

RESEARCH ARTICLE

Comparison of Online and Offline Learning During The COVID-19 Pandemic using Naïve Bayes Method and C4.5

(Perbandingan Pembelajaran Online dan Offline pada Masa Pandemi Covid-19 Menggunakan Metode Naive Bayes 4.5)

Andini Cahya Aulia^{*)}, Mohamat Fatekurohman, I Made Tirta

Jurusan Matematika, Fakultas MIPA, Universitas Jember, Jl. Kalimantan 37, Jember 68121 Indonesia

ABSTRACT

Learning is a process of interaction between educators and students who meet the elements of learning carried out in an educational environment, so that learning can develop student's abilities, interests and talents optimally. In today's era learning is done online and inversely with offline. The purpose of this study is to analyze the comparison of percentages and classification results as well as the results of learning evaluations using the Naïve Bayes method and C4.5. This test is carried out with 4 variables and a comparison of the two methods. The results showed that the accuracy of Naïve Bayes was 74.07% and C4.5. of 77.77% so that the comparison results show that the level of accuracy of the C4.5 method is better than Naïve Bayes. The resulting importance variables are time and effectiveness as well as the results of the classification of learning decisions, namely the offline category as many as 16 data on the Naïve Bayes method and 19 data on the Decision Tree algorithm C4.5 method from 27 input testing data.

Pembelajaran adalah proses interaksi antara pendidik dengan peserta didik yang memenuhi unsur-unsur pembelajaran yang dilakukan dalam suatu lingkungan pendidikan, sehingga pembelajaran dapat mengembangkan kemampuan minat dan bakat siswa secara optimal. Pada era zaman sekarang pembelajaran dilakukan secara *online* dan berbanding terbalik dengan *offline*. Tujuan dari penelitian ini menganalisa perbandingan presentase dan hasil klasifikasi serta hasil evaluasi pembelajaran menggunakan metode *Naïve Bayes* dan C4.5. Pengujian ini dilakukan dengan 4 variabel dan perbandingan kedua metode. Hasil penelitian didapatkan bahwa nilai akurasi *Naïve Bayes* sebesar 74,07% dan C4.5. sebesar 77,77% sehingga hasil perbandingan didapatkan bahwa tingkat akurasi metode C4.5 lebih baik daripada *Naïve Bayes*. Variabel *importance* yang dihasilkan yaitu waktu dan keefektifan serta hasil klasifikasi keputusan pembelajaran yaitu kategori *offline* ukuran data 16 pada metode *Naïve Bayes* dan sebanyak 19 data pada metode *Decision Tree* algoritma C4.5 dari 27 *input data testing*.

Keywords: Classification, Comparison, Naïve Bayes, C4.5.

^{*)}Corresponding author:

Andini Cahya Aulia

E-mail: auliaandini447@gmail.com

PENDAHULUAN

Perkembangan kondisi pembelajaran sekarang yaitu salah satunya di negara Indonesia yang sedang terdampak kasus COVID-19 mengakibatkan tindakan pembelajaran dilakukan secara *online* dan berbanding terbalik dengan *offline*. Metode pembelajaran *online* dan *offline* mempunyai kelebihan serta kekurangan masing-masing, maka di akhir pembelajaran selalu dilakukan evaluasi mengenai metode pembelajaran yang sedang dipakai. Permasalahan yang dialami yaitu hasil evaluasi dan keputusan pembelajaran *online* dan *offline*.

Hasil evaluasi dan keputusan dapat diselesaikan dengan menggunakan perbandingan teknik data

mining dengan algoritma Naïve Bayes dan algoritma C4.5. Menurut [1], membahas tentang prediksi hasil tingkat pemahaman mahasiswa terhadap mata kuliah berdasarkan posisi duduk menggunakan metode *Naïve Bayes*. Nilai dari segi akurasi sebesar 88,24% maka dengan nilai akurasi yang melebihi minimal dapat dikatakan metode tersebut berhasil digunakan untuk memprediksi hasil dari pemahaman materi kuliah berdasarkan posisi duduk, sedangkan menurut [2], tentang memprediksi kepuasan mahasiswa terhadap kinerja dosen. Hasil pengujian memiliki tingkat nilai akurasi sebesar 94,62% maka metode tersebut berhasil digunakan untuk memprediksi kepuasan mahasiswa terhadap kinerja dosen. Metode *Naïve Bayes*

