

COMPARISON OF NAIVE BAYES, DECISION TREE, AND RANDOM FOREST ALGORITHMS IN CLASSIFYING LEARNING STYLES OF UNIVERSITAS KRISTEN INDONESIA TORAJA STUDENTS

Melki Garonga^{*1}, Rita Tanduk²

¹Informatics Engineering, Faculty of Engineering, Universitas Kristen Indonesia

²Indonesian Language and Literature Education, Faculty of Teacher Training and Education, Universitas Kristen Indonesia

Email: [1melkigaronga@ukitoraja.ac.id](mailto:melkigaronga@ukitoraja.ac.id), [2ritatanduk@ukitoraja.ac.id](mailto:ritatanduk@ukitoraja.ac.id)

(Article received: May 1, 2023; Revision: May 27, 2023; published: December 23, 2023)

Abstract

Learning style is an individual's habit or way of absorbing, processing, and managing information. This factor is very important in achieving learning goals. However, in reality, learning styles are often overlooked in the learning process, which can lead to suboptimal absorption of lessons and affect the quality of education. Various models have been developed by educational experts to identify students' learning styles, one of which is the VAK model (Visualization Auditory Kinesthetic) for grouping learning styles. This study compares algorithms in classifying learning styles using the VAK model. The results showed that the most dominant learning style was kinesthetic with a percentage of 46.9% or 478 students. The algorithm modeling showed that Naive Bayes had the highest accuracy with a value of 75%, while Random Forest had the lowest accuracy with a value of 59%. This suggests that Naive Bayes is more suitable for classifying students' learning styles. In conclusion, understanding students' learning styles is crucial for effective education. The VAK model is one way to identify learning styles, and Naive Bayes is a suitable algorithm for classifying students' learning styles. By considering learning styles, educators can tailor their teaching methods to better suit their students' needs and improve the quality of education.

Keywords: *Auditory, Kinesthetic, Learning style, Naive Bayes, Random Forest, Visual*

PERBANDINGAN ALGORITMA NAIVE BAYES, DECISION TREE DAN RANDOM FOREST DALAM KLASIFIKASI GAYA BELAJAR MAHASISWA UNIVERSITAS KRISTEN INDONESIA TORAJA

Abstrak

Gaya belajar adalah kebiasaan atau cara individu dalam menyerap, memproses, dan mengelola informasi. Faktor ini sangat penting untuk mencapai tujuan pembelajaran. Namun, dalam kenyataannya, gaya belajar sering diabaikan dalam proses pembelajaran oleh karena itu, pelajaran tidak dapat diserap secara maksimal dan dapat mempengaruhi kualitas pendidikan. Berbagai model dikembangkan oleh ahli pendidikan untuk mengidentifikasi gaya belajar peserta didik salah satu model yang digunakan untuk pengelompokan gaya belajar dengan VAK (*Visualization Auditory Kinesthetic*). Pada penelitian ini membandingkan algoritma dalam klasifikasi gaya belajar dengan menggunakan model VAK (*Visualization Auditory Kinesthetic*) dimana hasil yang diperoleh Gaya belajar paling dominan yaitu gaya belajar dengan model kinesthetic dengan persentase 46,9% atau 478 mahasiswa, Pemodelan algoritma dengan nilai akurasi tertinggi pada algoritma *Naive Bayes* dengan nilai akurasi 75% dan terendah pada algoritma *Random Forest* nilai 59% ini menunjukkan bahwa algoritma *Naive Bayes* lebih cocok digunakan untuk klasifikasi gaya belajar mahasiswa.

Kata kunci: *Auditory, Gaya belajar, Kinesthetic, Naive Bayes, Random Forest, Visual*

1. PENDAHULUAN

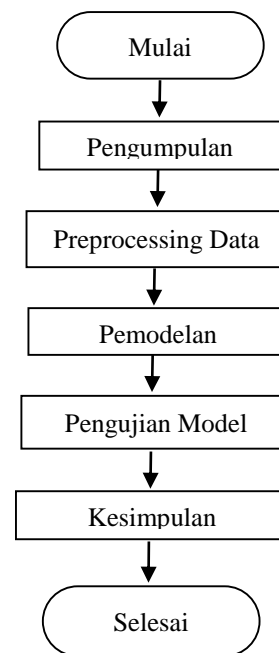
Universitas Kristen Indonesia (UKI) Toraja adalah sebuah lembaga pendidikan tinggi yang terletak di wilayah Tana Toraja, Sulawesi Selatan. UKI Toraja memiliki 11 Program Studi dengan jumlah mahasiswa mencapai 8000 mahasiswa.

