

## **Deteksi Gaya Belajar Siswa SMA pada Virtual Based Learning Environment(VBLE) dengan Decision Tree C4.5 dan Naive Bayes**

**Aang Kisnu Darmawan<sup>1,\*</sup>, Kurdianto<sup>1</sup>, Bakir<sup>1</sup>, Masdukil Makruf<sup>2</sup>**

<sup>1</sup>Fakultas Teknik, Prodi Sistem Informasi, Universitas Islam Madura, Pamekasan, Indonesia

<sup>2</sup>Fakultas Teknik, Prodi Teknik Informatika, Universitas Islam Madura, Pamekasan, Indonesia

Email: <sup>1,\*</sup>[ak.darmawan@gmail.com](mailto:ak.darmawan@gmail.com), <sup>2</sup>[si.kurdianto@gmail.com](mailto:si.kurdianto@gmail.com), <sup>3</sup>[bakir.madura@gmail.com](mailto:bakir.madura@gmail.com), <sup>4</sup>[masdukil.makruf@gmail.com](mailto:masdukil.makruf@gmail.com)

Email Penulis Korespondensi: [ak.darmawan@gmail.com](mailto:ak.darmawan@gmail.com)

**Abstrak**—Deteksi atau identifikasi gaya belajar siswa memiliki peranan yang signifikan dalam meningkatkan proses belajar mengajar. Guru yang menyadari gaya belajar siswanya yang beragam dapat melakukan adaptasi pembelajaran atau menyusun instruksi pembelajaran dengan tepat sedangkan siswa yang menyadari gaya belajarnya dapat melakukan adaptasi metode belajarnya sehingga menghasilkan luaran proses pembelajaran yang lebih baik. Akan tetapi mendeteksi gaya belajar siswa merupakan hal yang kompleks dan menantang serta terdapat berbagai faktor yang dapat mempengaruhi akurasi deteksi. Berbagai kajian pustaka tentang deteksi gaya belajar siswa mencatat terdapat berbagai permasalahan yang terjadi pada Virtual Based Learning Education(VBLE) yaitu kurangnya penyelidikan tentang kesesuaian gaya belajar mengajar sehingga sulit meningkatkan motivasi belajar siswa, instruksi tradisional beberapa guru yang masih bertentangan dengan sistem gaya belajar siswa sehingga sulit untuk beradaptasi dengan lingkungan belajar. Kebanyakan siswa tidak mengetahui pembelajaran yang efektif dan tidak memberikan perhatian yang diperlukan untuk gaya belajar mereka. Penelitian ini bertujuan dua(2) hal, pertama mendeteksi gaya belajar siswa sekolah menengah atas swasta(SMAS) pada VBLE dengan algoritma Decision Tree(DT) C4.5 dan Naive Bayes(NB) dan kedua melakukan komparasi performa algoritma DT dan NB untuk konteks deteksi gaya belajar siswa sekolah menengah atas swasta pada VBLE. Pengumpulan data dilakukan dengan kuesioner google form pada 252 responden terdiri dari alumni dan siswa SMAS melalui 30 pertanyaan tentang gaya belajar. Gaya belajar siswa diidentifikasi menjadi gaya belajar visual, auditori dan kinestetik. Untuk analisa data digunakan tool google colab dengan pemrograman python. Sedangkan proses evaluasi performa algoritma diukur menggunakan parameter accuracy, precision, recall/sensitivity, f-measure dan time consume. Hasil penelitian menunjukkan bahwa Algoritma DT dan NB bagus dalam memprediksi gaya belajar di SMAS, tetapi algoritma NB(98%) sedikit lebih baik daripada algoritma DT(96%) dalam hal accuracy, precision, recall/sensitivity, f-measure. sedangkan untuk time consume DT(0,00505 detik) jauh lebih cepat daripada NB(3,2895 detik). Penelitian ini memberikan **kontribusi** dalam dua(2) hal, pertama secara saintifik dengan menguji algoritma DT dan NB untuk deteksi gaya belajar siswa SMAS pada konteks VBLE dan secara praktis memberikan rekomendasi pada stakeholder pendidikan yaitu guru, siswa, wali murid dan pemangku kebijakan pendidikan untuk lebih memberikan perhatian pada deteksi gaya belajar siswa.

**Kata Kunci:** Gaya Belajar Siswa; Virtual-Based Learning Environment; Decision Tree; Naïve Bayes; Machine Learning

**Abstract**—A Detection or identification of student learning styles has a significant role in improving the teaching and learning process. Teachers who are aware of the diverse learning styles of their students can make learning adaptations or prepare learning instructions appropriately, while students who are aware of their learning styles can adapt their learning methods so as to produce better learning process outcomes. However, detecting student learning styles is complex and challenging and there are various factors that can affect the accuracy of detection. Various literature studies on detecting student learning styles note that there are various problems that occur in Virtual Based Learning Education (VBLE), namely the lack of investigations regarding the suitability of teaching and learning styles making it difficult to increase student learning motivation, the traditional instructions of some teachers which are still in conflict with the student learning style system so that it is difficult to adapt to the learning environment and most students do not know effective learning and do not pay the necessary attention to their learning style. This study aims at two (2) things, the first is to detect the learning styles of private high school students (SMAS) in VBLE with the Decision Tree (DT) C4.5 and Naive Bayes (NB) algorithms and the second is to compare the performance of DT and NB algorithms for the context of style detection. private high school students study on VBLE. Data collection was carried out using a google form questionnaire on 252 respondents consist of alumni and students through 30 questions about learning styles. Student learning styles are identified as visual, auditory and kinesthetic learning styles. For data analysis, the Google Colab tool is used with Python programming. While the process of evaluating the performance of the algorithm is measured using the parameters accuracy, precision, recall/sensitivity, f-measure and time consumption. The results showed that the DT and NB algorithms were good at predicting SMAS learning styles, but the NB algorithm (98%) was slightly better than the DT algorithm (96%) in terms of accuracy, precision, memory/sensitivity, f-measure, while for time consumption DT (0,00505 seconds) much faster than NB(3,2895 seconds). This research contributes in two (2) ways, first scientifically by testing the DT and NB algorithms for detecting learning styles of SMAS students in the VBLE context and practically providing recommendations to education stakeholders namely teachers, students, student guardians and education policymakers to further pay attention to the detection of student learning styles.