merupakan merupakan pendekatan statistik untuk melakukan inferensi induksi pada persoalan klasifikasi [3]. Metode ini menggunakan probabilitas bersyarat sebagai dasarnya. Metode *Decision tree* (C4.5) merupakan salah satu metode klasifikasi yang menggunakan representasi struktur pohon. Konsep dari pohon keputusan *Decision tree* ini kemudian akan dihasilkan *rule-rule* solusi permasalahan [4]. Metode *Naïve Bayes* dan *Decision tree* algoritma C4.5 sama-sama merupakan metode data *mining*. Kedua metode sangat sering digunakan untuk penelitian-penelitian terutama pada penelitian yang membahas tentang klasifikasi.

Pada paper ini menggunakan data *mining* dengan metode *Naïve Bayes* dan *Decision tree* (C4.5). Data yang digunakan yaitu data kuesioner dari pendapat mahasiswa angkatan 2018 dan 2019. Penelitian ini dilakukan dengan tujuan membandingkan hasil kedua metode tersebut dan memprediksi hasil keputusan pembelajaran berdasarkan pilihan mahasiswa yaitu pembelajaran *online* atau *offline*.

a. Data Mining

Data *mining* merupakan suatu proses penggalian data atau penyaringan data dengan memanfaatkan kumpulan data dengan ukuran yang cukup besar melalui serangkaian proses untuk mendapatkan informasi dari data tersebut. Pengumpulan data *mining* bukan hanya sebagai tempat terkumpul saja, tetapi mencakup analisis dan prediksi dari informasi yang ditampilkan. Data yang dikumpulkan disimpan dalam *database*, kemudian diproses sehingga dapat dijadikan untuk pengambilan keputusan informasi yang akan digunakan [3]. Salah satu jenis data *mining* yang digunakan yaitu klasifikasi (*classification*) yaitu melakukan generalisasi struktur yang diketahui untuk diaplikasikan pada data-data baru.

b. Naïve Bayes

Dasar metode *Naïve Bayes* yaitu pada asumsi penyederhanaan bahwa nilai atribut secara kondisional saling bebas jika diberikan nilai *output*, maka nilai *output* serta probabilitas mengamati secara bersama merupakan produk dari probabilitas individu. Kelebihan menggunakan metode *Naïve Bayes* adalah membutuhkan jumlah data *testing* yang kecil untuk menentukan estimasi parameter yang diperlukan dalam proses pengklasifikasian. *Naïve Bayes* dapat bekerja jauh lebih baik dalam suatu dunia nyata yang kompleks daripada yang diharapkan. Klasifikasi *Naïve*

Bayes diasumsikan bahwa ada atau tidak ciri tertentu dari sebuah kelas tidak berhubungan dengan ciri dari kelas lain. Persamaan dari teorema *Naïve Bayes* adalah:

$$P(H|X) = \frac{P(X|H).P(H)}{P(X)} \quad (1)$$

Keterangan:

X : Data dengan kelas yang belum diketahui

H : Hipotesis data X adalah suatu kelas spesifik

$P(H|X)$: Probabilitas hipotesis data H berdasarkan kondisi X

$P(H)$: Probabilitas hipotesis H

$P(X|H)$: Probabilitas X berdasarkan kondisi pada hipotesis H

$P(X)$: Probabilitas X

Keterangan diatas untuk menjelaskan teorema *Naïve Bayes*, dapat diketahui bahwa proses klasifikasi perlu sejumlah petunjuk untuk menentukan kelas yang cocok bagi sampel yang dianalisis. Teorema Bayes diatas disesuaikan seperti persamaan berikut:

$$P(C|F_1 \dots F_n) = \frac{P(C)P(F_1 \dots F_n|C)}{P(F_1 \dots F_n)} \quad (2)$$

c. Decision Tree (C4.5)