Dalam menyampaikan materi pembelajaran kepada mahasiswa, penting bagi dosen untuk memahami bahwa setiap mahasiswa memiliki gaya belajar yang berbeda-beda dengan memahami gaya belajar dari setiap mahasiswa, dosen dapat menyusun materi pembelajaran yang sesuai dengan preferensi belajar mereka, sehingga akan memudahkan mahasiswa

dalam memahami dan mengingat materi yang diberikan. Hal ini akan meningkatkan efektivitas pembelajaran dan membantu mahasiswa mencapai hasil belajar yang lebih baik[1]. Gaya belajar adalah kebiasaan atau cara individu dalam menyerap, memproses, dan mengelola informasi. Faktor ini sangat penting untuk mencapai tujuan pembelajaran yang diinginkan[2]. Namun, dalam kenyataannya, gaya belajar sering diabaikan dalam proses pembelajaran oleh karena itu, pelajaran tidak dapat diserap secara maksimal dalam jangka waktu lama, dapat mempengaruhi kualitas pendidikan[3]. Berbagai model dikembangkan oleh ahli pendidikan untuk mengidentifikasi gaya belajar peserta didik salah satu model yang digunakan untuk pengelompokan gaya belajar dengan VAK (*Visualization Auditory Kinesthetic*) yaitu teknik yang memanfaatkan potensi yang dimiliki siswa dengan mengkombinasikan gaya belajar melihat, mendengar dan bergerak agar semua model belajar terpenuhi[4].

Perkembangan teknologi informasi secara khusus pada bidang kecerdasan buatan dan *machine learning* memungkinkan kita untuk menemukan solusi untuk berbagai permasalahan[5]. *Data mining* implementasi kecerdasan buatan yang memiliki serangkaian proses untuk menggali nilai tambah dari suatu kumpulan data berupa pengetahuan yang selama ini tidak diketahui secara manual[6]. Dalam bidang data mining, terdapat berbagai metode dan algoritma yang digunakan untuk melakukan klasifikasi. diantaranya adalah, *Naive Bayes*, *Decision Tree* dan *Random Forest* [7]. Ikbal Andrian Prasetyo dkk telah mengimplementasikan metode *Decision Tree* dalam pemilihan gaya belajar siswa di sekolah dasar. Hasil pengujian menunjukkan bahwa algoritma *Decision Tree* dapat membantu mengidentifikasi masing-masing gaya belajar setiap siswa dengan tingkat akurasi sebesar 99.78% [8]. Wiyli Yustanti dalam penelitian Analisis Algoritma Klasifikasi untuk Memprediksi Karakteristik Mahasiswa pada Pembelajaran Daring hasil komparasi kinerja menunjukkan bahwa algoritma *Support Vector Machine (SVM)* lebih unggul dengan nilai F-1 Score 92,8% dan AUC sebesar 99,01% [9]. Oscario, dkk melakukan penelitian dengan judul Penerapan Algoritma C4.5 Untuk Memprediksi Kecocokan Gaya Belajar Bagi Siswa Siswi Sekolah Dasar (Studi Kasus : SD SariputraJambi) dengan hasil Persentasi hasil akurasi decision tree dengan menggunakan Use Training Set dengan persentasi akurasi Correctly Classified Instances sebesar 93.5484 % [10]. Pada penelitian ini akan membandingkan algoritma *Naive Bayes*, *Decision Tree* dan *Random Forest* untuk mendapatkan nilai optimal dalam klasifikasi gaya belajar mahasiswa UKI Toraja dilihat dari nilai *accuracy*, *recall*, dan *precision*.

2. METODE PENELITIAN



Gambar 1. Tahapan Penelitian

2.1. Pengumpulan data

Tahapan dimulai dengan pengambilan *data sample* yang dikumpulkan menggunakan teknik *random sampling* yang berasal dari mahasiswa Universitas Kristen Indonesia Toraja dengan membagikan kuesioner yang terdiri dari 35 pertanyaan mengenai 3 kategori gaya belajar kepada mahasiswa yang bersedia menjadi responden. Jumlah data yang didapatkan dari penelitian ini yaitu 1019 data yang akan dibagi menjadi *data training* dan *data testing*. Data yang telah didapatkan akan diinput ke dalam aplikasi *Microsoft Excel*. Data dapat dilihat pada table 1 dibawah ini.