**Keywords:** Student Learning Style; Virtual-Based Learning Environment; Decision Tree; Naïve Bayes; Machine Learning

### **1. PENDAHULUAN**

Deteksi gaya belajar siswa merupakan proses mengidentifikasi pendekatan belajar yang disukai siswa untuk semua situasi belajar[1]. Penting untuk mempelajari deteksi gaya belajar siswa karena penelitian menunjukkan bahwa siswa memperoleh lebih banyak pengetahuan, menyimpan lebih banyak informasi, dan tampil lebih baik ketika gaya mengajar guru cocok dengan gaya belajar siswa[1]. Gaya belajar mengacu pada pendekatan belajar yang disukai siswa untuk semua situasi belajar sedangkan gaya mengajar mengacu pada perilaku guru, keyakinan dan metode pembelajaran yang dipilih digunakan untuk menyajikan pelajaran kepada siswa[2]. Keragaman dalam preferensi gaya belajar mengajar akan menghasilkan kesesuaian dan ketidaksesuaian antara gaya mengajar guru dan gaya belajar siswa[3]. Ada berbagai

pendekatan untuk mendeteksi gaya belajar siswa, termasuk *Bayesian Network* dan *Tree Augmented Naive Bayesian Network*[4]. Dengan mengidentifikasi gaya belajar siswa, guru dapat merencanakan strategi pembelajaran dan kegiatan relevan yang sesuai dengan gaya belajar siswa, yang dapat meningkatkan kualitas proses pembelajaran dan meningkatkan prestasi akademik[1], [4], [5]. Penting untuk belajar tentang deteksi gaya belajar siswa karena penelitian terbaru menunjukkan bahwa kecocokan antara gaya mengajar dan belajar membantu memotivasi proses belajar siswa[6]. Ketika guru mengidentifikasi gaya mengajar mereka sendiri serta gaya belajar siswa mereka, mereka dapat memperoleh hasil yang lebih baik di kelas[6]. Memahami preferensi gaya belajar siswa penting untuk keberhasilan dalam belajar[7].

Namun, mendeteksi gaya belajar siswa merupakan hal yang kompleks dan menantang karena terdapat berbagai faktor yang dapat mempengaruhi akurasi deteksi gaya belajar siswa. Berdasarkan hasil penelitian yang telah diberikan, terdapat beberapa masalah yang dapat diidentifikasi untuk prediksi gaya belajar siswa pada jenjang pendidikan SMA. Salah satu permasalahannya adalah kurangnya penyelidikan mengenai kesesuaian gaya belajar mengajar[1]. Karena kesenjangan ini, sulit untuk meningkatkan motivasi belajar siswa[1]. Selain itu, instruksi tradisional yang digunakan oleh beberapa guru mungkin bertentangan dengan sistem gaya belajar siswa, sehingga sulit bagi mereka untuk beradaptasi dengan lingkungan belajar[8]. Masalah lain adalah bahwa siswa mungkin tidak mengetahui pembelajaran yang efektif dan tidak dapat memberikan perhatian yang diperlukan untuk gaya belajar mereka[5]. Selain itu, beberapa peneliti menunjukkan bahwa gaya belajar, gaya mengajar, dan penilaian saling terkait satu sama lain, dan guru dapat merencanakan strategi pembelajaran dan kegiatan yang relevan secara efektif jika mereka mengetahui gaya belajar siswa[9]. Oleh karena itu, penting untuk mengidentifikasi gaya belajar siswa dan mencocokkannya dengan metodologi pengajaran yang diadopsi oleh anggota fakultas untuk meningkatkan kualitas proses pembelajaran[9], [10].

*Virtual based Learning Environment (VBLE)* atau Pembelajaran berbasis lingkungan virtual mengacu pada penggunaan lingkungan virtual, seperti dunia virtual, laboratorium virtual, dan realitas virtual, untuk memfasilitasi pembelajaran[11]–[13]. Pendidikan pembelajaran berbasis virtual penting karena memberikan kesempatan kepada siswa untuk belajar dengan cara yang lebih interaktif dan menarik, yang dapat meningkatkan hasil belajar[12], [14]. Namun, ada juga tantangan yang terkait dengan pendidikan pembelajaran berbasis virtual, seperti kurangnya investigasi mengenai pencocokan gaya belajar mengajar, yang dapat mempersulit peningkatan motivasi belajar siswa[11], [13]. Selain itu, siswa mungkin menghadapi hambatan yang membatasi akses ke pakar konten dan pertumbuhan skolastik, khususnya di daerah pedesaan[12]. Selain itu, pembelajaran online yang bersifat virtual memiliki masalah yang melekat dan menghadirkan hambatan pembelajaran yang perlu diatasi oleh siswa[11]. Oleh karena itu, penting untuk melakukan penelitian untuk mengatasi tantangan tersebut dan meningkatkan efektivitas pendidikan pembelajaran berbasis virtual. Penelitian dapat membantu mengidentifikasi strategi pengajaran dan pembelajaran yang efektif yang dapat digunakan dalam lingkungan pembelajaran berbasis virtual, serta cara untuk meningkatkan akses pembelajaran berbasis virtual untuk semua siswa, terlepas dari lokasi atau latar belakang mereka. Dengan melakukan penelitian, kita dapat meningkatkan kualitas pendidikan pembelajaran berbasis virtual dan memastikan bahwa siswa memiliki akses ke pendidikan berkualitas tinggi yang sesuai dengan kebutuhan dan gaya belajar mereka.

