

Klasifikasi Gaya Belajar VARK Siswa Sekolah Dasar Menggunakan K-Nearest Neighbor dan Naive Bayes Berbasis Kuesioner

Deden Adi Mardian Lesmana¹, Fathoni Mahardika², Dani Indra Junaedi³

Jurusan Informatika, Universitas Sebelas April, Jl. Angkrek Situ No.19, Situ, Kec. Sumedang Utara, Kabupaten Sumedang, Jawa Barat 45323 Indonesia^{1,2,3}

220660121158@student.unsap.ac.id¹, fathoni@unsap.ac.id², dani@unsap.ac.id³

Abstract – Learning styles play a crucial role in determining how students comprehend information, particularly in elementary school, where they are at the concrete thinking stage of development. This study aimed to map student learning preferences using the VARK (Visual, Aural, Read/Write, Kinesthetic) model and to develop a classification model based on the K-Nearest Neighbor (KNN) algorithm. Forty fourth-sixth grade students participated by completing a 16-item VARK questionnaire. Data were processed through cleaning, Min-Max normalization, and four-dimensional feature formation. The KNN model was tested using varying k values through Stratified 5-Fold Cross Validation. The results showed that the Kinesthetic and Visual categories were the most dominant learning preferences at SDN Cipatat. The KNN model performed best at $k = 1$, with an average accuracy of 95%, a precision of 0.81, a recall of 0.80, and an F1-score of 0.79. Confusion matrix analysis showed that the Kinesthetic and Visual categories were more predictable, while the Aural and Read/Write categories had greater feature overlap. These findings suggest that a data-driven approach can provide an objective picture of students' learning preferences and support more adaptive learning strategies.

Keywords: *Learning Style, VARK, K-Nearest Neighbor (KNN), Educational Data Mining, Elementary School.*

Abstrak – Gaya belajar memiliki peranan penting dalam menentukan bagaimana siswa memahami informasi, khususnya pada jenjang sekolah dasar yang berada pada tahap perkembangan berpikir konkret. Penelitian ini bertujuan memetakan preferensi belajar siswa menggunakan model VARK (Visual, Aural, Read/Write, Kinesthetic) serta membangun model klasifikasi berbasis algoritma K-Nearest Neighbor (KNN). Sebanyak 40 siswa kelas IV–VI berpartisipasi dengan mengisi 16 item kuesioner VARK. Data diolah melalui proses pembersihan, normalisasi Min–Max, dan pembentukan fitur empat dimensi. Model KNN diuji menggunakan variasi nilai k melalui Stratified 5-Fold Cross Validation. Hasil penelitian menunjukkan bahwa kategori Kinesthetic dan Visual merupakan preferensi belajar yang paling dominan di SDN Cipatat. Model KNN memberikan performa terbaik pada $k = 1$, dengan akurasi rata-rata 95%, precision 0,81, recall 0,80, dan F1-score 0,79. Analisis confusion matrix memperlihatkan bahwa kategori Kinesthetic dan Visual lebih mudah diprediksi, sementara Aural dan Read/Write memiliki tumpang tindih fitur yang lebih besar. Temuan ini menunjukkan bahwa pendekatan berbasis data dapat memberikan gambaran objektif mengenai preferensi belajar siswa serta mendukung strategi pembelajaran yang lebih adaptif.

Kata Kunci: *Gaya Belajar, VARK, K-Nearest Neighbor (KNN), Penambangan Data Pendidikan, Sekolah Dasar.*

I. PENDAHULUAN

Pendidikan dasar adalah masa paling penting dalam perjalanan belajar seorang anak. Pada tahap ini, anak mulai membangun cara berpikir, berinteraksi, serta mengenali potensi dirinya [1]. Karena itu, keberhasilan belajar di

tingkat dasar tidak cukup hanya dengan memberikan materi pelajaran, tetapi juga bergantung pada bagaimana guru memahami cara belajar setiap anak

Setiap siswa memiliki cara sendiri dalam menyerap dan mengingat pelajaran. Ada yang lebih mudah memahami lewat gambar, ada

yang lebih suka mendengar penjelasan, dan ada pula yang lebih aktif jika belajar sambil bergerak. Perbedaan ini dikenal sebagai gaya belajar, yaitu cara khas seseorang dalam menerima dan mengolah informasi. Jika guru memahami perbedaan ini, proses belajar akan terasa lebih hidup dan bermakna bagi siswa. Salah satu pendekatan yang banyak digunakan untuk mengenali gaya belajar adalah model VARK, yang membagi cara belajar menjadi empat kategori: visual, aural (pendengaran), read/write (membaca dan menulis), serta kinestetik. Model ini membantu guru untuk melihat kecenderungan siswa dan menyesuaikan cara mengajar agar lebih sesuai dengan kebutuhan mereka.