Tabel 1. Data Gaya Belajar V-A-K mahasiswa Universitas Kristen Indonesia Toraja

No	Nama	P1	...	P35
1	Belma Wansu P. Jumaedi	Tidak	...	Ya
2	Patandak	Tidak	...	Tidak
3	Ansel Aldi	Ya	...	Ya
4	David Ravael	Ya	...	Ya
5	Hewitno Azfriel	Ya	...	Ya
6	Misel Mikael Pata'	Ya	...	Ya
7	Alson	Ya	...	Tidak
8	Mika Samaa	Ya	...	Ya
9	Bartholomeus	Ya	...	Ya
1019	Felyati Limbong	Ya	...	Tidak

2.2. Preprocessing

Data cleaning proses menghapus atribut yang tidak diperlukan dalam melakukan klasifikasi data. Tujuan dari pembersihan data adalah untuk menghilangkan noise dan data yang tidak konsisten, tidak relevan, tidak hilang, dan tidak redundan. Dalam proses ini, data yang tidak diperlukan akan dihapus atau diubah agar lebih mudah diproses dan dianalisis. Hal ini akan membantu meningkatkan kualitas data dan meningkatkan akurasi hasil klasifikasi data. Setelah dilakukan pembersihan data, data akan direstrukturisasi ulang agar lebih mudah diinterpretasikan dan dianalisis [11]

Transformation Data adalah proses mengubah data yang telah dipilih agar sesuai untuk diproses melalui teknik data mining. Proses coding dalam Knowledge Discovery in Database (KDD) merupakan proses kreatif yang sangat tergantung pada jenis atau pola informasi yang akan dicari dalam basis data. Dengan kata lain, coding melibatkan konversi data mentah menjadi format yang cocok untuk analisis dan pengenalan pola. Ini merupakan tahap penting dalam proses data mining, karena membantu mempersiapkan data untuk analisis lebih lanjut dan ekstraksi wawasan yang berguna. Data dapat dilihat pada tabel 2 dibawah ini.

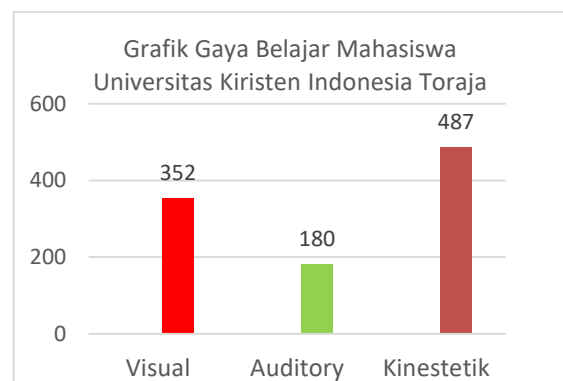
Tabel 2. Data Gaya Belajar Proses Encoding

No	Nama	P1	...	P35
1	Aurianti	0	...	1
2	Ellin	0	...	0
3	Riselina S	1	...	1
4	Marce	1	...	1
5	Karnita Ending	1	...	1
6	Maylin Tiku	1	...	1
7	Lilian Yulius	1	...	0
8	Grace S. Moga	1	...	1
9	Valencia Bunga	1	...	1
1019	Felyati Limbong	1	...	0

Data selection merupakan proses pembagian kelas data menjadi tiga kategori berdasarkan model VAK (*Visualization Auditory Kinesthetic*). Tabel sebaran data dapat dilihat pada tabel 3 dibawah ini.

Tabel 3. Sebaran Data Berdasarkan Kelas VAK

No	Nama	P1	...	P35	KET
1	Belma Wansu P.	0	...	1	V
2	Jumaedi Patandak	0	...	0	A
3	Ansel Aldi	1	...	1	V
4	David Ravael	1	...	1	V
5	Hewitno Azfriel	1	...	1	V
6	Misel Mikael Pata'	1	...	1	K
7	Alson	1	...	0	A
8	Mika Samaa	1	...	1	V
9	Bartholomeus	1	...	1	K
1019	Felyati Limbong	1	...	0	V



Gambar 2 Grafik Data Gaya Belajar Kelas V-A-K

Berdasarkan table 2.3 diatas data akan dibagi menjadi dua bagian yaitu *data training* 80% dan *data testing* berjumlah 20% . Pembagian dataset dapat dilihat pada tabel 4 dibawah ini.

Tabel 4. Pembagian Dataset

<i>Training</i>	80%	815
<i>Testing</i>	20%	204
Total	100%	1019

2.3. Pemodelan

Tahapan pemodelan algoritma merupakan tahapan training data dengan dataset yang telah disediakan. Proses training dilakukan dengan 3(tiga) algoritma klasifikasi *Decision Tree*, *Random frorest* dan *Naive Bayes* menggunakan aplikasi *orange*.

