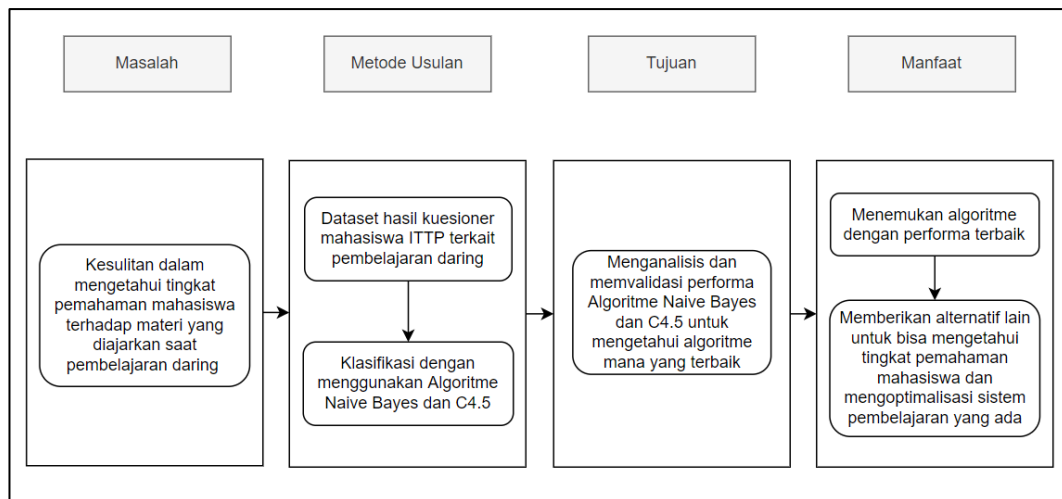


BAB III

METODE PENELITIAN

3.1 Kerangka Berpikir

Penelitian ini memiliki kerangka berpikir yang dimulai dengan penentuan masalah, metode yang diusulkan, tujuan, dan manfaat seperti pada Gambar 3.1. Permasalahan yang ada berupa kesulitannya seorang pengajar dalam memantau dan mengklasifikasikan mana mahasiswa yang sudah paham atau belum dengan materi yang diajarkan saat pembelajaran daring. Lalu, metode yang diusulkan berupa penggunaan Algoritme Naïve Bayes dan C4.5 yang berdasarkan penelitian sebelumnya bisa digunakan untuk klasifikasi. Kedua algoritme dibandingkan agar bisa dianalisis dan divalidasi performanya sehingga dapat diketahui mana algoritme terbaik yang bisa digunakan untuk kasus pada penelitian ini. Di samping itu, hal ini bisa menjadi alternatif lain untuk mempermudah pengajar dalam mengklasifikasi tingkat pemahaman mahasiswanya.



Gambar 3.1 Kerangka Berpikir

3.2 Subjek dan Objek Penelitian

Subjek penelitian yang diamati yaitu Algoritme Naïve Bayes dan C4.5 untuk klasifikasi. Kedua algoritme yang ada dibandingkan untuk bisa diketahui

mana yang memiliki akurasi terbaik terhadap objek yang diamati. Objeknya berfokus pada hasil survei mahasiswa Institut Teknologi Telkom Purwokerto terhadap pembelajaran daring yang pernah diikuti pada tahun 2021. Hasil survei tersebut dalam format file CSV, dengan atribut pertanyaan yang memiliki hubungan dengan pembelajaran daring. Label kelas yang diklasifikasi adalah kemampuan pemahaman belajar dengan nilai 1 (Ya) dan 0 (Tidak). Nilai 1 dapat diartikan responden bisa memahami materi yang diajarkan selama pembelajaran daring dan nilai 0 berarti belum bisa paham.

3.3 Teknik Pengambilan Data

Dataset merupakan kumpulan data yang dapat menunjang keberhasilan sebuah penelitian [49]. Pengambilan data pada penelitian ini dilakukan dengan cara penyebaran kuesioner kepada mahasiswa di Institut Teknologi Telkom Purwokerto yang pernah mengikuti pembelajaran daring pada tahun 2021. Penyebaran kuesioner dilakukan sejak 19 April 2022 sampai 21 Juni 2022 melalui platform Google Form dengan teknik *sampling: purposive sampling* yaitu peneliti menentukan kriteria sampelnya sendiri dengan disesuaikan pada tujuan penelitian [50]. Pengambilan sampel dilakukan dengan menggunakan rumus Slovin [51][52][53].