Seiring dengan berkembangnya teknologi, pengolahan data pendidikan kini bisa dilakukan dengan bantuan *Educational Data Mining* (EDM). Melalui EDM, data dari kuesioner atau hasil belajar siswa dapat diolah menjadi informasi yang membantu guru mengambil keputusan yang lebih tepat. Salah satu metode yang populer dalam analisis data pendidikan adalah *K-Nearest Neighbor* (KNN) [4]. Metode ini sering digunakan dalam machine learning karena kesederhanaannya [5], mudah dipahami [6], tidak memerlukan asumsi kompleks, dan mampu memberikan hasil klasifikasi yang mudah dipahami oleh guru maupun peneliti [7], [8]. KNN merupakan algoritma supervised learning yang digunakan untuk klasifikasi maupun regresi, yang prinsip kerjanya membandingkan data baru dengan data yang sudah ada berdasarkan jarak terdekat (lazy learning) [9]–[11].

Beberapa penelitian menunjukkan bahwa KNN dapat digunakan untuk mengenali pola belajar dan mengelompokkan siswa berdasarkan preferensi belajarnya [12]. Dengan memilih parameter yang tepat, algoritma ini dapat membantu menghasilkan model klasifikasi yang stabil dan berguna sebagai dasar pengambilan keputusan dalam kegiatan pembelajaran.

Penelitian ini dilakukan untuk menerapkan algoritma KNN dalam mengelompokkan gaya belajar siswa berdasarkan data kuesioner model VARK dan mengevaluasi kinerjanya. Evaluasi performa model dilakukan menggunakan metrik akurasi, presisi, recall, F1-score, serta analisis confusion matrix. Penggunaan beberapa

indikator ini memungkinkan peneliti menilai kinerja model secara lebih menyeluruh, terutama dalam memetakan kategori gaya belajar yang memiliki sebaran data berbeda.

II. TINJAUAN PUSTAKA

A. Konsep Gaya Belajar VARK

Gaya belajar merupakan cara khas yang dimiliki setiap individu dalam menerima, mengolah, dan memahami informasi selama proses pembelajaran. Perbedaan gaya belajar menyebabkan setiap siswa memiliki kecenderungan yang berbeda dalam memahami materi yang diberikan oleh guru. Menurut Fleming, model VARK merupakan salah satu pendekatan yang banyak digunakan untuk mengidentifikasi preferensi belajar seseorang berdasarkan empat kategori utama, yaitu *Visual*, *Aural*, *Read/Write*, dan *Kinesthetic* [1].

Gaya belajar *Visual* (V) menunjukkan kecenderungan siswa lebih mudah memahami informasi melalui gambar, diagram, grafik, warna, dan berbagai bentuk tampilan visual lainnya. Siswa dengan gaya belajar ini biasanya lebih mudah mengingat materi yang disajikan dalam bentuk ilustrasi atau peta konsep.

Gaya belajar *Aural* (A) atau auditori merupakan kecenderungan belajar melalui pendengaran. Siswa lebih mudah memahami materi dengan cara mendengarkan penjelasan guru, diskusi kelompok, maupun media audio.

Gaya belajar *Read/Write* (R) menggambarkan preferensi belajar melalui aktivitas membaca dan menulis. Siswa dengan karakteristik ini lebih menyukai penggunaan buku, catatan, ringkasan, maupun latihan menulis untuk memperdalam pemahaman materi.

Sementara itu, gaya belajar *Kinesthetic* (K) merupakan kecenderungan belajar melalui pengalaman langsung, praktik, percobaan, atau aktivitas fisik. Siswa dengan gaya belajar ini lebih mudah memahami konsep apabila terlibat secara aktif dalam kegiatan pembelajaran.

Pengenalan gaya belajar VARK dapat membantu guru menentukan strategi pembelajaran yang lebih sesuai dengan karakteristik siswa sehingga proses belajar mengajar menjadi lebih efektif dan adaptif.

B. Educational Data Mining (EDM)

Educational Data Mining (EDM) merupakan bidang ilmu yang memanfaatkan teknik penambangan data untuk menganalisis informasi yang berasal dari lingkungan pendidikan. EDM bertujuan menemukan pola, hubungan, dan informasi tersembunyi dari data pendidikan sehingga dapat digunakan sebagai dasar pengambilan keputusan yang lebih objektif [3].