1) Klasifikasi

Klasifikasi analisis data yang menentukan kelas dan lebel dari sebuah kelas sampel data yang akan di klasifikasikan. Metode ini merupakan metode *supervised learning* dengan menghubungkan data training dengan data target. Metode klasifikasi mencoba mendapatkan kemiripan antara atribut

masuk dengan atribut keluaran guna menghasilkan model yang termasuk dalam tahapan pelatihan.

2) *Decision Tree*

Algoritma *Decision Tree* adalah metode pengklasifikasian atau regresi yang menggunakan model pohon keputusan. Model ini menggambarkan setiap atribut atau fitur sebagai node dalam pohon, dan setiap cabang pohon mewakili kemungkinan nilai dari atribut tersebut. Setiap daun di pohon mewakili kelas atau nilai target yang diprediksi [12]. Sebelum memilih akar pertama dalam pembuatan model *Decision Tree*, langkah pertama yang harus dilakukan adalah menghitung nilai entropy dari dataset. Entropy mengukur seberapa homogen atau heterogen dataset, dan nilainya berkisar antara 0 dan 1. Semakin homogen dataset, semakin rendah nilai entropy, dan sebaliknya. Berikut adalah cara menghitung entropy algoritma *Decision Tree* C4.5 :

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

Keterangan:

S : Himpunan Kasus,
n : Jumlah partisi S
 p_i : Proporsi dari S_i terhadap S

Setelah nilai entropy dihitung, langkah selanjutnya adalah menghitung nilai gain untuk setiap atribut dalam dataset. Nilai gain mengukur seberapa besar pengaruh atribut dalam memisahkan dataset menjadi kelas-kelas yang berbeda. Atribut dengan nilai gain tertinggi akan dipilih sebagai akar pertama dalam pembuatan model *decision Tree*.

3) *Random Forest*

Random Forest pengembangan dari algoritma *Decision Tree* memiliki beberapa kelebihan, yaitu dapat meningkatkan hasil akurasi, efisien untuk penyimpanan sebuah data serta mempunyai proses seleksi fitur dimana mampu mengambil fitur terbaik sehingga dapat meningkatkan performa terhadap model klasifikasi. Dengan adanya seleksi fitur tentu *Random Forest* dapat bekerja pada big data dengan parameter yang kompleks secara efektif.[13]. *Random Forest* adalah algoritma yang terdiri dari beberapa predictor tree atau biasa disebut decision tree. Setiap tree dalam forest tersebut bergantung pada nilai random vector yang diambil secara acak dan merata pada semua tree. Hasil prediksi dari *Random Forest* didapatkan melalui hasil terbanyak dari setiap individual decision tree, yaitu dengan

menggunakan voting untuk klasifikasi dan rata-rata untuk regresi. Dengan kata lain, *Random Forest* menggabungkan hasil prediksi dari banyak decision tree untuk mendapatkan hasil yang lebih akurat dan konsisten [14].

4) *Naive Bayes*

Naive Bayes Classifier adalah sebuah metode klasifikasi yang didasarkan pada Teorema Bayes. Metode ini menggunakan probabilitas dan statistik untuk memprediksi peluang terjadinya suatu kejadian di masa depan berdasarkan pengalaman di masa lalu [13].

$$P(C_i|X) = (P(X|C_i) * P(C_i)) / P(X) \quad (2)$$

Keterangan:

$P(C_i|X)$ = Probabilitas a posteriori dari kelas C_i untuk fitur X
 $P(X|C_i)$ = Probabilitas fitur X dalam kelas C_i
 $P(C_i)$ = Probabilitas a priori dari kelas C_i
 $P(X)$ = Probabilitas fitur X

5) *Confusion matrix*

Confusion matrix adalah metode perhitungan untuk menganalisis kualitas model klasifikasi yang di bentuk. Dalam perhitungan nilai *Confusion matrix* dikenal istilah *True Positive* (TF) yaitu nilai yang berlabel positive yang diprediksi benar, *True Negative* (TN) yaitu nilai berlabel negative yang prediksi benar, *False Positif* (FP) yaitu nilai berlabel *positive* yang prediksi salah oleh model, *False negatif* (FN) yaitu nilai berlabel negatif yang prediksi salah oleh model. [16].