Decision tree dengan algoritma C4.5 merupakan salah satu metode klasifikasi yang menggunakan representasi struktur pohon (*tree*) dimana setiap *node* mempresentasikan atribut, cabangnya mempresentasikan nilai dari atribut dan daun mempresentasikan kelas. Algoritma C4.5 merupakan algoritma yang digunakan untuk membentuk pohon keputusan. Pohon keputusan merupakan pemetaan mengenai alternatif-alternatif pemecahan masalah yang dapat diambil dari masalah tersebut. Pohon tersebut juga memperlihatkan faktor-faktor kemungkinan yang akan mempengaruhi alternatif keputusan tersebut dan disertai dengan estimasi hasil akhir yang akan didapat [4]. Algoritma C4.5 ini mempunyai input berupa *training samples* dan *testing samples*. Pembuatan pohon keputusan untuk memilih atribut sebagai akar dapat menggunakan persamaan berikut:

$$Entropy(S) = \sum_{i=1}^n -p_i * \log_2 p_i \quad (3)$$

Dapat disederhanakan:

$$Entropy(S) = -p + \log_2 p - p - \log_2 p \quad (4)$$

Menghitung nilai *Gain*:

$$Gain(S, A) = Entropy(S) - \sum_{i=1}^s \frac{|S_i|}{|S|} \quad (5)$$

Keterangan:

S : himpunan kasus

A : Atribut

n : jumlah partisi atribut

|*S_i*| : jumlah kasus pada partisi ke - *i*

|*S*| : jumlah kasus dalam *S*

d. Confusion Matrix

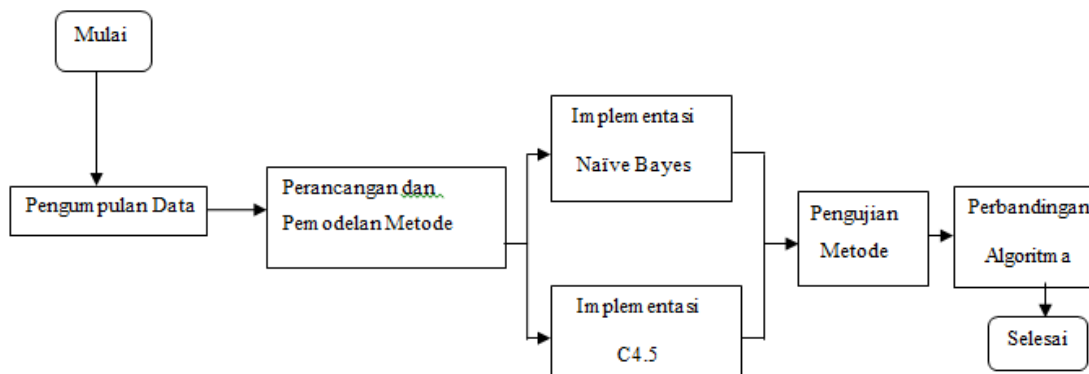
Confusion Matrix adalah suatu metode yang digunakan untuk perhitungan nilai akurasi pada konsep data mining. *Confusion Matrix* dapat digambarkan pada tabel yang menyatakan jumlah semua data uji yang benar dan salah diklasifikasikan. Nilai yang dihasilkan melalui metode *Confusion Matrix* yaitu berupa evaluasi *accuracy*, *precision*, *Recall* dan *specificity*.

METODE PENELITIAN

Proses penelitian ini dimulai dari pengumpulan data, analisis data, pemecahan masalah hingga penarikan kesimpulan. Penelitian ini dilakukan untuk membandingkan nilai tingkat akurasi antara metode *Naïve Bayes* dan metode *Decision tree* algoritma C4.5 serta memprediksi hasil keputusan pembelajaran berdasarkan pilihan mahasiswa yaitu pembelajaran *online* atau *offline*. Pada penelitian ini pengumpulan data menggunakan kuesioner dengan responden mahasiswa aktif kuliah lingkup Universitas Jember angkatan 2018-2019 dan sudah pernah melaksanakan kuliah tatap muka. Analisis data yang dilakukan berdasarkan hasil responden dengan mengolah data menggunakan metode *Naïve Bayes* dan metode *Decision Tree* algoritma C4.5. Pemecahan masalah yang dilakukan yaitu perbandingan kedua metode, hasil keputusan serta hasil evaluasi pembelajaran *online* dan *offline*.