Penerapan EDM mencakup berbagai aspek, seperti prediksi prestasi belajar, analisis perilaku siswa, identifikasi kesulitan belajar, hingga klasifikasi gaya belajar. Dengan memanfaatkan algoritma machine learning, data yang sebelumnya hanya berupa kumpulan informasi dapat diolah menjadi pengetahuan yang berguna bagi guru maupun pihak sekolah.

Dalam penelitian ini, konsep EDM digunakan untuk mengolah data hasil kuesioner VARK sehingga dapat diketahui kecenderungan gaya belajar siswa melalui proses klasifikasi menggunakan algoritma *K-Nearest Neighbor*.

C. Algoritma *K-Nearest Neighbor* (KNN)

K-Nearest Neighbor (KNN) merupakan salah satu algoritma supervised learning yang digunakan untuk menyelesaikan permasalahan klasifikasi maupun regresi. Prinsip dasar algoritma ini adalah menentukan kelas suatu data baru berdasarkan mayoritas kelas dari sejumlah tetangga terdekat (*nearest neighbor*) pada data latih [5].

KNN termasuk ke dalam kategori lazy learning karena proses pembelajaran tidak dilakukan pada saat pelatihan model, tetapi ketika proses prediksi dilakukan. Kedekatan antar data biasanya dihitung menggunakan metode Euclidean Distance yang dirumuskan sebagai berikut:

$$d(x,y)=\sqrt{\sum(x_i-y_i)^2} \quad (1)$$

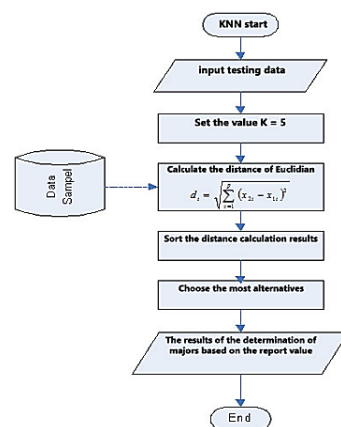
Nilai parameter *k* menunjukkan jumlah tetangga terdekat yang digunakan dalam proses klasifikasi. Pemilihan nilai *k* yang tepat sangat berpengaruh terhadap performa model. Nilai *k* yang terlalu kecil dapat menyebabkan model sensitif terhadap noise, sedangkan nilai *k* yang terlalu besar dapat mengurangi kemampuan model dalam membedakan karakteristik antar kelas.

Algoritma KNN memiliki beberapa kelebihan, antara lain mudah diimplementasikan, tidak memerlukan asumsi distribusi data, serta mampu memberikan hasil klasifikasi yang baik pada dataset berukuran kecil hingga menengah. Oleh karena itu, algoritma ini banyak diterapkan dalam berbagai penelitian di bidang pendidikan, termasuk untuk klasifikasi gaya belajar siswa.

III. ANALISA DAN PERANCANGAN MODEL

Studi ini mengadopsi pendekatan kuantitatif dengan memanfaatkan algoritma Machine Learning untuk melakukan klasifikasi gaya belajar VARK pada siswa Sekolah Dasar. Sumber data penelitian berasal dari hasil pengisian kuesioner gaya belajar VARK yang dilakukan pada siswa kelas IV–VI di SDN Cipatat, Kecamatan Buahdua, Kabupaten Sumedang, dengan total 40 responden. Data kuesioner tersebut digunakan sebagai dasar dalam pembentukan dataset penelitian.

Tahap awal penelitian mencakup proses pra-pemrosesan data untuk memastikan kualitas data yang digunakan. Proses ini meliputi pemeriksaan data tidak lengkap, pemilihan atribut yang relevan, serta penyesuaian format data agar siap digunakan dalam proses klasifikasi. Data kuesioner selanjutnya direpresentasikan dalam bentuk numerik berupa skor Visual, Aural, Read/Write, dan Kinesthetic. Untuk algoritma *K-Nearest Neighbor*, dilakukan normalisasi data guna menyamakan skala antar fitur. Setelah tahap pra-pemrosesan, dataset digunakan dalam proses pelatihan dan evaluasi model klasifikasi menggunakan metode cross validation.



Gambar 1 Flowchart Metode Penelitian

Gambar 1 menunjukkan alur metode penelitian yang dimulai dari pengumpulan data kuesioner VARK, dilanjutkan dengan pra-pemrosesan data, pembentukan fitur dan normalisasi, proses klasifikasi menggunakan algoritma K-Nearest Neighbor, hingga tahap evaluasi performa model.