$$accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (3)$$

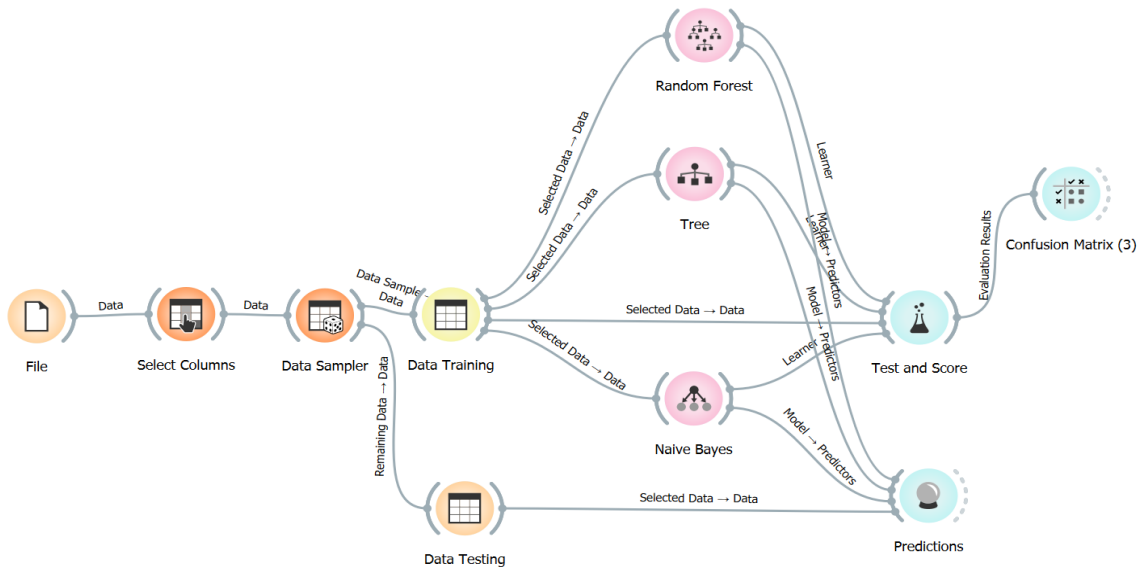
3. HASIL DAN PEMBAHASAN

Berdasarkan tabel 2 maka di peroleh hasil dari data 1019 responden terbagi dalam 3 kelas gaya paling dominan yaitu gaya belajar dengan model *kinesthetic* dengan persentase 46,9% atau 478 mahasiswa dan yang paling rendah adalah *auditory* dengan persentase 17,66% atau 180 mahasiswa. Dari data – data tersebut kemudian dimodelkan dalam 3 algoritma menggunakan aplikasi *orange*. Adapun tipe data yang digunakan pada data dalam algoritma dapat dilihat pada tabel 5 dibawah ini.

Tabel 5. Tipe Data

Atribut	Type	Role	Values
Nama	<i>Text</i>	<i>Meta</i>	-
Fakultas	<i>Text</i>	<i>Meta</i>	-
Angkatan	<i>Numeric</i>	<i>Meta</i>	-
P1	<i>Categorical</i>	<i>feature</i>	1, 0
P2	<i>Categorical</i>	<i>feature</i>	1, 0
P3	<i>Categorical</i>	<i>feature</i>	1, 0
P4	<i>Categorical</i>	<i>feature</i>	1, 0
P5	<i>Categorical</i>	<i>feature</i>	1, 0
P6	<i>Categorical</i>	<i>feature</i>	1, 0
P7	<i>Categorical</i>	<i>feature</i>	1, 0
P8	<i>Categorical</i>	<i>feature</i>	1, 0
P9	<i>Categorical</i>	<i>feature</i>	1, 0
P10	<i>Categorical</i>	<i>feature</i>	1, 0
P11	<i>Categorical</i>	<i>feature</i>	1, 0
P12	<i>Categorical</i>	<i>feature</i>	1, 0
P13	<i>Categorical</i>	<i>feature</i>	1, 0
P14	<i>Categorical</i>	<i>feature</i>	1, 0
P15	<i>Categorical</i>	<i>feature</i>	1, 0
P16	<i>Categorical</i>	<i>feature</i>	1, 0
P17	<i>Categorical</i>	<i>feature</i>	1, 0
P18	<i>Categorical</i>	<i>feature</i>	1, 0
P19	<i>Categorical</i>	<i>feature</i>	1, 0
P20	<i>Categorical</i>	<i>feature</i>	1, 0
P21	<i>Categorical</i>	<i>feature</i>	1, 0
P22	<i>Categorical</i>	<i>feature</i>	1, 0
P23	<i>Categorical</i>	<i>feature</i>	1, 0
P24	<i>Categorical</i>	<i>feature</i>	1, 0
P25	<i>Categorical</i>	<i>feature</i>	1, 0
P26	<i>Categorical</i>	<i>feature</i>	1, 0
P27	<i>Categorical</i>	<i>feature</i>	1, 0
P28	<i>Categorical</i>	<i>feature</i>	1, 0
P29	<i>Categorical</i>	<i>feature</i>	1, 0
P30	<i>Categorical</i>	<i>feature</i>	1, 0
P31	<i>Categorical</i>	<i>feature</i>	1, 0
P32	<i>Categorical</i>	<i>feature</i>	1, 0
P33	<i>Categorical</i>	<i>feature</i>	1, 0
P34	<i>Categorical</i>	<i>feature</i>	1, 0
P35	<i>Categorical</i>	<i>feature</i>	1, 0
KET	<i>Categorical</i>	<i>target</i>	A,K,V

Pemodelan pada penelitian ini dilakukan dengan cara membuat koneksi antara *data training* dan algoritma. Proses koneksi akan tampil seperti gambar 3 dibawah ini.