Tabel 1. Tabel Confusion Matrix

Correct Clasification	Classified as	
	Predicted “+”	Predicted “-”
Actual “+”	True Positives	False Negatives
Actual “-”	False Positives	True Negatives



Gambar 1. Kerangka Masalah

HASIL DAN PEMBAHASAN

Pengumpulan dan Perancangan Data

Pengumpulan data pada penelitian ini menggunakan teknik *sampling probability*. Jenis teknik *sampling probability* pada penelitian ini adalah metode *Cluster Random Sampling* untuk menentukan sampel

minimum yang akan dihasilkan pada pengisian kuesioner di setiap jurusan yang ada di Universitas Jember. Data yang digunakan pada penelitian ini adalah data primer, dengan mengisi kuesioner oleh mahasiswa Universitas Jember angkatan 2018 dan 2019 sebanyak 35.425 mahasiswa aktif semester ganjil tahun ajaran 2021/2022. Data yang terkumpul sebanyak 540 responden sesuai dengan jumlah *cluster* minimal.

Tabel 2. Jumlah Mahasiswa dan Hasil Pengumpulan Data Tiap Cluster

Fakultas	Jumlah Mahasiswa	Jumlah Pengumpulan Data Tiap Cluster
Hukum	2552	34
Ilmu Soisal dan Ilmu Politik	3568	44
Pertanian	3094	41
Ekonomi dan Bisnis	3733	42
Keguruan dan Ilmu Pendidikan	8182	98
Ilmu Budaya	2040	23
Teknologi Pertanian	1642	24
Kedokteran Gigi	632	9
Matematika dan Ilmu Pengetahuan Alam	1608	109
Teknik	3308	41
Kesehatan Masyarakat	1235	17
Farmasi	685	16
Keperawatan	923	10
Ilmu Komputer	1528	25
Kedokteran	693	7
Jumlah	35425	540

Tabel 3. Data Penelitian

Responden	Waktu	Interaksi	Keefektifan	Media	Keputusan
Responden 1	Teratur	Tidak Baik	Efektif	Baik	Offline
Responden 2	Tidak Teratur	Tidak Baik	Tidak Efektif	Tidak Baik	Online
Responden 3	Teratur	Baik	Efektif	Baik	Offline
Responden 4	Tidak Teratur	Tidak Baik	Tidak Efektif	Baik	Online
Responden 5	Teratur	Baik	Tidak Efektif	Baik	Offline
Responden 6	Tidak Teratur	Tidak Baik	Tidak Efektif	Baik	Online
Responden 7	Teratur	Tidak Baik	Efektif	Baik	Offline
Responden 8	Teratur	Tidak Baik	Tidak Efektif	Tidak Baik	Online
Responden 9	Teratur	Tidak Baik	Efektif	Baik	Offline
Responden 10	Teratur	Baik	Efektif	Baik	Offline
Responden 11	Tidak Teratur	Baik	Efektif	Baik	Offline
Responden 12	Tidak Teratur	Baik	Tidak Efektif	Baik	Online
Responden 13	Tidak Teratur	Baik	Tidak Efektif	Tidak Baik	Online
Responden 14	Teratur	Baik	Tidak Efektif	Baik	Offline
Responden 15	Teratur	Baik	Efektif	Baik	Offline

Data yang terkumpul dari responden kuesioner berbentuk numerik yang akan dirubah menjadi bentuk kategori setiap variabel. Pada kuesioner terdapat 16 pertanyaan dengan dua kategori pensekoran, pertanyaan urutan 1 sampai 15 memiliki kategori Setuju skor 1 dan Tidak Setuju skor 2, sedangkan pertanyaan urutan 16 memiliki kategori *online* skor 1 dan *offline* skor 2. Hasil numerik setiap pertanyaan

diambil rata-rata untuk dijadikan data penelitian variabel.

Pengujian dan Perbandingan Metode

Hasil penelitian klasifikasi menggunakan metode *Naïve Bayes* dan metode *Decision Tree* algoritma C4.5 dengan data penelitian hasil responden Mahasiswa Universitas Jember sebanyak 540 responden yang terdiri dari dua kelas kategori yaitu kelas *online* dan

offline. Penelitian ini menggunakan jumlah data training sebesar 95% dan jumlah data testing 5%.

a. Model *Naïve Bayes*

Pemodelan *Naïve Bayes* untuk klasifikasi penentuan keputusan pembelajaran kuliah online dan offline menggunakan input data training. Model *Naïve Bayes* yang dibuat yaitu menentukan nilai probabilitas prior pada suatu data berdasarkan persamaan (2). Pemodelan *Naïve Bayes* menghasilkan nilai probabilitas