Berbagai peneliti sebelumnya telah melakukan identifikasi gaya belajar siswa diantaranya Sianturi et al(2022) tentang Analisis Perilaku Siswa Untuk Mendeteksi Gaya Belajar Menggunakan Metode *Decision Tree*, *Naïve Bayes*, dan *K-Nearest Neighbor* Pada LMS Moodle[2], penelitian ini menyajikan sistem deteksi gaya belajar berdasarkan perilaku belajar di LMS Sekolah Vokasi Kristen Petra Surabaya untuk mata pelajaran Administrasi Sistem Jaringan menggunakan tiga metode: *Decision Tree*, *Naïve Bayes*, dan *K-Nearest Neighbor*. Studi ini menunjukkan bahwa metode *Decision Tree* dapat mendeteksi dan memprediksi gaya belajar dengan akurasi tinggi dengan lebih baik. Akurasi yang diperoleh dengan menggunakan uji split kereta 80:20 adalah 0,96, dan uji *K-Fold 10 Cross-Validation* memperoleh akurasi 0,98. Waktu pemrosesan untuk metode *Decision Tree* adalah 0,000998 detik untuk uji split kereta 80:20 dan 0,04033 detik untuk uji *K-Fold 10 Cross-Validation*. Kemudian Mehenaoui et al(2022)[3] tentang Analisis Perilaku Belajar untuk Mengidentifikasi Gaya Belajar Peserta Didik Berdasarkan Teknik *Machine Learning*, makalah ini menyajikan hasil percobaan yang dilakukan untuk mengidentifikasi gaya belajar peserta didik secara otomatis berdasarkan pola perilaku belajar dalam lingkungan belajar online. Eksperimen menggunakan enam pengklasifikasi yang berbeda untuk mendeteksi gaya belajar, dan hasilnya dievaluasi menggunakan metrik *recall*, presisi, akurasi, dan *F1-score*. Temuan menunjukkan bahwa adalah mungkin untuk mendeteksi gaya belajar secara otomatis dengan akurasi tinggi menggunakan teknik pembelajaran mesin. *Support Vector Machines (SVM)* menunjukkan kemampuan hebat dalam memprediksi gaya belajar untuk semua dimensi Model Gaya Pembelajaran *Felder dan Silverman (FSLSM)* dengan akurasi rata-rata 88%. Kemudian Hasibuat et al(2022) tentang Deteksi gaya belajar dengan data pengetahuan awal menggunakan algoritma SVM, K-NN dan *Naïve Bayes*[15], hasil penelitian ini menunjukkan bahwa algoritma yang diusulkan untuk mendeteksi pola pembelajaran otonom menggunakan pengetahuan sebelumnya memiliki akurasi yang tinggi. Algoritma *Naïve Bayes* memiliki nilai akurasi tertinggi 91,48%, diikuti oleh K-NN dengan 89,39%, dan SVM dengan 87,31%. Kemudian penelitian Altamimi et al(2022) tentang prediksi gaya belajar siswa dengan menggunakan teknik regresi[16], paper ini mengusulkan pendekatan probabilistik untuk mendeteksi gaya belajar yang disukai siswa menggunakan analisis regresi. Studi ini membandingkan model regresi dan model klasifikasi untuk mendeteksi gaya belajar dan menyimpulkan bahwa teknik regresi lebih akurat dan representatif untuk skenario dunia nyata. Hasil penelitian menunjukkan bahwa pendekatan yang diusulkan dapat secara efektif mendeteksi gaya belajar siswa, bahkan ketika mereka tidak memiliki gaya dominan atau campuran gaya belajar. Studi ini dapat membantu lembaga pendidikan untuk melibatkan gaya belajar dalam proses pengajaran. Penelitian oleh Rasshed et al(2021) tentang Deteksi gaya belajar dalam sistem *E-learning* menggunakan teknik pembelajaran mesin[17], paper ini mengumpulkan kumpulan data 498 sampel dari siswa yang mengambil kursus

online dan mengidentifikasi atribut baru untuk mendeteksi gaya belajar pelajar. Berbagai algoritma klasifikasi diterapkan pada kumpulan data, dan akurasi dibandingkan. Para penulis mengamati pola yang menarik dalam perilaku pelajar sambil mempelajari berbagai jenis konsep dalam situasi yang berbeda.