Gambar 3. Desain Widget Klasifikasi

Berdasarkan gambar 3 diatas, dapat dilihat proses pembangunan model langkah awal yang dilakukan pada aplikasi *orange* yaitu menginput *file* yang bertipe excel selanjutnya dibagi 2 yaitu data training dan data testing dengan data *sampler*. Data training digunakan untuk membangun model dengan komposisi 80% dari keseluruhan data dan data testing untuk menguji model yang terbentuk dengan komposisi 20% selanjutnya membangun model dengan menggunakan algoritma dengan *Decision Tree*, *Random Forest* dan *Naive Bayes* selanjutnya untuk mengetahui akurasi model yang terbentuk maka menggunakan confusion matrix dapat dilihat pada gambar berikut.

		Predicted			Σ
		A	K	V	
Actual	A	1	19	18	38
	K	0	82	12	94
	V	0	34	37	71
Σ		1	135	67	203

Gambar 4. Hasil *Confusion Matrix* Algoritama *Random Forest*

$$accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

$$accuracy = \frac{1 + 82 + 37}{1 + 0 + 0 + 19 + 82 + 34 + 18 + 12 + 37}$$

$$accuracy = \frac{120}{203}$$

$$accuracy = 0,59$$

		Predicted			Σ
		A	K	V	
Actual	A	23	7	8	38
	K	13	62	19	94
	V	15	14	42	71
Σ		51	83	69	203

Gambar 5. Hasil *Confusion Matrix* Algoritama *Decision Tree*

$$accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

$$accuracy = \frac{23 + 62 + 42}{23 + 13 + 15 + 7 + 62 + 14 + 8 + 19 + 42}$$

$$accuracy = \frac{127}{203}$$

$$accuracy = 0,62$$

		Predicted			Σ
		A	K	V	
Actual	A	25	9	4	38
	K	7	78	9	94
	V	2	20	49	71
Σ		34	107	62	203

Gambar 6. Hasil *Confusion Matrix* Algoritama *Naive Bayes*

$$accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

$$accuracy = \frac{25 + 78 + 49}{25 + 7 + 2 + 9 + 78 + 20 + 4 + 9 + 49}$$

$$accuracy = \frac{152}{203}$$

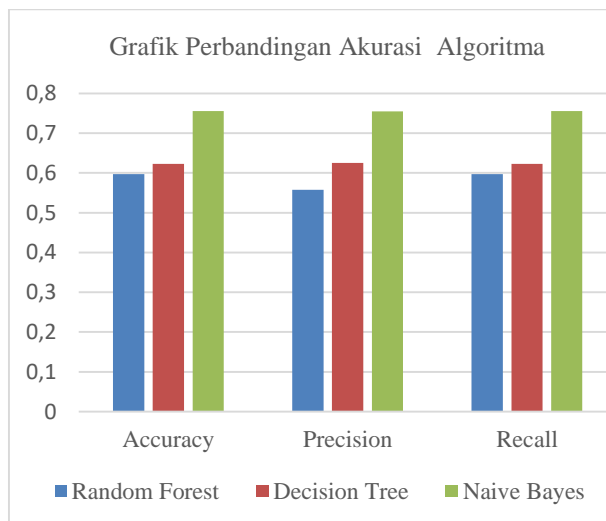
$$accuracy = 0,75$$

Dari gambar 4, 5 dan 6 maka diperoleh hasil pengujian dari ketiga algoritma yang digunakan untuk membangun model klasifikasi gaya belajar dapat dilihat pada table 5

Tabel 6. Hasil akurasi masing – masing metode

Metode	Accuracy	Precision	Recall
<i>Random Forest</i>	0,597	0,558	0,597
<i>Decision Tree</i>	0,623	0,625	0,623
<i>Naive Bayes</i>	0,756	0,755	0,756

Berdasarkan Tabel 6. diperoleh nilai akurasi tertinggi adalah 0,756 dengan metode *Naive Bayes*, dan terendah adalah 0,597 dengan metode *Random Forest*. Kemudian untuk *precision* tertinggi adalah 0,755 dengan *Naive Bayes* dan terendah adalah 0,558 dengan metode *Random Forest*. Lalu untuk *recall* tertinggi diperoleh 0,756 dengan metode *Naive Bayes* dan terendah 0,597 dengan metode *Random Forest*.



