Setuju
| C4 = Sangat Tidak Setuju: Tidak Stres {Sangat Stres=0, Tidak Stres=1}
| C4 = Setuju
| | C2 = Sangat Setuju: Sangat Stres {Sangat Stres=4, Tidak Stres=0}
| | C2 = Setuju
| | | C3 = Sangat Setuju: Sangat Stres {Sangat Stres=1, Tidak Stres=0}
| | | C3 = Setuju: Sangat Stres {Sangat Stres=6, Tidak Stres=3}
| | | C3 = Tidak Setuju: Tidak Stres {Sangat Stres=0, Tidak Stres=1}
| | C2 = Tidak Setuju: Tidak Stres {Sangat Stres=0, Tidak Stres=1}
    
```

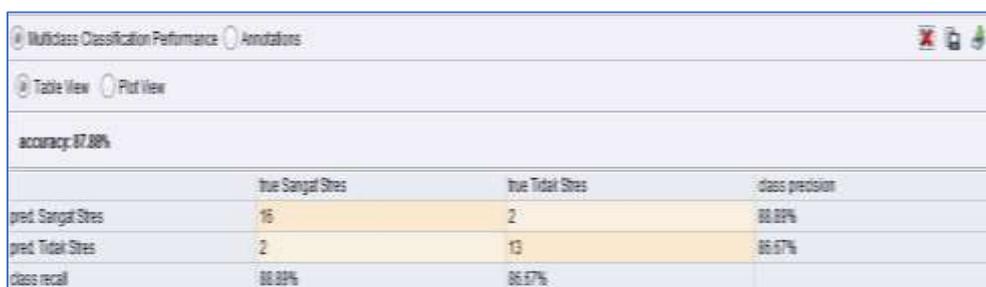
Gambar 5. Rules Decision Tree Pada RapidMiner

4. Validasi Decision Tree

Hasil penerapan Algoritma C4.5 menggunakan *software Rapidminer* dengan operator *X Validation* diperoleh nilai *Accuracy* yaitu sebesar 87.88% dan *classification_error* sebesar 12% artinya bahwa *rule* yang dihasilkan tingkat kebenarannya mendekati 100%, seperti yang terlihat pada Gambar 6.



Gambar 6. Performance Vector Algoritma C4.5



Gambar 7. Nilai Akurasi Algoritma C4.5

Berdasarkan gambar 7 dapat dijelaskan bahwa label Sangat Stres mempunyai nilai 16 dengan *Class Precision* 88,89% dan label Tidak Stres mempunyai nilai 2 dengan *Class Precision* sebesar 86,67%

KESIMPULAN DAN SARAN

Berdasarkan uraian pembahasan sebelumnya peneliti dapat memberikan beberapa kesimpulan bahwa permasalahan dalam menentukan tingkat stress pada Mahasiswa Akhir dapat diselesaikan menggunakan teknik *data mining*, yaitu dengan Algoritma C4.5. Menghasilkan 20 *rules* dan Tingkat akurasi yang dihasilkan oleh metode tersebut adalah 87.88%. Berdasarkan perhitungan menggunakan Algoritma C4.5 maka didapatkan faktor yang paling dominan adalah Interpersonal dengan nilai *gain* sebesar 0,328180116.

DAFTAR REFERENSI

Aditama, D. (2017). Hubungan Antara Spiritualitas dan Stres pada Mahasiswa yang Mengerjakan Skripsi. *Jurnal EL-Tarbawi*, 10(2), 39–62.

Ambarwati, P. D., Pinilih, S. S., & Astuti, R. T. (2019). Gambaran Tingkat Stres Mahasiswa. *Jurnal Keperawatan Jiwa*, 5(1), 40–47. <https://doi.org/10.26714/jkj.5.1.2017.40-47>

- Arminarahmah, N., GS, A. D., Bhawika, G. W., Dewi, M. P., & Wanto, A. (2021). Mapping the Spread of Covid-19 in Asia Using Data Mining X-Means Algorithms. *IOP Conf. Series: Materials Science and Engineering*, 1071(1), 012018. <https://doi.org/10.1088/1757-899X/1071/1/012018>
- Asmaul Husnah Nasrullah. (2018). penerapan metode C45 untuk klasifikasi mahasiswa drop out. *ILKOM Jurnal Ilmiah*, 10, 244–250.
- Bahri, S., Marisa Midyanti, D., Hidayati, R., Sistem Komputer, J., & Mipa, F. (2018). Perbandingan Algoritma Naive Bayes dan C4.5 Untuk Klasifikasi Penyakit Anak. *Seminar Nasional Aplikasi Teknologi Informasi (SNATi)*, 24–31.
- Fadli, F., & Butar, B. B. (2019). Penerapan Decision Tree Menggunakan Algoritma C4.5 Untuk Deteksi Demam Berdarah Pada RS. IMC Bintaro. *Indonesian Journal on Software Engineering (IJSE)*, 5(1), 75–86. <https://doi.org/10.31294/ijse.v5i1.5866>
- Febriani, S., & Sulistiani, H. (2021). ANALISIS DATA HASIL DIAGNOSA UNTUK KLASIFIKASI GANGGUAN. 2(4), 89–95.
- Febriyati, N. A., GS, A. D., & Wanto, A. (2020). GRDP Growth Rate Clustering in Surabaya City uses the K- Means Algorithm. *International Journal of Information System & Technology*, 3(2), 276–283.