Penelitian oleh Ramírez-Correa, et al(2021) tentang Mengidentifikasi Profil Gaya Belajar Sarjana Teknik Menggunakan Teknik *Machine Learning*[18], Makalah ini bertujuan untuk menghasilkan model prediktif untuk dengan cepat mendeteksi profil gaya belajar sarjana di lingkungan belajar universitas hibrida. Para peneliti menggunakan algoritma pengelompokan *K-means* untuk mengidentifikasi profil gaya belajar siswa dan algoritma C4.5 pohon keputusan untuk memprediksi keanggotaan siswa ke kelompok yang diidentifikasi sebelumnya. Hasil penelitian adalah model prediktif yang dengan sedikit pertanyaan, mendeteksi profil siswa dengan akurasi 82,93%. Prediksi ini memungkinkan penyesuaian cepat metode pengajaran dalam lingkungan belajar hibrida. Namun, dua keterbatasan penelitian ini harus dinyatakan. Pertama, analisis dilakukan pada sampel yang relatif kecil dalam satu unit, yang tidak secara langsung memperkirakan hasilnya. Kedua, serangkaian atribut siswa yang terbatas digunakan untuk memprediksi profil gaya belajar mereka, dan, sebagai hasilnya, atribut lain dapat memprediksi profil ini dengan akurasi yang lebih besar. Penelitian oleh Sani Ibrahim et al(2020) tentang Deteksi Gaya Belajar Menggunakan *K-Means Clustering*[19], paper ini mengevaluasi kinerja pengelompokan K-means dalam mendeteksi gaya belajar peserta didik dalam lingkungan belajar online. Hasil eksperimen menunjukkan bahwa ada perbedaan dalam pembelajaran, yang dapat digunakan untuk mengkarakterisasi peserta didik berdasarkan gaya belajar mereka. Penelitian Hmedna et al(2020) tentang Model prediktif untuk mengidentifikasi gaya belajar di lingkungan MOOC[20], Makalah ini menganalisis data dari kursus edX “pembelajaran statistik” untuk memprediksi gaya belajar peserta didik menggunakan teknik pembelajaran mesin. Hasil penelitian menunjukkan bahwa pohon keputusan berkinerja terbaik untuk memprediksi gaya belajar peserta didik dengan akurasi lebih tinggi dari 98% dan penurunan risiko *overfitting* data pelatihan. Studi ini juga menemukan bahwa peserta didik dengan preferensi kuat untuk gaya belajar tertentu berinteraksi dengan platform lebih dari peserta didik lainnya. Namun, jumlah peserta didik tertinggi dikelompokkan sebagai sangat lemah, yang dapat dijelaskan oleh tingginya tingkat putus sekolah yang diderita MOOC. Penelitian oleh Khan, et al(2019) tentang Pemodelan Siswa Otomatis untuk Deteksi Gaya Belajar dan Status Afektif dalam Sistem Manajemen Pembelajaran Berbasis Web[21], paper ini mengusulkan pendekatan otomatis untuk mengidentifikasi gaya belajar dan keadaan afektif dalam lingkungan belajar berbasis web. Pendekatan ini dievaluasi dan ditemukan sesuai untuk mengidentifikasi gaya belajar dan keadaan afektif. Alat yang disebut ASLim telah dikembangkan berdasarkan pendekatan ini, yang dapat digunakan oleh guru untuk mengidentifikasi gaya belajar dan keadaan afektif siswa mereka. Kemudian penelitian Ling Xiao, et al(2018) tentang Deteksi gaya belajar siswa menggunakan *tree augmented naive bayes*[4], hasil eksperimen makalah ini menunjukkan bahwa *tree augmented naive bayes* mengungguli *Bayesian network* yang umum dalam mendeteksi gaya belajar siswa di lingkungan belajar online.

Berdasarkan penelusuran pustaka yang dilakukan oleh penulis, terdapat gap penelitian bahwa masih sangat sedikit penelitian tentang deteksi gaya belajar siswa dalam konteks *Virtual based Learning Environment*(VBLE). GAP atau *research gap* pada penelitian ini dapat dijelaskan sebagai berikut: (1)Kurangnya penelitian tentang pendeteksian gaya belajar siswa pada VBLE : Meskipun penggunaan VBLE semakin meningkat, penelitian untuk mendeteksi gaya belajar siswa dalam VBLE masih minim. Oleh karena itu, penelitian ini dapat memberikan kontribusi penting dalam mengisi kesenjangan tersebut. (2)Keterbatasan penggunaan metode klasifikasi Beberapa penelitian sebelumnya telah menggunakan metode klasifikasi untuk menganalisis gaya belajar siswa, namun masih terdapat keterbatasan dalam penggunaan metode tersebut. Oleh karena itu, penelitian ini dapat memperluas penggunaan metode klasifikasi dan membandingkan kinerja kedua metode klasifikasi yang berbeda tersebut. (3)Perkembangan VBLE yang selalu berubah: Teknologi VBLE terus berkembang dan berubah seiring waktu. Oleh karena itu, penelitian ini dapat membantu mengembangkan teknik pendeteksian gaya belajar siswa yang lebih efektif dan efisien yang dapat beradaptasi dengan perubahan teknologi VBLE di masa mendatang.

Penelitian ini bertujuan dua(2) hal, pertama mendeteksi gaya belajar siswa menengah atas swasta(SMAS) pada VBLE dengan algoritma *Decision Tree(DT) C4.5 dan Naive Bayes(NB)* dan kedua melakukan komparasi algoritma DT dan NB untuk konteks deteksi gaya belajar siswa menengah atas swasta pada VBLE. Pengumpulan data dilakukan dengan kuesioner *google form* pada 252 responden terdiri dari alumni dan siswa SMAS melalui 30 pertanyaan tentang gaya belajar. Gaya belajar siswa diidentifikasi menjadi gaya belajar visual, auditori dan kinestetik. Untuk analisa data digunakan *tool google colab* dengan pemrograman *python*. Sedangkan proses evaluasi performa algoritma diukur menggunakan parameter *accuracy, precision, recall/sensitivity, f-measure* dan *time consumse*.