Gambar 7 Grafik Perbandingan akurasi Algoritma

Dari gambar 7. Grafik Perbandingan akurasi Algoritma *Naive Bayes* lebih tinggi dibandingkan dengan *Random Forest* dan *Decision Tree* dilihat dari nilai *accuracy*, *recall*, dan *precision*

3. DISKUSI

Penelitian yang dilakukan oleh S Rama dan Jahring [17] dengan judul Klasifikasi Gaya Belajar Mahasiswa Menggunakan Metode *Naive Bayes Classifier* hasil dari penelitian ini menunjukkan nilai akurasi algoritma *Naive Bayes* dalam klasifikasi gaya belajar mahasiswa dengan persentase 90% ini sejalan dengan penelitian ini dimana algoritma *Naive Bayes* tertinggi dengan tingkat akurasi 75%.

Penelitian yang dilakukan Dinda Novita Sari [18] dengan judul penerapan data mining untuk

klasifikasi gaya belajar siswa menggunakan algoritma c4.5 dengan hasil algoritma c4.5 memiliki akurasi tertinggi pada klasifikasi 3 fold *cross validation* dengan akurasi 81% berbeda dengan penelitian ini nilai akurasi algoritma c4.5 menurun karena dipengaruhi jumlah dataset dan banyaknya atribut yang digunakan.

Hal ini menunjukkan bahwa kecocokan data dengan algoritma yang digunakan dapat mempengaruhi tingkat akurasi yang dicapai.

4. KESIMPULAN

Pada penelitian ini , klasifikasi gaya belajar dengan model VAK (*visual, auditory, kinesthetic*) dengan jumlah data 1019 data yang diambil dari responden mahasiswa UKI Toraja dengan menggunakan kuisioner dapat disimpulkan bahwa.

1. Gaya belajar paling dominan yaitu gaya belajar dengan model *kinesthetic* dengan persentase 46,9% atau 478 mahasiswa.
2. Pemodelan algoritma dengan nilai akurasi tertinggi pada algoritma *Naive Bayes* dengan nilai akurasi 75% dan terendah pada algoritma *Random Forest* nilai 59% ini menunjukkan bahwa algoritma *Naive Bayes* lebih cocok digunakan untuk klasifikasi gaya belajar mahasiswa .
3. Hasil penelitian ini dapat diinterpretasikan sebagai pentingnya memahami gaya belajar mahasiswa, karena hal ini dapat membantu para pendidik dalam merancang pembelajaran yang lebih efektif dan efisien.

UCAPAN TERIMA KASIH

Ucapan terima kasih disampaikan kepada Universitas kristen Indonesia Toraja (UKI TORAJA) yang memberikaan bantuan dana dan memberikan kesempatan untuk melakukan melakukan pengambilan data tentang gaya belajar mahasiswa.

DAFTAR PUSTAKA

- [1] Y. Wahyuni, "Identifikasi Gaya Belajar (Visual, Auditorial, Kinestetik) Mahasiswa Pendidikan Matematika Universitas Bung Hatta," *J. Penelit. dan Pembelajaran Mat.*, vol. 10, no. 2, pp. 128–132, 2017, doi: 10.30870/jppm.v10i2.2037.
- [2] T. F. Prasetyo and M. Iqbal, "Sistem Pakar Identifikasi Gaya Belajar Mahasiswa Berbasis Web," *Semin. Nas. Sains dan Teknol. 2016 Fak. Tek. Univ. Muhammadiyah Jakarta*, no. November, pp. 1–7, 2016, [Online]. Available: <https://jurnal.umj.ac.id/index.php/semnastek/article/view/776>.