Model klasifikasi yang digunakan dalam penelitian ini adalah algoritma K-Nearest Neighbor (KNN), yang diterapkan untuk mengklasifikasikan gaya belajar VARK siswa Sekolah Dasar berdasarkan data kuesioner. Kinerja model dievaluasi menggunakan beberapa indikator evaluasi, meliputi *accuracy*, *precision*, *recall*, dan *F1-score*. Penggunaan beberapa metrik evaluasi bertujuan untuk memberikan gambaran performa model secara menyeluruh pada setiap kategori gaya belajar yang dianalisis.

A. Sumber dan Pengambilan Data

Dalam penelitian ini, data diperoleh dari hasil pengisian kuesioner gaya belajar VARK yang disebarkan kepada siswa Sekolah Dasar di SDN Cipatat, Kecamatan Buahdua, Kabupaten Sumedang. Responden penelitian berjumlah 40 siswa yang berasal dari kelas IV hingga VI.

Kuesioner yang digunakan mengacu pada model gaya belajar VARK (Visual, Aural, Read/Write, dan Kinesthetic). Setiap responden mengisi sejumlah pertanyaan yang merepresentasikan preferensi belajar mereka. Jawaban responden kemudian dikonversi ke dalam bentuk skor numerik untuk masing-masing kategori gaya belajar, sehingga data yang diperoleh dapat diproses menggunakan algoritma *Machine Learning*.

Data hasil kuesioner selanjutnya disusun dalam bentuk dataset yang terdiri dari empat atribut utama, yaitu skor Visual, Aural, Read/Write, dan Kinesthetic, yang digunakan sebagai fitur dalam proses klasifikasi gaya belajar siswa.

TABLE 1
CONTOH DATA KUESIONER GAYA BELAJAR VARK SISWA

No	Kode Siswa	Visual (V)	Aural (A)	Read/Write (R)	Kinesthetic (K)
1	S01	2	1	1	2
2	S02	1	2	1	2
3	S03	3	0	1	2
4	S04	1	1	2	1
...
40	S40	2	1	0	3

Tabel 1 menunjukkan contoh struktur dataset kuesioner gaya belajar VARK siswa Sekolah Dasar. Setiap responden direpresentasikan dalam empat atribut numerik, yaitu Visual, Aural, Read/Write, dan Kinesthetic, yang digunakan sebagai fitur dalam proses klasifikasi menggunakan algoritma K-Nearest Neighbor.

Pada penelitian ini, atribut yang digunakan sebagai fitur masukan terdiri atas skor Visual (V), Aural (A), Read/Write (R), dan Kinesthetic (K) yang diperoleh dari hasil pengisian kuesioner VARK. Sementara itu, kelas (label) ditentukan berdasarkan kategori gaya belajar yang memiliki skor tertinggi pada setiap responden. Dengan demikian, setiap data siswa direpresentasikan dalam bentuk empat atribut numerik (V, A, R, K) dan satu label kelas berupa Visual, Aural, Read/Write, atau Kinesthetic.

Dataset yang digunakan dalam penelitian ini terdiri dari 40 responden siswa Sekolah Dasar yang diperoleh melalui pengisian kuesioner gaya belajar VARK. Dataset tersebut merepresentasikan variasi karakteristik gaya belajar siswa pada jenjang Sekolah Dasar di lingkungan penelitian.

Sebelum dilakukan proses klasifikasi, data kuesioner terlebih dahulu melalui tahap pra-pemrosesan data. Tahapan ini meliputi pemeriksaan dan penanganan data yang tidak lengkap, pemilihan atribut yang relevan, serta penyesuaian format data agar dapat diolah secara numerik. Selain itu, dilakukan normalisasi pada variabel numerik guna

menyamakan skala antar fitur, khususnya untuk mendukung perhitungan jarak pada algoritma K-Nearest Neighbor. Proses pra-pemrosesan dilakukan untuk memastikan dataset memiliki kualitas yang layak sehingga dapat diolah secara optimal dalam proses klasifikasi gaya belajar VARK siswa.

B. Desain Penelitian

Desain penelitian disusun agar seluruh proses berjalan terarah dan setiap tahap saling mendukung. Untuk mencapai tujuan penelitian, beberapa langkah utama diterapkan secara berurutan. Penjelasan lengkapnya adalah sebagai berikut.