$$n = \frac{N}{1 + N(e)^2} \dots \dots \dots (3.1)$$

Keterangan:

n = Ukuran sampel

N = Jumlah populasi

e = Tingkat kesalahan tidak terwakilinya populasi (*Margin of error*)

Populasi mahasiswa Institut Teknologi Telkom Purwokerto sebanyak 4200 orang. Jika dihitung menggunakan rumus Slovin dengan nilai $e = 5\%$ atau 0,05 maka didapat ukuran sampel yang dibutuhkan minimal sebesar 365 orang untuk

bisa mewakili populasi. Semakin kecil nilai e maka ukuran sampel semakin akurat. Untuk perhitungan ukuran sampelnya sebagai berikut:

$$n = \frac{N}{1 + N(e)^2} = \frac{4200}{1 + 4200(0,05)^2} = \frac{4200}{11,5} = 365$$

Kemudian, pengumpulan dataset dilakukan dengan langkah awal menentukan instrumen dari kuesioner yang akan disebar. Adapun atribut yang ada pada kuesioner seperti yang tertera pada Tabel 3.1.

Tabel 3. 1 Keterangan Dataset

Indikator	Variabel	Nama Field	Bentuk Data
Data Responden	Nama	Nama	Teks
	NIM	NIM	Teks
	Program Studi	Program Studi	Kategorik
	Jenis Kelamin	Jenis Kelamin	Kategorik
Kondisi Kesehatan	P1	Pembelajaran daring tidak membuat mata saya terasa perih	Numerik
	P2	Pembelajaran daring tidak membuat saya sakit pinggang	Numerik
	P3	Pembelajaran daring tidak membuat saya sakit kepala atau pusing	Numerik
	P4	Pembelajaran daring tidak membuat saya merasa kesemutan atau kebas	Numerik
	P5	Pembelajaran daring tidak membuat saya merasa stres	Numerik
Motivasi	P6	Pembelajaran daring meningkatkan minat saya dalam belajar mandiri (asinkronus)	Numerik
	P7	Saya tertarik dan bersemangat mengikuti pembelajaran daring	Numerik
	P8	Pembelajaran daring meningkatkan rasa ingin tahu saya dalam mengeksplor materi perkuliahan	Numerik
Cara Pengajaran	P9	Pemberian tugas selama pembelajaran daring meningkatkan penguasaan terhadap materi	Numerik
	P10	Materi yang diajarkan jelas dan memiliki tampilan yang menarik	Numerik

Kuesioner ini terdiri dari 14 kolom yang mana 4 bagian pertama tentang data responden, 5 bagian tentang pertanyaan yang berhubungan dengan kondisi kesehatan, 3 bagian tentang motivasi, dan 2 bagian tentang cara pengajaran. Pada variabel nama dan NIM, nilainya disesuaikan dari inputan setiap individu. Untuk variabel Program Studi terdiri dari beberapa pilihan yaitu S1 Teknik Elektro, S1 Teknik Telekomunikasi, D3 Teknik Telekomunikasi, S1 *Software Engineering*, S1 Teknik Biomedis, S1 Teknik Informatika, S1 Sistem Informasi, S1 Sains Data, S1 Teknik Industri, S1 Desain Komunikasi Visual, S1 Teknik Logistik, S1 Bisnis Digital, dan S1 Desain Produk. Variabel Jenis Kelamin memiliki *value* laki-laki dan perempuan. P1 sampai P10 adalah variabel untuk mewakili setiap pertanyaan memiliki *value* berupa sangat tidak setuju, tidak setuju, setuju, dan sangat setuju yang secara otomatis oleh Google Form akan ditransformasi menjadi angka yaitu sangat tidak setuju = 1, tidak setuju = 2, setuju = 3, dan sangat setuju = 4.

Kuesioner yang sudah jadi disebarakan kepada beberapa responden sebagai sampel awal untuk pengujian validitas menggunakan *software* SPSS. Berdasarkan nilai r pada tabel *product moment* maka pada penelitian ini r tabel yang dipakai yaitu 0,361 dengan nilai signifikan 5% karena jumlah respondennya 30 orang. Untuk hasil uji validitas dapat dilihat pada Tabel 3.2.