- [3] F. D. Widayanti, "Pentingnya Mengetahui Gaya Belajar Siswa Dalam Kegiatan Pembelajaran Di Kelas," *Erud. J. Educ. Innov.*, vol. 2, no. 1, 2013, doi: 10.18551/erudio.2-1.2.
- [4] R. K. Dewi, W. Sunarno, and S. Budiawanti, "Pengaruh Model Somatic Auditory Visualization Intellectually (SAVI) dan Visualization Auditory Kinesthetic (VAK) dengan Pendekatan Ilmiah Terhadap Hasil Belajar Fisika Siswa SMA Ditinjau dari Gaya Belajar," *J. Mater. dan Pembelajaran Fis.*, vol. 11, no. 2, p. 75, 2021, doi: 10.20961/jmpf.v11i2.49008.
- [5] E. N. Fitri *et al.*, "DECISION TREE SIMPLIFICATION THROUGH FEATURE SELECTION," *J. Tek. Inf.*, vol. 4, no. 2, 2023.
- [6] Y. Mardi, "Data Mining : Klasifikasi Menggunakan Algoritma C4.5," *Edik Inform.*, vol. 2, no. 2, pp. 213–219, 2017, doi: 10.22202/ei.2016.v2i2.1465.
- [7] F. S. Pamungkas, B. D. Prasetya, and I. Kharisudin, "Perbandingan Metode Klasifikasi Supervised Learning pada Data Bank Customers Menggunakan Python," *Prism. Pros. Semin. Nas. Mat.*, vol. 3, pp. 692–697, 2020, [Online]. Available: <https://journal.unnes.ac.id/sju/index.php/prisma/article/view/37875>.
- [8] U. Islam, N. Sunan, G. Djati, and M. F. Rachman, "Implementasi Metode Decision Tree Untuk," *J. Inform.*, vol. 10, no. 1, 2020.
- [9] W. Yustanti and N. Rochmawati, "Analisis Algoritma Klasifikasi untuk Memprediksi Karakteristik Mahasiswa pada Pembelajaran Daring," *JEPIN (Jurnal Edukasi dan Penelit. Inform.*, vol. 8, no. 1, pp. 57–61, 2022.
- [10] Oscario, Jasmir, And Yudi, "Penerapan Algoritma C4.5 Untuk Memprediksi Kecocokan Gaya Belajar Bagi Siswa Siswi Sekolah Dasar (Studi Kasus : Sd Sariputra Algoritma C4.5 Untuk Memprediksi Kecocokan Gaya Belajar Bagi Siswa Siswi Sekolah Dasar (Studi Kasus : SD Sariputra Jambi)," *J. Process.*, vol. 14, no. 2, pp. 141–152, 2019, doi: 10.33998/processor.2019.14.2.637.
- [11] M. I. P. Hant and H. Hendry, "Data Mining Technique Using Naïve Bayes Algorithm To Predict Shopee Consumer Satisfaction Among Millennial Generation," *J. Tek. Inform.*, vol. 3, no. 4, pp. 829–838, 2022, doi: 10.20884/1.jutif.2022.3.4.295.
- [12] Robianto ; Sampe Hotlan Sitorus ; Uray Ristian, "Penerapan Metode Decision Tree Untuk Mengklasifikasikan Mutu Buah Jeruk Berdasarkan Fitur Warna Dan Ukuran," *J. Komput. dan Apl.*, vol. 9, no. 01, pp. 76–86, 2021.
- [13] R. Supriyadi, W. Gata, N. Maulidah, and A. Fauzi, "Penerapan Algoritma Random Forest Untuk Menentukan Kualitas Anggur Merah," *E-Bisnis J. Ilm. Ekon. dan Bisnis*, vol. 13, no. 2, pp. 67–75, 2020, doi: 10.51903/e-bisnis.v13i2.247.
- [14] D. Y. Kardono, Y. M. Pranoto, and E. Setyati, "Prediksi Kecocokan Jurusan Siswa SMK Dengan Support Vector Machine dan Random Forest," *Teknika*, vol. 12, no. 1, pp. 11–17, 2023, doi: 10.34148/teknika.v12i1.567.
- [15] N. Agustina, A. Adrian, and M. Hermawati, "Implementasi Algoritma Naïve Bayes Classifier untuk Mendeteksi Berita Palsu pada Sosial Media," *Fakt. Exacta*, vol. 14, no. 4, pp. 1979–276, 2021, doi: 10.30998/faktorexacta.v14i4.11259.
- [16] Gde Agung Brahmana Suryanegara, Adiwijaya, and Mahendra Dwifabri Purbolaksono, "Peningkatan Hasil Klasifikasi pada Algoritma Random Forest untuk Deteksi Pasien Penderita Diabetes Menggunakan Metode Normalisasi," *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 5, no. 1, pp. 114–122, 2021, doi: 10.29207/resti.v5i1.2880.
- [17] S. Ramadandi and J. Jahring, "Student Learning Style Classification Using Naïve Bayes Classifier Method," *J. Teknol. dan Inf.*, vol. 10, no. 2, pp. 170–179, 2020, doi: 10.34010/jati.v10i2.3096.
- [18] D. N. Sari, H. Oktavianto, and I. Saifudin3, "Penerapan Data Mining Untuk Klasifikasi Gaya Belajar Siswa Menggunakan Algoritma C4.5 Application," *J. Smart Teknol.*, vol. 3, no. 2, pp. 184–190, 2022.