1. Pengumpulan Data Melalui Kuesioner VARK

Tahapan dimulai dengan penyebaran kuesioner VARK kepada siswa. Kuesioner ini berisi situasi-situasi sederhana yang menggambarkan bagaimana mereka biasanya belajar atau memproses informasi. Dari sini diperoleh data dasar mengenai kecenderungan gaya belajar masing-masing responden. Tahap ini menjadi fondasi utama karena seluruh analisis selanjutnya bergantung pada akurasi jawaban yang diberikan siswa.

2. Pemeriksaan Kelengkapan dan Pembersihan Data

Setelah kuesioner terkumpul, setiap lembar diperiksa satu per satu untuk memastikan tidak ada jawaban yang kosong atau terduplikasi. Data yang tidak memenuhi kriteria kelengkapan tidak dilanjutkan ke tahap analisis. Langkah ini penting untuk menjaga kualitas data agar proses klasifikasi tidak terpengaruh oleh kesalahan input atau data yang tidak konsisten.

3. Pembentukan Fitur dan Normalisasi Data

Jawaban siswa kemudian diubah menjadi bentuk numerik melalui vektor empat dimensi [V,A,R,K]. Vektor ini merepresentasikan jumlah pilihan untuk kategori Visual, Aural, Read/Write, dan Kinesthetic. Setelah itu, nilai-nilai tersebut dinormalisasi menggunakan *Min-Max Scaling* agar skala antar atribut berjalan seimbang. Normalisasi membantu model KNN bekerja lebih optimal karena setiap fitur berada pada rentang nilai yang setara.

4. Penerapan Algoritma KNN untuk Klasifikasi

Tahap berikutnya adalah menerapkan algoritma KNN untuk menentukan kategori

gaya belajar berdasarkan kedekatan antar data. Beberapa nilai k —seperti 3, 5, 7, dan 9—diuji untuk menemukan parameter yang paling sesuai dengan karakter data. Perhitungan jarak antar titik data menggunakan rumus Euclidean yang mengukur seberapa besar perbedaan antara profil satu siswa dengan siswa lainnya.

5. Evaluasi Performansi Model

Setelah proses klasifikasi dilakukan, model diuji menggunakan metode *Stratified 5-Fold Cross Validation*. Teknik ini membagi data dalam lima kelompok dengan komposisi kelas yang tetap seimbang pada tiap fold. Melalui cara ini diperoleh gambaran performa model yang lebih stabil. Pengukuran dilakukan menggunakan beberapa metrik seperti akurasi, precision, recall, F1-score, dan analisis menggunakan confusion matrix.

6. Analisis Hasil dan Penyusunan Rekomendasi

Tahap akhir adalah menganalisis pola yang muncul dari hasil klasifikasi. Distribusi gaya belajar siswa diidentifikasi untuk mengetahui kecenderungan dominan pada masing-masing kelas. Dari temuan tersebut, disusunlah rekomendasi pembelajaran yang dapat membantu guru menyesuaikan strategi mengajar dengan kebutuhan belajar siswa.

C. Metode Pengembangan Model

Metode pengembangan model yang digunakan dalam penelitian ini disesuaikan dengan proses pengembangan model *Machine Learning*. Tahapan pengembangan model mengikuti alur penelitian eksperimental yang terstruktur ditunjukkan pada Gambar 2, yang meliputi pengumpulan data, pra-pemrosesan data, pelatihan model, pengujian model, serta evaluasi hasil klasifikasi.

Tahap pengumpulan data dilakukan melalui penyebaran kuesioner gaya belajar VARK kepada siswa. Data yang diperoleh kemudian diproses melalui tahap pembersihan data, pembentukan fitur numerik, dan normalisasi data. Selanjutnya, model klasifikasi dilatih menggunakan algoritma K-Nearest Neighbor dengan beberapa variasi nilai parameter k . Model yang telah dilatih kemudian dievaluasi menggunakan metode *Stratified K-Fold Cross Validation* untuk memperoleh hasil pengujian yang lebih stabil

dan representatif. Tahap akhir penelitian adalah analisis performa model berdasarkan metrik evaluasi klasifikasi dan interpretasi hasil klasifikasi gaya belajar siswa.



Gambar 2. Diagram Alur Penelitian

IV. IMPLEMENTASI DAN PEMBAHASAN

Bab ini menjabarkan hasil penelitian yang diperoleh berdasarkan implementasi model klasifikasi gaya belajar VARK menggunakan algoritma K-Nearest Neighbor (KNN). Pembahasan difokuskan pada proses implementasi model, hasil pengujian yang diperoleh, serta interpretasi hasil dalam konteks pembelajaran di Sekolah Dasar.