Tabel 3. 2 Hasil Uji Validitas

Variabel	Item	r_{hitung}	r_{tabel}	Keterangan
Variabel P1-P10 Berpengaruh pada Tingkat Pemahaman Pembelajaran Daring	P1	0,823	0,361	Valid
	P2	0,834	0,361	Valid
	P3	0,896	0,361	Valid
	P4	0,826	0,361	Valid
	P5	0,847	0,361	Valid
	P6	0,764	0,361	Valid
	P7	0,539	0,361	Valid
	P8	0,814	0,361	Valid
	P9	0,854	0,361	Valid
	P10	0,812	0,361	Valid

Lalu, dilakukan penyebaran kuesioner lagi untuk bisa mendapatkan jumlah data sesuai dengan ukuran sampel. Dataset yang sudah terkumpul masih berupa data mentah sebanyak 419 *record data* seperti pada Tabel 3.3. Pada atribut P1 sampai P10 jawabannya berupa pilihan yaitu sangat tidak setuju = 1, tidak setuju = 2, setuju = 3, dan sangat setuju = 4.

Tabel 3. 3 Sampel Data Kuesioner

No	Nama	NIM	Program Studi	Jenis Kelamin	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10
1	Widi Afandi	19102127	S1 Teknik Informatika	Laki-laki	2	2	3	2	4	4	3	2	2	3
2	Firma mukarromah	19102210	S1 Teknik Informatika	Perempuan	2	2	1	3	1	3	3	3	2	3
3	Risa Riski Amalia	19102079	S1 Teknik Informatika	Perempuan	2	3	3	3	3	3	2	3	2	2
4	Amanda Aziz P. Ramadhani	20106052	S1 Teknik Industri	Perempuan	2	4	3	2	3	2	3	3	4	4
5	Annida Nur Islami	19102240	S1 Teknik Informatika	Perempuan	1	1	1	1	1	3	3	1	3	3
6	Rania Nur Hikmah	20104002	S1 Software Engineering	Perempuan	2	3	3	3	3	4	4	4	4	4
...
419	Darul Qomariyah	21109007	S1 Teknik Logistik	Perempuan	3	3	2	3	3	3	3	3	3	3

Kemudian, dengan data sebanyak 419 dilakukan uji reliabilitas menggunakan *software* SPSS. Berdasarkan pengujian reliabilitas ini didapatkan nilai Cronbach's Alpha yaitu 0,857 yang mana lebih besar dari 0,6 seperti pada Tabel 3.4. Oleh karena itu, kuesioner ini dinyatakan reliabel.

Tabel 3. 4 Hasil Uji Reliabilitas

Cronbach's Alpha	Jumlah Item Pertanyaan	Keterangan
0,857	10	Reliabel

3.4 Alat dan bahan penelitian

Alat:

1. *Laptop Acer Swift 3 Intel(R) Core(TM) i5-8250U CPU @ 1.60GHz (8 CPUs), ~1.8GHz Memory 8192MB RAM*
2. *Jupyter Notebook*
3. *Google Form, SPSS, Weka, RapidMiner*

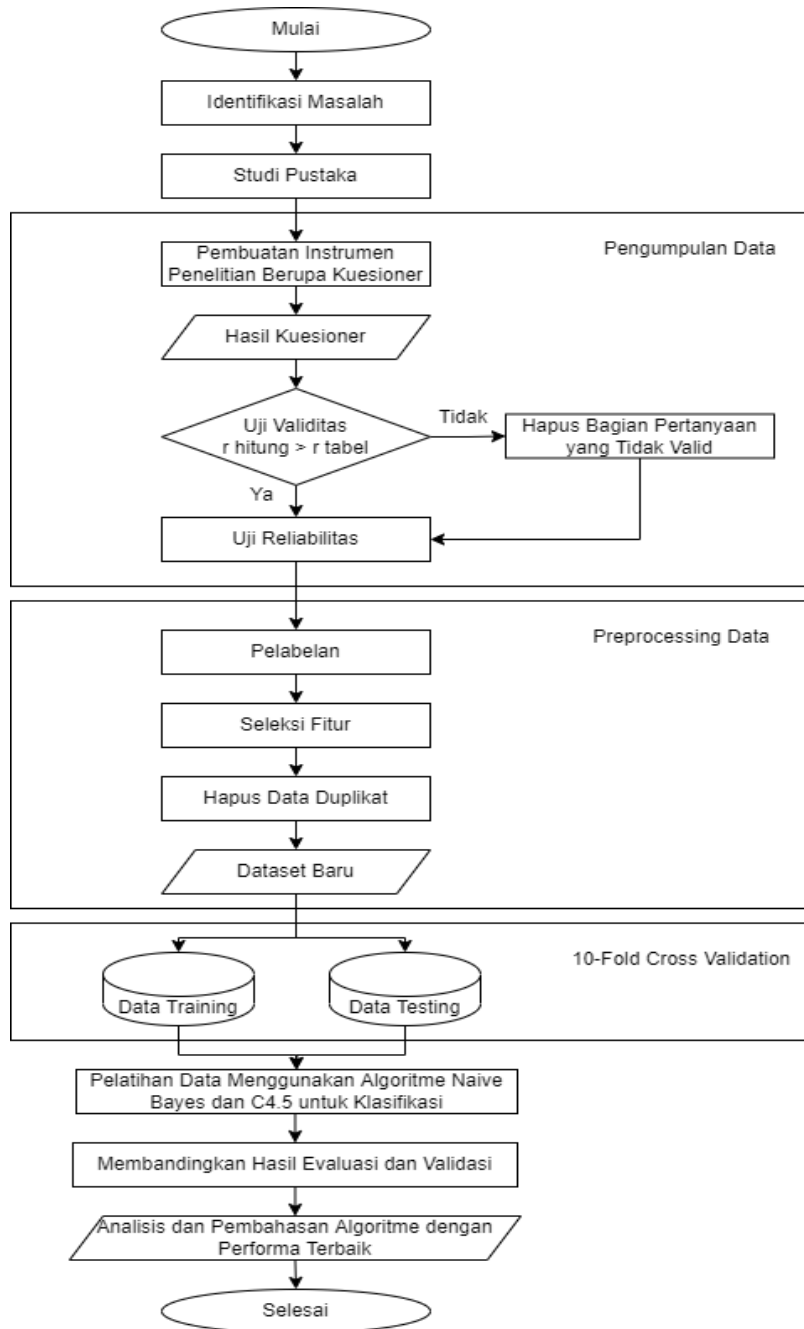
Bahan:

1. Dataset dari platform Google Form

3.5 Diagram Alir Penelitian

Tahapan penelitian ini diawali dengan menentukan topik kemudian perumusan masalah, tujuan, dan manfaat. Setelah itu, dilakukan studi literatur untuk mendapatkan acuan dalam melakukan penelitian. Kemudian, mengumpulkan dataset melalui penyebaran kuesioner. Diambil 30 data sebagai sampel awal untuk diuji validitas dan reliabilitas. Apabila saat uji validitas belum ada atribut pertanyaan yang tidak valid maka atribut tersebut dihapus lalu dilanjutkan dengan uji reliabilitas. Apabila sudah valid maka bisa langsung melakukan uji reliabilitas. Lalu, melakukan *preprocessing data* yang terdiri dari pelabelan, seleksi fitur, dan penghapusan duplikasi. Kemudian, terbentuklah data bersih. Data tersebut kemudian dibagi menjadi data latih dan data uji dengan percobaan pembagian data sebanyak 10 kali. Data-data itu diimplementasikan pada kedua algoritme yang dipakai sebagai bahan penelitian yaitu Naïve Bayes

dan C4.5. Hasil dari pelatihan dan pengujian dari proses klasifikasi kemudian dibandingkan untuk dievaluasi dan divalidasi. Selain itu, dilakukan analisis dan pembahasante rhadap hasilnya sehingga dapat diketahui algoritme dengan performa terbaik berdasarkan alasan yang tepat. Adapun alur tahapan penelitiannya seperti pada Gambar 3.2.



Gambar 3.2 Tahapan Penelitian

3.6 Preprocessing Data

3.6.1 Pelabelan

Menurut pakar pembelajaran daring yaitu Bapak Wahyu Andi Saputra, S. Pd, M. Eng bahwa setiap faktor yang dijadikan tolak ukur pada penelitian ini, memiliki nilai yang dapat mengklasifikasikan mahasiswa tergolong paham atau tidak. Apabila menggunakan skala Likert, semakin besar nilainya maka semakin mewakili faktor yang ada. P1 sampai P5 merupakan faktor dari segi kondisi kesehatan. Apabila kondisi kesehatannya baik maka akan lebih mudah dalam memahami pembelajaran yang diajarkan saat pembelajaran daring. Adapun untuk batas nilai seseorang dikatakan kondisi kesehatannya baik sehingga bisa memiliki pemahaman yang baik pula dalam menangkap materi sebagai berikut:

P1 = 1 (Sangat Tidak Setuju)

P2 = 2 (Tidak Setuju)

P3 = 4 (Sangat Setuju)

P4 = 4 (Sangat Setuju)

P5 = 1 (Sangat Tidak Setuju)

Kemudian, P6 sampai P8 adalah faktor motivasi diri. Semakin besar motivasi yang dimiliki maka peluang untuk bisa memahami materi dengan benar juga semakin besar. Berikut batas nilai seseorang dikatakan memiliki motivasi yang bagus sehingga bisa memiliki pemahaman yang baik dalam menangkap materi yang diajarkan sebagai berikut:

P6 = 2 (Tidak Setuju)

P7 = 3 (Setuju)

P8 = 3 (Setuju)

P9 dan P10 adalah faktor cara pengajaran. Apabila cara pengajaran dapat menarik perhatian mahasiswa untuk memperhatikan dan melaksanakan apa yang diajarkan maka peluang untuk mahasiswa dapat memahami materi juga semakin besar. Berikut batasan nilai jika seseorang dianggap menikmati cara pengajaran yang diberikan pengajar:

P9 = 4 (Sangat Setuju)

P10 = 4 (Sangat Setuju)

Nilai-nilai yang telah disebutkan di atas juga merupakan batasan untuk seseorang dikatakan belum bisa memahami materi dengan baik. Semakin kecil nilainya maka semakin kurang pemahaman yang bisa dimiliki. Apabila semua nilai itu ditotal maka didapatkan nilai 28. Nilai ini bisa dijadikan patokan paham tidaknya mahasiswa terhadap materi yang diajarkan. Batas nilai maksimal yaitu 40 dan minimal yaitu 10. Apabila lebih dari sama dengan 28 maka bisa dikatakan sudah mampu memahami materi yang diajarkan saat pembelajaran daring. Jika kurang dari 28 maka digolongkan belum bisa memahami materi pembelajaran daring. Untuk menampung dua kategori tersebut (Ya dan Tidak) maka dibuat kelas label yang diberi nama 'Kemampuan Memahami Materi' setelah kolom P10.

3.6.2 Seleksi Fitur

Seleksi Fitur pada penelitian ini menggunakan *software* WEKA dengan *Attribute Evaluator* berupa *Correlation Attribute Eval* [54]. Filter *Correlation Attribute Eval* digunakan untuk mengevaluasi korelasi antara kelas dan atribut lainnya. Langkah ini adalah penting karena dapat menemukan atribut yang paling dekat hubungannya yang mempengaruhi kelas dan mengabaikan atribut yang kurang terkait dari model. *Correlation Attribute Eval* mengevaluasi nilai atribut dengan mengukur korelasi antara atribut dan kelas yang dijadikan label. Semua atribut yang ada dihitung nilai korelasinya dan didapatkan nilai seperti pada Gambar 3.3. Lalu, atribut yang ada diurutkan nilai korelasinya dari yang terbesar hingga terkecil. Apabila nilai korelasinya kurang dari sama dengan nol maka atribut tersebut tidak memiliki korelasi dengan kelas dan jika lebih dari nol maka memiliki korelasi. Semakin besar nilai korelasi maka semakin besar pengaruhnya terhadap kelas. Berdasarkan data yang dihasilkan dan informasi dari pakar tentang pembelajaran daring yaitu Bapak Wahyu Andi Saputra, S. Pd, M. Eng bahwa atribut Nama, Program Studi, Jenis Kelamin dan NIM kurang berpengaruh pada pemahaman belajar mahasiswa sehingga keempat atribut tersebut sebaiknya dihapus. Jadi, atribut yang akan terus digunakan untuk penelitian selanjutnya yaitu P1,P2,P3,P4,P5,P6,P7,P8,P9, dan P10.

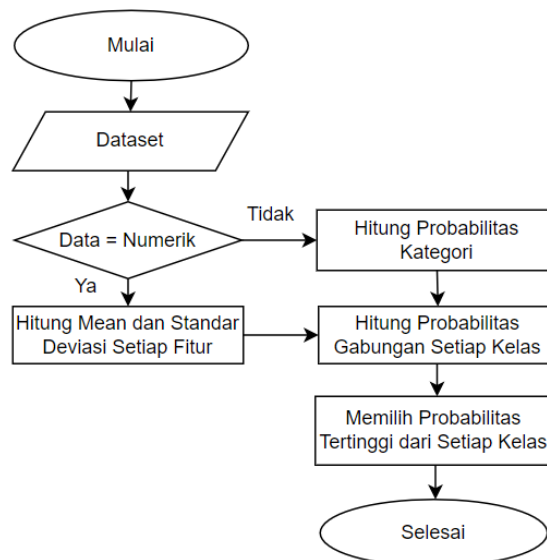
3.7 10-Fold Cross Validation

Penelitian ini menggunakan metode *Cross Validation* dengan nilai $k=10$ untuk membagi data menjadi data latih dan data uji.