- Fitriani, E., Aryanti, R., Saepudin, A., & Ardiansyah, D. (2020). Penerapan Algoritma C4.5 Untuk Klasifikasi Penempatan Tenaga Marketing. *Paradigma - Jurnal Komputer Dan Informatika*, 22(1), 72–78. <https://doi.org/10.31294/p.v22i1.6898>
- Gultom, W. T. C., Wanto, A., Gunawan, I., Lubis, M. R., & Kirana, I. O. (2021). Application of The Levenberg Marquardt Method In Predict The Amount of Criminality in Pematangsiantar City. *Journal of Computer Networks, Architecture, and High-Performance Computing*, 3(1), 21–29. <https://doi.org/10.47709/cnahpc.v3i1.926>
- Hanafiah, M. A., & Wanto, A. (2020). Implementation of Data Mining Algorithms for Grouping Poverty Lines by District/City in North Sumatra. *International Journal of Information System & Technology*, 3(2), 315–322.
- Hutagalung, J., Ginantra, N. L. W. S. R., Bhawika, G. W., Parwita, W. G. S., Wanto, A., & Panjaitan, P. D. (2021). COVID-19 Cases and Deaths in Southeast Asia Clustering using K-Means Algorithm. *Journal of Physics: Conference Series*, 1783(1), 012027. <https://doi.org/10.1088/1742-6596/1783/1/012027>
- Irnanda. (2020). Penerapan Klasifikasi C4.5 Dalam Meningkatkan Kecakapan Berbahasa Inggris dalam Masyarakat. *Seminar Nasional Teknologi Komputer & Sains (SAINTEKS)*, 304–308.
- Merry, & Mamahit, H. C. (2020). Stres Akademik Mahasiswa Aktif Angkatan 2018 dan 2019 Universitas Swasta di DKI Jakarta. *Jurnal Konseling Indonesia*, 6(1), 6–13.
- Muhamad, Windarto, A. P., & Suhada, S. (2019). Penerapan Algoritma C4.5 Pada Klasifikasi Potensi Siswa Drop Out. *KOMIK (Konferensi Nasional Teknologi Informasi Dan Komputer)*, 3(1), 753–760. <https://doi.org/10.30865/komik.v3i1.1688>
- Nurellisa, L., & Fitrihanah, D. (2020). Analisis Rekomendasi Calon Debitur Motor pada PT.XYZ menggunakan Algoritma C 4.5. *Jurnal Teknologi Informasi Dan Ilmu*

Komputer, 7(4), 673. <https://doi.org/10.25126/jtiik.2020742080>

- Oktavia, R., Hardinata, J. T., & Irawan, I. (2020). Penerapan Metode Algoritma K-means Dalam Pengelompokan Angka Harapan Hidup Saat Lahir Menurut Provinsi. *Kesatria: Jurnal Penerapan ...*, 1(4), 154–161.
- Phie, C., & Kasmara, Y. (2021). *Sistem Monitoring dan Deteksi Stres Pada Anak Berbasis Wearable*. 5(10), 943–949.
- Rahayu, I. A. S. (2018). *PENERAPAN METODE CHI-SQUARE AUTOMATIC INTERACTION DETECTION (CHAID) DAN ALGORITMA C4.5 PADA KLASIFIKASI STRES AKADEMIK MAHASISWA FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM*. 1–7.
- Sadewo, M. G., Windarto, A. P., Damanik, I. S., & Hartama, D. (2018). Penerapan C4.5 Untuk Memprediksi Kepuasan Pasien Terhadap Sistem Informasi Manajemen Rumah Sakit. *Ready Star*, 1(1), 21–30.
- Sagita, D. D., Fairuz, S. U. N., & Aisyah, S. (2021). Perbedaan Stres Akademik Mahasiswa Berdasarkan Jenis Kelamin. *Jurnal Konseling Andi Matappa*, 5(1), 9–16.
- Setiana, D. G. A. A., & Cristin Wiyani, R. E. (2017). Terhadap Stres Pada Lansia. *Jurnal Kebidanan Dan Keperawatan*, 13(2), 192–202.
- Sinaga, T. H., Wanto, A., Gunawan, I., Sumarno, S., & Nasution, Z. M. (2021). Implementation of Data Mining Using C4.5 Algorithm on Customer Satisfaction in Tirta Lihou PDAM. *Journal of Computer Networks, Architecture, and High-Performance Computing*, 3(1), 9–20. <https://doi.org/10.47709/cnahpc.v3i1.923>
- Sudarsono, B. G., & Lestari, S. P. (2020). Diagnosa Tingkat Depresi Mahasiswa Akhir Terhadap Penelitian Ilmiah Menggunakan Algoritma K-Nearest Neighbor. *Jurnal Media Informatika Budidarma*, 4(4), 1094–1099. <https://doi.org/10.30865/mib.v4i4.2448>
- Suhandiah, S., Ayuningtyas, A., & Sudarmaningtyas, P. (2021). Tugas Akhir dan Faktor Stres Mahasiswa. *JAS-PT (Jurnal Analisis Sistem Pendidikan Tinggi Indonesia)*, 5(1), 65–74. <https://doi.org/10.36339/jaspt.v5i1.424>
- Tapidingan, Y. C., & Paseru, D. (2020). Comparative Analysis of Classification Methods of KNN and Naïve Bayes to Determine Stress Level of Junior High School Students. *Indonesian Journal of Information Systems*, 2(2), 80–89. <https://doi.org/10.24002/ijis.v2i2.3035>
- Wanto, A., Siregar, M. N. H., Windarto, A. P., Hartama, D., Ginantra, N. L. W. S. R., Napitupulu, D., ... Prianto, C. (2020). *Data Mining : Algoritma dan Implementasi*. Yayasan Kita Menulis.