5.1 Implementasi Model Klasifikasi

Implementasi model dilakukan dengan menerapkan algoritma K-Nearest Neighbor pada dataset hasil kuesioner gaya belajar VARK siswa Sekolah Dasar. Data yang digunakan telah melalui tahap pra-pemrosesan, meliputi pembersihan data, pembentukan fitur numerik Visual, Aural, Read/Write, dan Kinesthetic, serta normalisasi data menggunakan metode Min-Max Scaling.

Model KNN diimplementasikan menggunakan bahasa pemrograman Python dengan bantuan pustaka *scikit-learn*. Beberapa nilai parameter k diuji untuk memperoleh konfigurasi model yang memberikan performa terbaik. Proses pengujian dilakukan menggunakan metode Stratified K-Fold Cross Validation guna memastikan hasil evaluasi yang stabil dan representatif terhadap distribusi kelas gaya belajar.

Hasil implementasi menunjukkan bahwa model mampu melakukan klasifikasi gaya belajar siswa berdasarkan kedekatan karakteristik fitur dengan data latih yang tersedia.

5.2 Hasil Pengujian Model

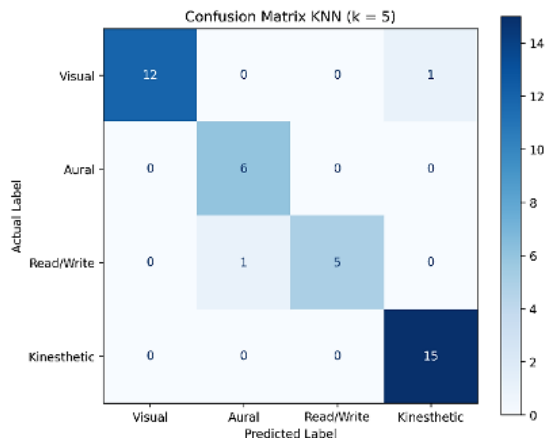
Berdasarkan hasil pengujian, nilai parameter $k=5$ menghasilkan performa klasifikasi terbaik dibandingkan nilai k lainnya. Model dengan konfigurasi tersebut mampu mencapai akurasi sebesar 95%, dengan nilai precision, recall, dan F1-score yang relatif seimbang.

TABLE 2
PERBANDINGAN PERFORMA NILAI K

Nilai k	Accuracy	Precision	Recall	F1-Score
1	97,5%	0,977	0,975	0,975
2	87,5%	0,879	0,875	0,874
3	95,0%	0,955	0,950	0,950
4	95,0%	0,954	0,950	0,950
5	95,0%	0,955	0,950	0,950

Berdasarkan hasil pengujian pada beberapa nilai parameter k, diperoleh bahwa algoritma K-Nearest Neighbor menunjukkan performa klasifikasi yang berbeda pada setiap konfigurasi. Nilai $k = 2$ menghasilkan akurasi terendah sebesar 87,5%, sedangkan nilai $k = 3, 4, \text{ dan } 5$ memberikan performa terbaik dengan accuracy mencapai 95%.

Hasil tersebut menunjukkan bahwa pemilihan nilai k berpengaruh terhadap performa model klasifikasi gaya belajar VARK. Nilai k yang terlalu kecil cenderung lebih sensitif terhadap variasi data sehingga meningkatkan kemungkinan kesalahan klasifikasi. Sebaliknya, penggunaan nilai k yang lebih stabil mampu menghasilkan prediksi yang lebih konsisten pada dataset gaya belajar siswa Sekolah Dasar. Berdasarkan hasil evaluasi menggunakan accuracy, precision, recall, dan F1-score, nilai $k = 1$ dipilih sebagai konfigurasi terbaik karena memberikan performa klasifikasi yang stabil dan seimbang pada seluruh metrik evaluasi.



Gambar 3 Confusion Matrix (Heatmap)

Berdasarkan confusion matrix pada hasil klasifikasi algoritma K-Nearest Neighbor dengan nilai $k = 1$, sebagian besar data berhasil diklasifikasikan dengan benar. Kategori Kinesthetic menunjukkan tingkat prediksi tertinggi dengan seluruh data berhasil dikenali secara tepat. Kategori Visual juga menunjukkan performa yang baik dengan hanya satu kesalahan klasifikasi ke kategori Kinesthetic.

Confusion matrix disusun berdasarkan hasil prediksi terhadap label gaya belajar dominan siswa. Label yang digunakan pada confusion matrix bukan merupakan atribut masukan, melainkan hasil kategorisasi akhir yang terdiri atas empat kelas, yaitu Visual, Aural, Read/Write, dan Kinesthetic.