3.8 Alur Algoritme

3.8.1 Naïve Bayes

Naïve Bayes merupakan salah satu algoritme dengan metode probabilitas yang mempunyai beberapa tahapan dalam proses eksekusinya. Alur algoritme Naïve Bayes dapat dilihat pada Gambar 3.4.

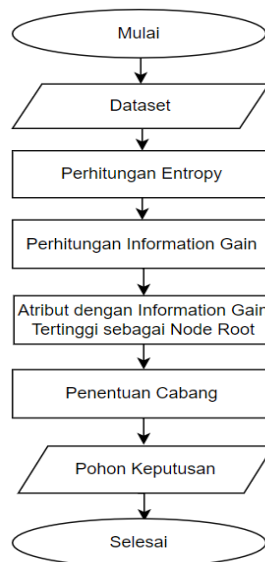


Gambar 3.4 Diagram Alur Algoritme Naive Bayes

Ketika dataset sudah dikumpulkan maka data tersebut dicek terlebih dahulu tipe datanya termasuk numerik atau kategori. Apabila datanya numerik maka atributnya dihitung dulu *mean* dan standar deviasinya. Jika kategori maka dihitung probabilitasnya, termasuk atribut yang menjadi kelasnya. Kemudian, melakukan perhitungan probabilitas gabungan dari setiap atribut. Hasil dari probabilitas gabungan dimasukkan ke dalam rumus perhitungan probabilitas akhir begitu juga hasil dari perhitungan probabilitas kelasnya. Hasil probabilitas akhir dari setiap kelas kemudian dibandingkan untuk diketahui mana yang memiliki nilai probabilitas tertinggi untuk dijadikan jawabannya.

3.8.2 C4.5

C4.5 adalah salah satu algoritme *decision tree* yang mempunyai beberapa tahapan dalam proses eksekusinya. Alur algoritme C4.5 dapat dilihat pada Gambar 3.5.



Gambar 3.5 Diagram Alur Algoritme C4.5

Ketika datasetnya sudah terkumpul maka dengan data tersebut dilakukan perhitungan *entropy* total dan *entropy* masing-masing jawaban dari setiap atribut. Lalu, dilakukan perhitungan *information gain* dari masing-masing atribut. Atribut yang memiliki *information gain* tertinggi maka akan dijadikan *node root*. Lalu, dilakukan perhitungan ulang dengan cara yang sama untuk menentukan cabang hingga terbentuklah pohon keputusan.

3.8 Analisis Data

Penelitian ini dilakukan dengan analisis data untuk mendapat hipotesis awal. Perbandingan performa Algoritme Naïve Bayes dan C4.5 dilakukan untuk mengklasifikasi tingkat pemahaman mahasiswa terhadap pembelajaran daring berdasarkan beberapa faktor yang dapat mempengaruhinya yaitu kondisi kesehatan, motivasi diri, dan cara pengajaran yang didapat. Hal ini merupakan percobaan awal untuk mengukur performa yang meliputi akurasi, *precision*, *recall*, RMSE, dan waktu komputasi. Percobaan ini dilakukan dengan menggunakan *software* RapidMiner. Untuk perhitungan waktu komputasi

dilakukan secara *real time* menggunakan *stopwatch*. Adapun atribut yang digunakan sebagai kelas labelnya yaitu Kemampuan Memahami Materi dengan kategori ‘Ya’ dan ‘Tidak’. Lalu, pembagian data latih dan data uji dilakukan dengan *k-fold cross validation* dengan $k=10$.

Algoritme Naïve Bayes dihitung nilai probabilitas dari kategori pada kelas yang ada. Adapun hasil dari percobaan awal ini seperti Tabel 3.6. Didapatkan mahasiswa yang data sebenarnya tidak paham berhasil diprediksi tidak paham juga dinyatakan sebagai *True Positive* (TP) sebanyak 273. Lalu, mahasiswa yang data sebenarnya ya paham tetapi diprediksi tidak paham dinyatakan sebagai *False Positive* (FP) sebanyak 2. Mahasiswa yang data sebenarnya tidak paham tetapi diprediksi ya paham dinyatakan sebagai *False Negative* (FN) sebanyak 7. Mahasiswa yang data sebenarnya termasuk ya paham dan berhasil diprediksi ya paham juga dinyatakan sebagai *True Negative* (TN) sebanyak 118.