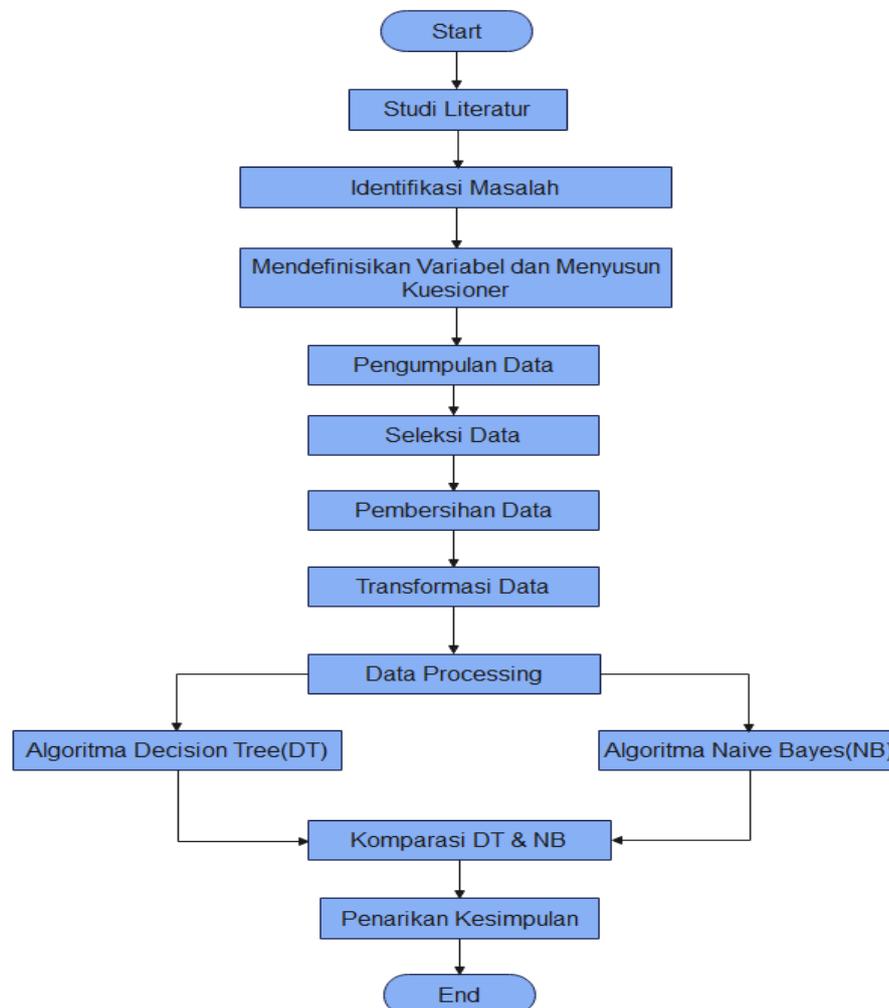
Penelitian tentang pendeteksian gaya belajar pada *Virtual Based Learning Environment* (VBLE) menggunakan metode klasifikasi seperti *Decision Tree C4.5 dan Naive Bayes* memiliki urgensi penting karena beberapa alasan berikut: (1)Peningkatan penggunaan VBLE: Selama pandemi COVID-19, penggunaan VBLE menjadi semakin penting untuk mengatasi keterbatasan pembelajaran di ruang kelas fisik. Oleh karena itu, pendeteksian gaya belajar siswa pada VBLE sangat penting untuk meningkatkan efektivitas pembelajaran pada lingkungan virtual. (2)Kompleksitas data: Data tentang gaya belajar siswa pada VBLE bisa sangat kompleks dan sulit dianalisis. Oleh karena itu, penggunaan metode klasifikasi seperti *Decision Tree C4.5 dan Naive Bayes* dapat membantu menganalisis data secara efektif dan menghasilkan hasil yang akurat. Dengan demikian, penelitian ini dapat memberikan kontribusi yang signifikan dalam meningkatkan efektivitas pembelajaran di lingkungan virtual dan memfasilitasi proses pembelajaran yang lebih efektif dan efisien bagi siswa.

Perbandingan atau perbandingan antara dua atau lebih metode klasifikasi seperti *Decision Tree C4.5* dan *Naive Bayes* untuk mendeteksi gaya belajar siswa dalam *Virtual Based Learning Environment (VBLE)* penting karena alasan berikut: (1) Mendapatkan hasil yang lebih akurat: Dengan membandingkan Dari kinerja dua atau lebih metode klasifikasi, kita dapat mengetahui metode mana yang memberikan hasil yang lebih akurat dalam mendeteksi gaya belajar siswa pada VBLE. (2) Meningkatkan interpretasi hasil: Dengan membandingkan beberapa metode klasifikasi, kita dapat membandingkan dan mengevaluasi kekuatan dan kelemahan masing-masing metode. Ini dapat membantu dalam interpretasi hasil dan memperkuat kesimpulan yang diambil dari analisis data. (3) Menghindari *overfitting* atau *underfitting*: Dalam beberapa kasus, satu metode klasifikasi mungkin cocok untuk satu set data, tetapi tidak cocok untuk set data lainnya. Dengan membandingkan beberapa metode klasifikasi, kita dapat meminimalkan risiko *overfitting* atau *underfitting* model kami dan memastikan bahwa metode klasifikasi yang dipilih cocok untuk dataset yang digunakan. Hasil yang ingin dicapai dari proses perbandingan ini adalah menentukan metode klasifikasi yang paling efektif dalam mendeteksi gaya belajar siswa pada VBLE, sehingga dapat membantu meningkatkan keefektifan pembelajaran pada lingkungan virtual.

Penelitian ini memberikan kontribusi dalam dua(2) hal, pertama secara saintifik dengan menguji algoritma DT dan NB untuk deteksi gaya belajar siswa SMAS pada konteks VBLE dan secara praktis memberikan rekomendasi pada *stakeholder* pendidikan yaitu guru, siswa, wali murid dan pemangku kebijakan pendidikan untuk lebih memberikan perhatian pada deteksi gaya belajar siswa.