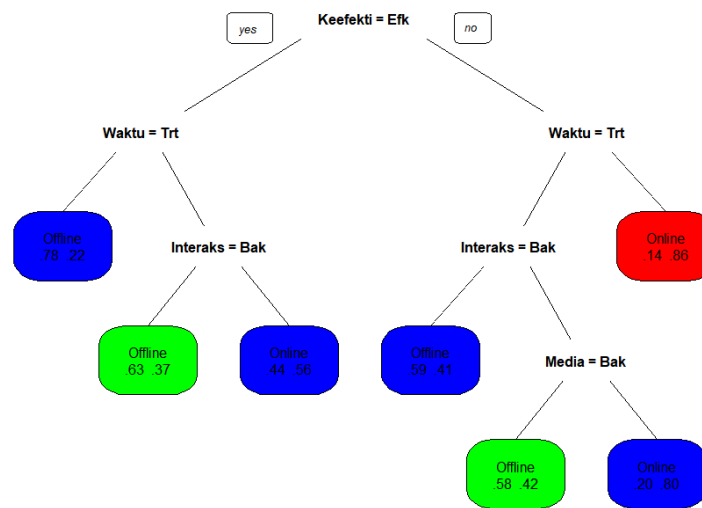
prior dan nilai probabilitas bersyarat. Hasil pemodelan *Naïve Bayes* sebagai berikut.

Tabel 4. Nilai Probabilitas *Prior*

Kelas	Nilai Probabilitas <i>Prior</i> $P(Y_i)$
Online	0,4483431
Offline	0,5516569

Tabel 5. Nilai Probabilitas Bersyarat

Variabel	Kategori	Nilai Probabilitas Bersyarat	
		Online	Offline
Waktu (X_1)	Teratur	0,4043478	0,7703180
	Tidak Teratur	0,5956522	0,2296820
Interaksi (X_2)	Baik	0,5608696	0,7844523
	Tidak Baik	0,4391304	0,2155477
Keefektifan (X_3)	Efektif	0,3478261	0,7420495
	Tidak Efektif	0,6521739	0,2579505
Media (X_4)	Baik	0,73043478	0,90106007
	Tidak Baik	0,26956522	0,09893993



Gambar 2. Pohon Keputusan

Perhitungan hasil nilai *probabilitas prior* pada Tabel 4 merupakan hasil perhitungan awal atau dugaan awal sebuah probabilitas untuk penentuan keputusan pembelajaran kuliah *online* atau *offline*. Pada Tabel 5 merupakan hasil perhitungan probabilitas bersyarat setiap variabel, probabilitas bersyarat digunakan untuk mengestimasi peluang bersyarat secara akurat pada setiap variabel penelitian.

b. *Decision Tree* algoritma C4.5

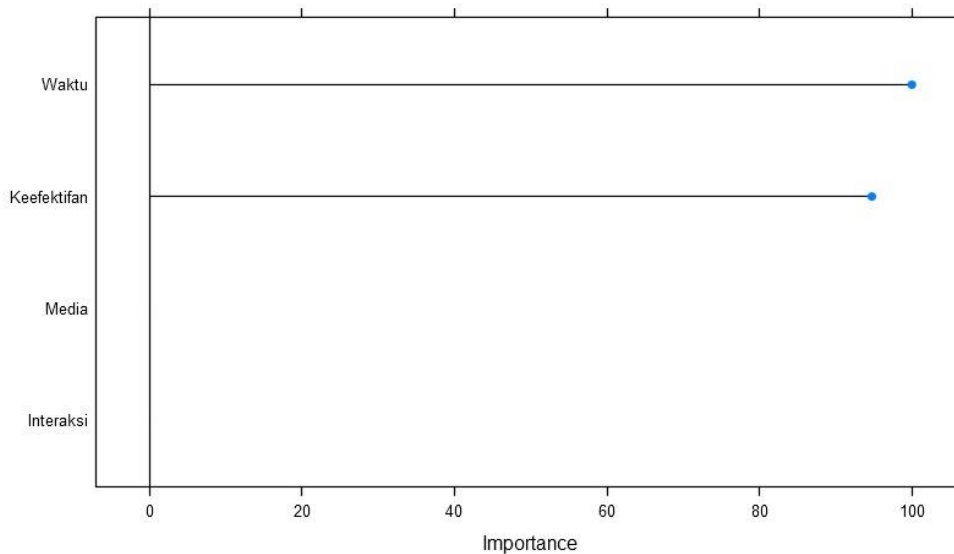
Pada penelitian metode *Decision tree* yang digunakan