Tabel 3. 6 Hasil Klasifikasi dengan Algoritme Naive Bayes

	True Tidak	True Ya	Class Precision
Pred. Tidak	273	2	99,27%
Pred. Ya	7	118	94,40%
Class Recall	97,50%	98,33%	
Akurasi	97,75%		
RMSE	0,132		
Waktu Komputasi	0,53 detik		

Algoritme C4.5 menggunakan *information gain* sebagai pembentuk *node* dengan *maximal depth* (*max depth*) terbaik adalah 8 seperti pada Tabel 3.7. Hasil performa dari percobaan awal ini dapat dilihat pada Tabel 3.8. Didapatkan mahasiswa yang data sebenarnya tidak paham berhasil diprediksi tidak paham juga dinyatakan sebagai *True Positive* (TP) sebanyak 262. Lalu, mahasiswa yang data sebenarnya ya paham tetapi diprediksi tidak paham dinyatakan sebagai *False Positive* (FP) sebanyak 21. Mahasiswa yang data sebenarnya tidak paham tapi diprediksi ya paham dinyatakan sebagai *False Negative* (FN) sebanyak 18.

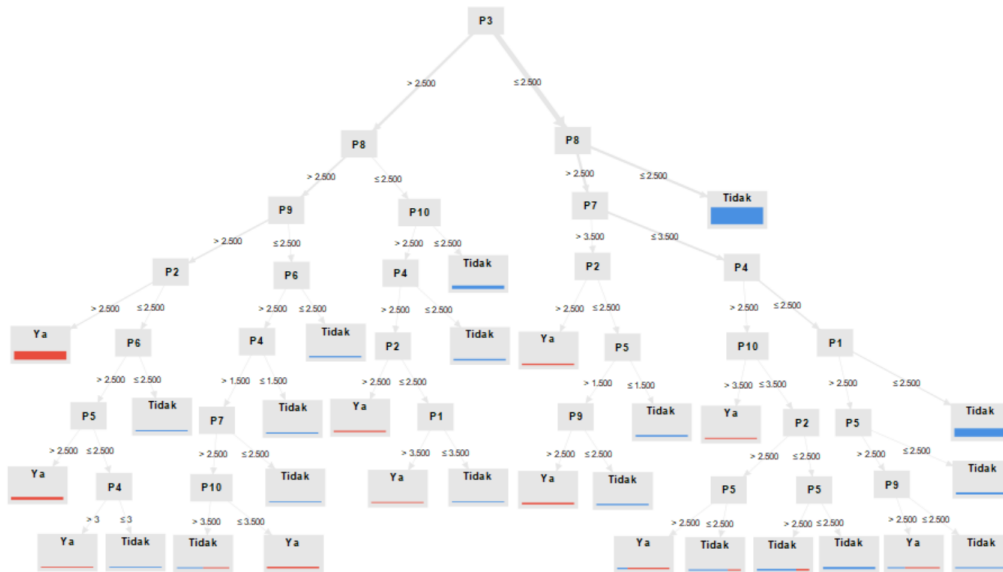
Mahasiswa yang data sebenarnya termasuk ya paham dan berhasil diprediksi ya paham juga dinyatakan sebagai *True Negative* (TN) sebanyak 99. Adapun pohon keputusan yang terbentuk dapat dilihat pada Gambar 3.6.

Tabel 3. 7 Percobaan Berdasarkan *Max Depth*

Percobaan Ke-	<i>Max Depth</i>	Akurasi
1	1	70,00%
2	2	83,00%
3	3	88,25%
4	4	87,75%
5	5	88,75%
6	6	88,75%
7	7	89,75%
8	8	90,25%
9	9	90,25%
10	10	90,00%

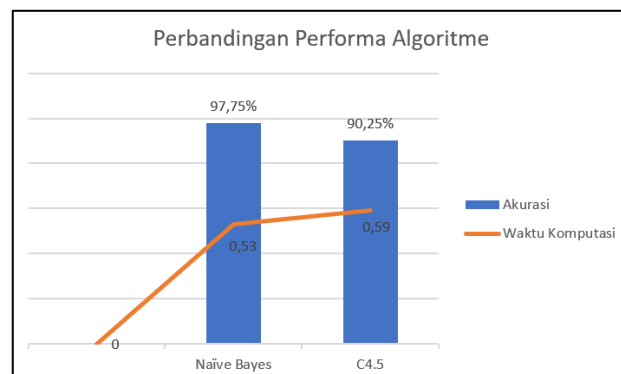
Tabel 3. 8 Hasil Klasifikasi dengan Algoritme C4.5

	True Tidak	True Ya	<i>Class Precision</i>
Pred. Tidak	262	21	92,58%
Pred. Ya	18	99	84,62%
<i>Class Recall</i>	93,57%	82,50%	
Akurasi	90,25%		
RMSE	0,290		
Waktu Komputasi	0,59 detik		