## 2. METODOLOGI PENELITIAN

Tahapan penelitian yang telah dilakukan dalam deteksi gaya belajar siswa sekolah menengah atas dilakukan mengikuti langkah-langkah penelitian dalam gambar 1 sebagai berikut:



Gambar 1. Tahapan Penelitian

### a. Studi Literatur

Studi literatur atau kajian pustaka dilakukan dengan cara mempelajari paper penelitian terkait dengan identifikasi dan deteksi gaya belajar siswa melalui database jurnal portal garuda kemdikbud dengan kata kunci deteksi/prediksi gaya belajar siswa dan database jurnal yang terindeks di scopus melalui *software Publish or Perish(PoP)*.

**b. Identifikasi Masalah**

Pada tahap ini dilakukan dengan melakukan observasi di Sekolah Menengah Atas Swasta (SMAS) Muhammadiyah 1 Pamekasan untuk dilakukan identifikasi masalah yang terjadi di Sekolah Menengah Atas Swasta (SMAS) berkaitan dengan gaya belajar siswa. Berdasarkan observasi didapatkan hasil bahwa permasalahan yang teridentifikasi adalah belum adanya informasi tentang gaya belajar siswa, guru saat ini mengajar tanpa memperhatikan keunikan gaya belajar siswa yang mengakibatkan rendahnya motivasi siswa dalam pembelajaran berbasis VBLE.

**c. Mendefinisikan Variabel dan Menyusun Kuesioner**

Tahap ini adalah menentukan variabel dan penentuan instrumen penelitian yaitu dengan menggunakan kuesioner. Penyusunan kuesioner ini terdiri dari nama, jenis kelamin dan pertanyaan dari 3 (tiga) kategori atau jenis gaya belajar visual, auditori dan kinestetik. Dan masing-masing gaya belajar tersebut terdiri dari 30 pertanyaan. Total pertanyaan gaya belajar sebanyak 90 item. Atribut atau indikator gaya belajar yang digunakan dalam tabel 1 sebagai berikut:

**Tabel 1.** Atribut atau indikator gaya belajar

Gaya Belajar	Kode	Indikator	Jumlah Item	No Item
Visual	A	Menggunakan visual untuk belajar	7	A1,A2,A3,A4,A5,A6,A7
	B	Memahami posisi yang baik, geometris, numerik, warna, dan rasa	7	B1,B2,B3,B4,B5,B6,B7
	C	Teratur dan Rapi	7	C1,C2,C3,C4,C5,C6,C7
	D	Tidak menyukai kebisingan	5	D1,D2,D3,D4,D5
	E	Instruksi verbal sulit diterima	4	E1,E2,E3,E4
Auditori	F	Mendengar merupakan cara belajar	7	F1,F2,F3,F4,F5,F6,F7
	G	Aktivitas lisan baik	5	F1,F2,F3,F4,F5
	H	Mempunyai kepekaan musik	7	H1,H2,H3,H4,H5,H6,H7
	I	Tidak terganggu dengan kebisingan	5	I1,I2,I3,I4,I5
	J	Aktivitas visual lemah	6	J1,J2,J3,J4,J5,J6
Kinestetik	K	Aktivitas fisik merupakan cara belajar	5	K1,K2,K3,K4,K5
	L	Peka terhadap ekspresi dan bahasa tubuh	6	L1,L2,L3,L4,L5,L6
	M	Banyak gerak dan fokus pada fisik	6	M1,M2,M3,M4,M5,M6
	N	Kurang rapi dan selalu ingin mencoba	6	N1,N2,N3,N4,J5,J6
	O	Aktivitas verbal lemah	7	O1,O2,O3,O4,O5,O6,O7
<b>Jumlah</b>			<b>90</b>	

**d. Pengumpulan Data**

Proses pengumpulan dilakukan dengan cara mengirim link dari *google form* yang berisi *questioner* kepada responden yaitu seluruh siswa Sekolah Menengah Atas Swasta (SMAS) Muhammadiyah 1 Pamekasan dari kelas X sampai dengan XII tahun ajaran 2022/2023 dan alumni tahun 2019/2020 sampai dengan 2021/2022 melalui media sosial sekaligus melakukan rekapitulasi hasil *questioner* dari *google form*. Penulis mengumpulkan informasi hasil kuesioner dari *google form* sebanyak 252 kuesioner yang diisi oleh siswa kelas X, XI dan XII dan alumni Sekolah Menengah Atas Swasta (SMAS) Muhammadiyah 1 Pamekasan. Atribut yang digunakan peneliti sebanyak 91 atribut atau indikator yang terdiri dari 1 (satu) atribut jenis kelamin dan 90 item tentang pertanyaan seputar gaya belajar siswa.

**e. Seleksi Data**

Pada tahap ini dilakukan pemilihan data hasil dari jawaban di kuesioner. Dalam hal ini akan dilakukan penghapusan kolom *timestamp* dan nama siswa. Tahap ini akan dipilih atribut yang akan dijadikan data dalam penelitian ini. Atribut yang digunakan peneliti sebanyak 92 atribut atau indikator yang terdiri dari 1 (satu) atribut jenis kelamin dan 90 item tentang pertanyaan seputar gaya belajar siswa serta 1 (satu) atribut output yang bernama Gaya belajar yang terdiri dari 3 (tiga) opsi yaitu visual, auditori, kinestetik.