Pada kategori Read/Write ditemukan satu kesalahan prediksi ke kategori Aural. Hal ini menunjukkan adanya kemiripan karakteristik antar fitur pada beberapa responden sehingga model mengalami kesulitan dalam membedakan kedua kategori tersebut.

Secara keseluruhan, confusion matrix menunjukkan bahwa algoritma KNN mampu mengenali pola gaya belajar siswa dengan baik dan menghasilkan performa klasifikasi yang stabil pada dataset penelitian.

5.3 Pembahasan Hasil Penelitian

Hasil penelitian menunjukkan bahwa sebagian besar siswa Sekolah Dasar memiliki kecenderungan gaya belajar Kinesthetic dan Visual. Temuan ini sejalan dengan karakteristik perkembangan kognitif siswa Sekolah Dasar yang lebih responsif terhadap

pembelajaran berbasis aktivitas langsung dan media visual.

Dari sisi metode, algoritma K-Nearest Neighbor terbukti mampu mengklasifikasikan gaya belajar siswa secara cukup akurat pada dataset berukuran kecil. Namun, performa model sangat dipengaruhi oleh distribusi data dan jumlah sampel pada masing-masing kategori. Ketidakseimbangan data menyebabkan model lebih mudah mengenali kelas mayoritas dibandingkan kelas minoritas.

Secara keseluruhan, implementasi model klasifikasi berbasis Machine Learning ini menunjukkan bahwa pendekatan berbasis data dapat digunakan sebagai alat bantu untuk memahami preferensi belajar siswa secara lebih objektif. Informasi yang dihasilkan dari model ini berpotensi mendukung guru dalam merancang strategi pembelajaran yang lebih adaptif dan sesuai dengan karakteristik siswa.

V. PENGUJIAN

Pada bagian ini dijelaskan langkah-langkah pengujian yang telah dilakukan untuk mengetahui apakah model klasifikasi gaya belajar VARK yang dibangun menggunakan algoritma K-Nearest Neighbor (KNN) dapat berfungsi dengan benar. Pengujian dilakukan dengan mengevaluasi kinerja model klasifikasi menggunakan data yang telah diproses sebelumnya. Hasil pengujian disajikan dalam bentuk tabel dan dianalisis berdasarkan metrik evaluasi yang digunakan.

6.1 Skenario Pengujian

Pengujian model dilakukan terhadap model klasifikasi KNN dengan beberapa skenario pengujian yang meliputi proses pelatihan model, proses prediksi gaya belajar, serta evaluasi performa model menggunakan data uji. Pengujian bertujuan untuk memastikan bahwa model mampu mengklasifikasikan gaya belajar siswa sesuai dengan karakteristik data yang diberikan.

6.2 Hasil Pengujian

Hasil pengujian model klasifikasi gaya belajar VARK menggunakan algoritma KNN disajikan pada Tabel 3.

TABLE 3
HASIL PENGUJIAN MODEL VARK

Skenario Pengujian	Tes #1	Tes #2	Hasil Tes
Pengujian Pra-pemrosesan Data	Data kuesioner diperiksa dan dibersihkan dari data kosong	Data berhasil direpresentasikan dalam fitur numerik V, A, R, K	Sesuai
Pengujian Normalisasi Data	Data dinormalisasi menggunakan Min-Max Scaling	Seluruh fitur berada pada rentang nilai yang sama	Sesuai
Pengujian Pelatihan Model KNN	Model dilatih menggunakan data latih	Model berhasil dibangun tanpa error	Sesuai
Pengujian Prediksi Gaya Belajar	Data uji dimasukkan ke dalam model Perhitungan	Model menghasilkan label gaya belajar	Sesuai
Pengujian Evaluasi Model	akurasi, precision, recall, dan F1-score	Nilai evaluasi berhasil ditampilkan	Sesuai

6.3 Pembahasan Pengujian

Berdasarkan hasil pengujian yang ditampilkan pada Tabel 3, model klasifikasi gaya belajar VARK menggunakan algoritma KNN dapat berjalan dengan baik pada setiap tahapan pengujian. Seluruh proses, mulai dari pra-pemrosesan data hingga evaluasi model, berhasil dilakukan sesuai dengan skenario pengujian yang dirancang.

Model mampu menerima data masukan berupa hasil kuesioner, melakukan normalisasi, serta menghasilkan prediksi gaya belajar siswa. Selain itu, model juga dapat menampilkan hasil evaluasi performa model menggunakan metrik klasifikasi yang telah ditentukan. Dengan demikian, dapat disimpulkan bahwa model yang dibangun telah berfungsi sesuai dengan tujuan penelitian.