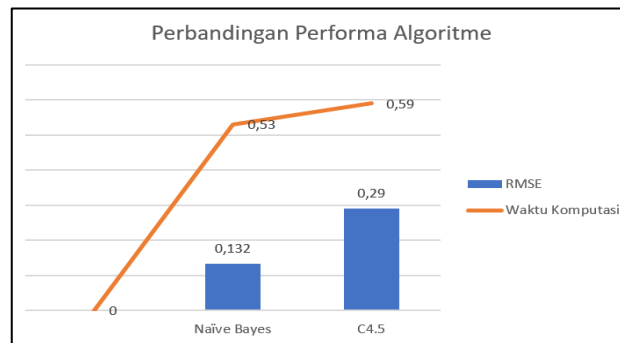


Gambar 3.6 Pohon Keputusan yang Terbentuk

Visualisasi perbandingan performa algoritmenya dapat dilihat pada Gambar 3.7 yang merupakan perbandingan dari segi akurasi dan waktu komputasi dan Gambar 3.8 yang merupakan perbandingan dari nilai RMSE dan waktu komputasi. Berdasarkan Gambar 3.7 Algoritma Naïve Bayes lebih unggul, dimana waktu komputasinya lebih rendah sehingga menghasilkan akurasi yang lebih tinggi. Gambar 3.8 menunjukkan bahwa saat nilai RMSE lebih rendah yaitu pada Algoritma Naïve Bayes, maka waktu komputasi yang dibutuhkan akan rendah juga karena *error* yang dihasilkan juga kecil.



Gambar 3.7 Perbandingan Algoritme Berdasarkan Akurasi dan Waktu Komputasi



Gambar 3.8 Perbandingan Algoritme Berdasarkan RMSE dan Waktu Komputasi

Tabel 3. 9 Perbandingan Performa Algoritme Naïve Bayes dan C4.5

Parameter	Naïve Bayes	C4.5
<i>Recall</i> Kelas Tidak	99,27%	93,57%
<i>Recall</i> Kelas Ya	94,40%	82,50%
<i>Precision</i> Prediksi Tidak	97,50%	92,58%
<i>Precision</i> Prediksi Ya	98,33%	84,62%
Akurasi	97,75%	90,25%
RMSE	0,132	0,290
Waktu Komputasi	0,53 detik	0,59 detik

Jika dibandingkan hasil performa Algoritme Naïve Bayes dan C4.5 maka didapatkan performa Algoritme Naïve Bayes lebih baik dari C4.5 seperti Tabel 3.9. Hal ini dikarenakan jumlah data yang ada termasuk jumlah data kecil [24] yang mana Naïve Bayes yang sifatnya sederhana sangat cocok untuk jumlah data seperti ini. Selain itu, tidak ada *missing value* dan semua data memiliki nilai sehingga nilai probabilitas tidak ada yang bernilai nol. Hal ini memungkinkan nilai akurasi menjadi lebih tepat dan nilai *error* rendah. Perhitungan waktu komputasi pada percobaan ini juga sedikit lebih cepat daripada C4.5 karena eksekusinya langsung menghitung keseluruhan atribut dan kriteria terhadap kelas label secara sederhana [13]. Sementara itu, C4.5 menghasilkan akurasi yang bagus juga namun tidak sebesar Naïve Bayes. Hal ini dikarenakan jumlah data yang tergolong kecil membuat cabang dari pohon keputusan yang terbentuk tidak bisa mengklasifikasikan secara spesifik. Ada anomali cabang yang terbentuk. Selain

itu, terjadi pengakumulasian jumlah *error* dari setiap tingkat dalam sebuah pohon keputusan yang menyebabkan akurasi menjadi rendah dan nilai *error* lebih banyak [16]. Percobaan ini juga menghasilkan waktu komputasi yang sedikit lebih lambat pada C4.5 daripada Naïve Bayes karena datasetnya dipanggil berulang untuk membuat pohon keputusannya. Oleh karena itu, pada penelitian ini Algoritme terbaik yang digunakan terhadap dataset yang ada yaitu Algoritme Naïve Bayes.

Adapun terkait hasil klasifikasinya didapatkan masih banyak mahasiswa yang tidak paham terhadap materi yang diajarkan melalui pembelajaran daring. Oleh sebab itu, diperlukan optimalisasi terhadap sistem pembelajaran yang diterapkan dengan memperhatikan aspek kesehatan, motivasi, dan cara pengajaran yang baik untuk mahasiswa.