yaitu dengan algoritma C4.5. Model *Decision tree* algoritma C4.5 akan dilakukan menggunakan input data *training*. Variabel keputusan (Y) dibangun pada pohon keputusan dengan semua variabel X sebanyak empat variabel yaitu X_1 adalah waktu, X_2 adalah interaksi, X_3 adalah keefektifan dan X_4 adalah media. Bentuk dari model *Decision Tree* yaitu pohon keputusan dengan *root* tertinggi yaitu keefektifan. Hasil model pohon keputusan metode *Decision Tree* algoritma C4.5 seperti pada Gambar 2.

Pohon keputusan pembelajaran online dan offline pada Gambar 2 sebagai berikut.

1. Jika keefektifan yang dihasilkan efektif dan waktu yang dihasilkan teratur maka keputusan *offline* sebesar 78%.
2. Jika keefektifan yang dihasilkan efektif, waktu yang dihasilkan tidak teratur, dan interaksi yang dihasilkan baik maka keputusan *offline* sebesar 63%.
3. Jika keefektifan yang dihasilkan efektif, waktu yang dihasilkan tidak teratur dan interaksi yang dihasilkan tidak baik maka keputusan *online* sebesar 56%.
4. Jika keefektifan yang dihasilkan tidak efektif, waktu yang dihasilkan teratur dan interaksi yang

- dihasilkan baik maka keputusan *offline* sebesar 59%.
 5. Jika keefektifan yang dihasilkan tidak efektif, waktu yang dihasilkan teratur, interaksi yang dihasilkan tidak baik dan media yang dihasilkan baik maka keputusan *offline* sebesar 58%.
 6. Jika keefektifan yang dihasilkan tidak efektif, waktu yang dihasilkan teratur, interaksi yang dihasilkan tidak baik dan media yang dihasilkan tidak baik maka keputusan *offline* sebesar 80%.
 7. Jika keefektifan yang dihasilkan tidak efektif dan waktu yang dihasilkan tidak teratur maka keputusan *online* sebesar 86%.
- c. *Variabel Importance*
 Hasil *variabel importance* sebagai berikut.



Gambar 3. Hasil *Variable Importance*

Tabel 6. Hasil Perbandingan Klasifikasi Metode *Naïve Bayes* dan Metode *Decision Tree* algoritma C4.5

No	<i>Naïve Bayes</i>	<i>Decision Tree (C4.5)</i>								
1.	Hasil <i>Confusion Matrix</i> <table border="1" style="width: 100%; border-collapse: collapse;"> <tr> <td style="text-align: center;">True Positive = 14</td> <td style="text-align: center;">False Positive = 2</td> </tr> <tr> <td style="text-align: center;">False Negative = 5</td> <td style="text-align: center;">True Negative = 6</td> </tr> </table>	True Positive = 14	False Positive = 2	False Negative = 5	True Negative = 6	Hasil <i>Confusion Matrix</i> <table border="1" style="width: 100%; border-collapse: collapse;"> <tr> <td style="text-align: center;">True Positive = 16</td> <td style="text-align: center;">False Positive = 3</td> </tr> <tr> <td style="text-align: center;">False Negative = 3</td> <td style="text-align: center;">True Negative = 5</td> </tr> </table>	True Positive = 16	False Positive = 3	False Negative = 3	True Negative = 5
True Positive = 14	False Positive = 2									
False Negative = 5	True Negative = 6									
True Positive = 16	False Positive = 3									
False Negative = 3	True Negative = 5									
2.	Hasil akurasi akhir penelitian perbandingan kuliah <i>online</i> dan <i>offline</i> sebesar 74,07%	Hasil akurasi akhir penelitian perbandingan kuliah <i>online</i> dan <i>offline</i> sebesar 77,77%								
3.	Hasil <i>precision</i> sebesar 87,5%	Hasil <i>precision</i> sebesar 84,21%								
4.	Hasil <i>sensitivity</i> sebesar 73,68%	Hasil <i>sensitivity</i> sebesar 84,21%								
5.	Hasil <i>specificity</i> sebesar 75%	Hasil <i>specificity</i> sebesar 62,5%								
6.	Hasil kelas klasifikasi dengan algoritma <i>Naïve Bayes</i> yaitu kelas <i>offline</i>	Hasil kelas klasifikasi dengan algoritma C4.5 yaitu kelas <i>offline</i>								