**f. Pembersihan Data**

Tahapan pembersihan data dilakukan dengan penghapusan data duplikat, pengosongan data, penghilangan *noise* data. Pada penelitian ini, penghapusan data dilakukan apabila salah satu atribut dibiarkan kosong atau terdapat *noise* pada data. Data yang sudah bersih akan dilakukan pemisahan yaitu *data training* dan *data testing*. Tujuan dari proses pembersihan data adalah untuk menghilangkan duplikasi data, mengidentifikasi data yang tidak konsisten, dan memperbaiki kesalahan pada data, seperti kesalahan ketik dan sebagainya agar siap diproses.

**g. Transformasi Data**

Prosedur transformasi data dilakukan dengan mengubah atribut menjadi format yang dapat dipahami oleh aplikasi. Untuk menerima label numerik, atribut merek, jenis, dan bahan mengubah nilai atribut dari diskrit menjadi numerik (label) pada program. Sebelum dimasukkan ke pemrograman *python*, data yang semula berbentuk *Microsoft Excel* diubah terlebih dahulu ke bentuk *csv* supaya bisa dipahami oleh pemrograman *python* di *google collab*.

**h. Data Processing**

Data processing dilakukan dengan membagi data menjadi data *training* dan data *testing* dengan ketentuan 80% data *training* dan 20% data *testing* maka dari 252 siswa, data latih(*training*) sebanyak 201 data dan data uji(*testing*) sebanyak 51 data. Pada penelitian ini menggunakan 2(dua) metode yaitu *Decision Tree C4.5* dan *Naive Bayes*. Data yang sudah dibersihkan akan diolah menggunakan pemrograman python di *google colab*.

i. **Komparasi Performa Algoritma *Decision Tree(DT)* dan *Naive Bayes(NB)***

Komparasi performa atau kinerja algoritma DT dan NB dilakukan dengan menghitung dan membandingkan nilai *Accuracy, Precision, Recall/Sensitivity, f-measure* dan *time consumse*.

j. **Penarikan Kesimpulan**

Penarikan kesimpulan dilakukan untuk menjawab pertanyaan penelitian yaitu deteksi gaya belajar siswa dan membandingkan kinerja algoritma DT dan NB untuk deteksi gaya belajar siswa sekolah menengah atas swasta(SMAS).

### 3. HASIL DAN PEMBAHASAN

Hasil penelitian berdasarkan tahapan penelitian sebagaimana dijelaskan pada bagian metodologi penelitian adalah sebagai berikut:

#### 3.1 Hasil *Preprocessing Data*

##### 3.1.1 Pelabelan Data (*Data Labelling*)

Pelabelan data dilakukan dengan memberi penamaan pada output gaya belajar. Dalam memberikan nama atau menentukan gaya belajar masing-masing menggunakan “Ya“ terbanyak dari masing-masing atribut atau indikator, hasilnya seperti tabel 2 dibawah ini :

**Tabel 2.** Pelabelan untuk Penentuan Gaya Belajar Siswa

No.	Nama Siswa	Jenis Kelamin	Visual		Auditori		Kinestetik		Gaya Belajar
			Ya	Tidak	Ya	Tidak	Ya	Tidak	
1	ABD. KHALIK	Laki-laki	0	30	19	11	24	6	Kinestetik
2	ABDUS SYUKUR	Laki-laki	0	30	28	2	1	29	Auditori
3	ACHMAT FAISOL	Laki-laki	0	30	0	30	30	0	Kinestetik
4	ARIF AIMAN	Laki-laki	22	8	0	30	0	30	Visual
5	ARSY AUGUSTIN ROHMANI	Perempuan	30	0	0	30	0	30	Visual
6	DENI APRILIANSYAH	Laki-laki	0	30	30	0	0	30	Auditori
7	FARHAN ROHIM	Laki-laki	21	9	0	30	0	30	Visual
8	HAMIMAH	Perempuan	0	30	29	1	0	30	Auditori
9	HERFI HANAFI	Laki-laki	0	30	1	29	18	12	Kinestetik
10	ISMATUL MAR'AH	Perempuan	23	7	0	30	0	30	Visual
..	..	..	..	..	..	..	..	..	..
240	HARUN ZADID	Laki-laki	21	9	0	30	0	30	Visual
241	HELMI HIDAYANTI	Perempuan	0	30	0	30	28	2	Kinestetik
242	ADAM WA'DU LAMAF'UULAA S	Laki-laki	0	30	28	2	0	30	Auditori
243	ILMA ZULFAINI	Perempuan	24	6	0	30	0	30	Visual
244	MOH. SYUFYAN ASSAURI	Laki-laki	0	30	29	1	7	23	Auditori
245	NOVI WARDIATUL JANNAH	Perempuan	21	9	1	29	0	30	Visual
246	NUR KHOMARIYAH	Perempuan	0	30	0	30	25	5	Kinestetik
247	OKTAVIA IRANI PUTRI	Perempuan	29	1	1	29	0	30	Visual
248	PUTRI AYU WAHYUNI	Perempuan	0	30	30	0	0	30	Auditori
249	RANGGA RAMADAN	Laki-laki	20	10	0	30	0	30	Visual
250	SIANATUL LAYLA	Perempuan	0	30	30	0	0	30	Auditori
251	SUCI ZAIHDATUL THALIB	Perempuan	0	30	0	30	27	3	Kinestetik
252	ACH. TEGUH MIRZA LUTFULLAH	Laki-laki	0	30	0	30	26	4	Kinestetik

*Pie chart* gaya belajar siswa yang dipilih oleh responden melalui hasil kuesioner di *google form* seperti dalam gambar 2 dibawah ini. Grafik ini memberikan deskripsi bahwa siswa MA Swasta memiliki gaya belajar yang terbanyak adalah Auditori(42%), Visual(36%) dan Kinestetik(22%).