VI. KESIMPULAN

Berdasarkan hasil analisis, perancangan, implementasi, dan pengujian model yang telah dilakukan, dapat disimpulkan bahwa model klasifikasi gaya belajar siswa Sekolah Dasar menggunakan model VARK dan algoritma K-Nearest Neighbor (KNN) telah berhasil dibangun dan berfungsi dengan baik.

Hasil penelitian menunjukkan bahwa algoritma KNN mampu mengklasifikasikan gaya belajar siswa berdasarkan data kuesioner VARK dengan tingkat akurasi yang cukup baik. Model dengan nilai parameter $k=4$ & $k=5$ memberikan performa terbaik dibandingkan nilai k lainnya, dengan hasil evaluasi yang relatif seimbang pada metrik akurasi, precision, recall, dan F1-score.

Pengujian model membuktikan bahwa seluruh tahapan proses, mulai dari pra-pemrosesan data, normalisasi, pelatihan model, hingga evaluasi performa, dapat dijalankan sesuai dengan skenario yang dirancang. Model mampu menghasilkan prediksi gaya belajar siswa secara konsisten berdasarkan karakteristik data yang diberikan.

Secara umum, hasil klasifikasi menunjukkan bahwa gaya belajar Kinesthetic dan Visual lebih dominan pada siswa Sekolah Dasar. Temuan ini dapat menjadi dasar bagi guru dalam merancang strategi pembelajaran yang lebih adaptif dan sesuai dengan karakteristik belajar siswa.

Sebagai pengembangan selanjutnya, penelitian ini masih dapat ditingkatkan dengan menambah jumlah responden, memperluas cakupan sekolah, serta membandingkan algoritma KNN dengan metode klasifikasi lain agar diperoleh hasil yang lebih optimal dan generalisasi yang lebih luas.

REFERENSI

- [1] N. D. Fleming, *Teaching and Learning Styles: VARK Strategies*. Christchurch: N.D. Fleming, 2006.
- [2] R. Dunn and K. Dunn, *Teaching Students Through Their Individual Learning Styles*. Reston: Reston Publishing, 1978.
- [3] R. S. Baker and P. S. Inventado, "Educational data mining and learning analytics," in *Learning Analytics*, Springer, 2014, pp. 61–75.
- [4] J. Han, M. Kamber, and J. Pei, *Data Mining: Concepts and Techniques*, 3rd ed. Morgan Kaufmann, 2012.
- [5] T. Cover and P. Hart, "Nearest neighbor pattern classification," *IEEE Transactions on Information Theory*, vol. 13, no. 1, pp. 21–27, 1967.
- [6] G. Shalev-Shwartz and S. Ben-David, *Understanding Machine Learning: From Theory to Algorithms*. Cambridge University Press, 2014.

- [7] L. Rokach and O. Maimon, *Data Mining with Decision Trees: Theory and Applications*. World Scientific, 2014.
- [8] M. A. Hall, "Correlation-based feature selection for machine learning," Ph.D. thesis, Univ. of Waikato, 1999.
- [9] S. B. Kotsiantis, "Supervised machine learning: A review of classification techniques," *Informatica*, vol. 31, no. 3, pp. 249–268, 2007.
- [10] M. Witten, E. Frank, and M. A. Hall, *Practical Machine Learning Tools and Techniques*. Morgan Kaufmann, 2016.
- [11] N. L. Singh and R. Kaur, "Application of KNN for classification of students based on learning behavior," *International Journal of Computer Science and Information Technologies*, vol. 5, no. 3, pp. 3484–3489, 2014.
- [12] M. Romero and B. Ventura, "Educational data mining: A review of the state of the art," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 40, no. 6, pp. 601–618, 2010.
- [13] P. A. Kirschner, "Cognitive load theory: Implications of cognitive load theory on learning and instruction," *Learning and Instruction*, vol. 12, no. 1, pp. 1–10, 2002.
- [14] S. Graf, T. Liu, and Kinshuk, "Identifying learning styles in learning management systems by using indicators and data mining techniques," *Interactive Learning Environments*, vol. 17, no. 1, pp. 89–102, 2009.
- [15] A. M. Ahmad, "Using KNN classification to identify student learning styles," *International Journal of Emerging Technologies in Learning (IJET)*, vol. 11, no. 4, pp. 49–55, 2016.
- [16] P. Cortez and A. Silva, "Using data mining to predict secondary school student performance," in *Proceedings of the 5th Future Business Technology Conference*, 2008.