Hasil *variable importance* berdasarkan Gambar 3 didapatkan dari ke empat variabel yang berpengaruh adalah variabel waktu dengan kategori Teratur dan Tidak Teratur serta variabel keefektifan dengan dua kategori Efektif dan Tidak Efektif.

d. Perbandingan Metode

Penelitian ini menggunakan jumlah data testing dan data training yang sama antara kedua metode, sehingga diperoleh hasil pada setiap metode. Proses analisis pada penelitian ini juga dilakukan perbandingan hasil antara kedua metode tersebut. Berikut ini dijabarkan masing-masing hasil kedua metode tersebut.

Berdasarkan Tabel 6 diperoleh nilai akurasi metode *Decision Tree* algoritma C4.5 lebih baik dibandingkan dengan metode *Naïve Bayes*. Nilai akurasi didapatkan perbedaan karena sebaran *Confusion Matrix* pada kedua metode tersebut. Hasil *Confusion Matrix* ditentukan dari prediksi setiap metode dan prediksi yang dihasilkan berdasarkan model kedua metode yang dibentuk. Pada metode *Naïve Bayes* model yang dibentuk berbentuk nilai probabilitas prior dalam dan nilai probabilitas bersyarat yaitu nilai peluang pada suatu data, sedangkan metode *Decision Tree* algoritma C4.5 model berbentuk pohon keputusan yang memiliki nilai disetiap *node* sehingga nilai tersebut yang mempengaruhi terhadap model yang dibuat sebanyak 7 *rule* dengan *root* yaitu keefektifan. Nilai akurasi kedua metode cukup baik untuk proses klasifikasi tetapi nilai yang dihasilkan masih tidak sesuai dengan ketentuan yaitu diatas 80% sehingga nantinya dapat dikatakan sangat baik karena pada penelitian ini data yang digunakan sulit untuk dikelompokkan dan variabel yang digunakan juga tidak semua berpengaruh dan hanya dua variabel yang berpengaruh menunjukkan variabel yang berpengaruh hanya waktu dan keefektifan. Nilai *precision* yang dihasilkan dari kedua metode tersebut sudah baik dimana hasil tersebut tingkat keakuratan yang dihasilkan sudah sesuai prediksi model dan tingkat keakuratan sebagaimana data positif sudah sesuai tetapi tidak sepenuhnya karena masih terdapat kesalahan record data yaitu *False Positive* (FP) artinya pada kelas *online* yang seharusnya kelas *offline* sebanyak 2 data pada metode *Naïve Bayes* sedangkan sebanyak 3 pada metode *Decision tree* algoritma C4.5 dan *False Negative* (FN) artinya pada kelas *offline* yang seharusnya kelas *online* sebanyak 5 data pada metode *Naïve Bayes* sedangkan sebanyak 3 data pada metode *Decision tree* algoritma C4.5. Nilai *sensitivity* pada metode

Decision Tree algoritma C4.5 lebih baik dari nilai *sensitivity* metode *Naïve Bayes* dengan hasil tersebut sesuai dengan model yang dibentuk dan data yang ditemukan terrecord sesuai dengan jumlah data *input*, tetapi tidak sepenuhnya terrecord secara benar karena masih terdapat record kesalahan yaitu terletak pada *False Positive* (FP) dan *False Negative* (FN). Nilai *specificity* cukup baik dimana hasil tersebut pada data yang digunakan terdapat data yang terrecord negative yaitu pada metode *Naïve Bayes False Negative* (FN) sebanyak 5 data dan *True Negative* (TN) sebanyak 6 data sedangkan pada metode *Decision tree* algoritma C4.5 *False Negative* (FN) sebanyak 3 data dan *True Negative* (TN) sebanyak 5 data. Hasil evaluasi data kuesioner berdasarkan hasil visualisasi data dengan pilihan responden menunjukkan bahwa responden memilih *offline* karena waktu yang teratur, interaksi yang baik, keefektifan yang optimal dan media yang baik, sedangkan responden memilih *online* yaitu kebalikan dari *offline* tetapi interaksi dan media menunjukkan hasil yang sama. Pada *variable importance* yang dihasilkan bahwa variabel yang berpengaruh terhadap keputusan pembelajaran yaitu waktu dengan indikator pelaksanaan jadwal dan durasi pembelajaran sesuai sks serta keefektifan dengan indikator situasi pembelajaran dan keaktifan mahasiswa serta kualitas bahan ajar. Hasil klasifikasi kedua metode tersebut menunjukkan hasil yang sama yaitu hasil klasifikasi kategori *offline* sebanyak 16 data pada metode *Naïve Bayes* sedangkan sebanyak 19 data pada metode *Decision Tree* algoritma C4.5 dari 27 data input data *testing*.

KESIMPULAN

Hasil perbandingan diperoleh, tingkat akurasi algoritma C4.5 sebesar 77,77% lebih baik dibandingkan dengan algoritma *Naïve Bayes* sebesar 74,07% karena berdasarkan perhitungan algoritma bahwa algoritma C4.5 alur perhitungan lebih lengkap dari algoritma *Naïve Bayes*. Terdapat dua jenis kategori pembelajaran yaitu *online* dan *offline* yaitu dengan data penelitian yang digunakan yaitu data primer sebanyak 540 responden. Hasil evaluasi data kuesioner berdasarkan hasil visualisasi data dengan pilihan responden menunjukkan bahwa responden memilih *offline* karena waktu yang teratur, interaksi yang baik, keefektifan yang optimal dan media yang baik, sedangkan responden memilih *online* yaitu kebalikan dari *offline* tetapi interaksi dan media menunjukkan

hasil yang sama. Pada *variable importance* yang dihasilkan bahwa variabel yang berpengaruh terhadap keputusan pembelajaran yaitu waktu dengan indikator pelaksanaan jadwal dan durasi pembelajaran sesuai sks serta keefektifan dengan indikator situasi pembelajaran dan keaktifan mahasiswa serta kualitas bahan ajar. Hasil klasifikasi yang dihasilkan antara metode *Naïve Bayes* dan metode *Decision Tree* algoritma C4.5 yaitu sama dengan kategori *offline* sebanyak 16 data pada metode *Naïve Bayes* dan sebanyak 19 data pada metode *Decision Tree* algoritma C4.5 dari 27 input data *testing*.

DAFTAR PUSTAKA

- [1] D.S. Siltonga, Saifullah, and D. Rafika. "Analisis metode *naïve bayes* dalam memprediksi tingkat pemahaman mahasiswa terhadap mata kuliah berdasarkan posisi duduk," Prosiding Seminar Nasional Riset Information Science (SENARIS), pp. 435, 2019.
- [2] A. Yuliana, and D.B. Pratomo. "Algoritma *decision tree* (c4.5) untuk memprediksi kepuasan mahasiswa terhadap kinerja dosen Politeknik TEDC Bandung," Seminar Nasional Inovasi Teknologi UN PGRI Kediri vol. 22, no. 384, 2017
- [3] I. Andriyanto, E. Santoso, and Suprpto. "Pemodelan sistem pakar untuk menentukan penyakit diabetes mellitus menggunakan metode *naïve bayes* Studi Kasus: Puskesmas Poncokusumo Malang," *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, vol. 2, no. 2, p. 881, 2018
- [4] E. Elisa, "Analisa dan penerapan algoritma c4.5 dalam data mining untuk mengidentifikasi faktor-faktor penyebab kecelakaan kerja konstruksi PT. Arupadhatu Adisesanti," *Sistem Infromasi, Universitas Putera Batam*, vol. 2, no. 1, p. 37, Juni 2017.
- [5] T. Suryaudin, "Implementasi data mining untuk menganalisa pola penjualan barang dengan menggunakan algoritma apriori," *Jurnal Simki-Techsain*, vol 2, no. 4, 2018.
- [6] R.T. Wulandari, *Data Mining*, Yogyakarta: Gava Media, 2017.