Gambar 2. Grafik Gaya Belajar Siswa

### 3.1.2 Pembersihan Data (Data Cleaning)

Pembersihan data dilakukan terhadap 252 data yang diperoleh. Pembersihan data dilakukan untuk menghilangkan duplikasi data, mengidentifikasi data yang tidak konsisten dan memperbaiki kekurangan pada data, seperti kesalahan ketik. Karena atribut khusus gaya belajar yang digunakan sebanyak 15 atribut atau 90 item maka untuk memudahkan dalam mentransformasi data ke *python* maka atribut tentang gaya belajar diubah menjadi kode item dari A1 sampai dengan O7. Berikut ini tabel 3 yang berisi data yang sudah dilakukan pembersihan :

Tabel 3. Data penelitian setelah melalui proses pembersihan data

Timestamp	Nama siswa	Jenis Kelamin	A1	A2	A3	....	O5	O6	O7	Gaya Belajar
13/03/2023 20:20:54	ABD. KHALIK	Laki-laki	Tidak	Tidak	Tidak	....	Tidak	Ya	Ya	Kinestetik
13/03/2023 20:29:13	ABDUS SYUKUR	Laki-laki	Tidak	Tidak	Tidak	....	Tidak	Tidak	Tidak	Auditori
13/03/2023 20:32:09	ACHMAT FAISOL	Laki-laki	Tidak	Tidak	Tidak	....	Ya	Ya	Ya	Kinestetik
13/03/2023 20:35:24	ARIF AIMAN	Laki-laki	Ya	Ya	Ya	....	Tidak	Tidak	Tidak	Visual
13/03/2023 20:38:37	ARSY AUGUSTIN ROHMANI	Perempuan	Ya	Ya	Ya	....	Tidak	Tidak	Tidak	Visual
.....	.....	.....	.....	.....	.....	....				
18/03/2023 13:07:32	PUTRI AYU WAHYUNI	Perempuan	Tidak	Tidak	Tidak	.....	Tidak	Tidak	Tidak	Visual
18/03/2023 13:09:22	RANGGA RAMADAN	Laki-laki	Ya	Ya	Ya	.....	Tidak	Tidak	Tidak	Auditori
18/03/2023 13:10:52	SIANATUL LAYLA	Perempuan	Tidak	Tidak	Tidak	.....	Tidak	Tidak	Tidak	Visual
18/03/2023 13:13:57	SUCI ZAIDHATUL THALIB	Perempuan	Tidak	Tidak	Tidak	.....	Tidak	Tidak	Tidak	Auditori
18/03/2023 13:16:07	ACH. TEGUH MIRZA LUTFULLAH	Laki-laki	Tidak	Tidak	Tidak	.....	Ya	Ya	Ya	Kinestetik

Setelah data sudah bersih maka selanjutnya dilakukan penghapusan kolom *timestamp* dan nama siswa sehingga hasilnya seperti pada tabel 4 dibawah ini :

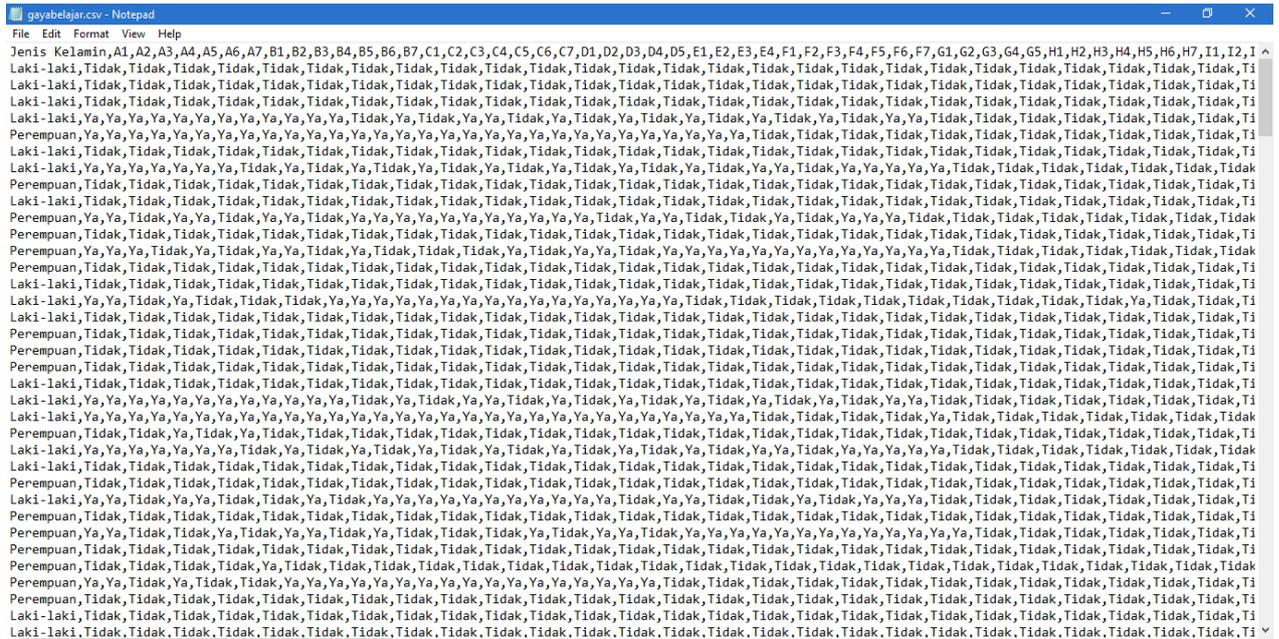
Tabel 4. Dataset Penelitian

Jenis Kelamin	A1	A2	A3	....	O5	O6	O7	Gaya Belajar
Laki-laki	Tidak	Tidak	Tidak	....	Tidak	Ya	Ya	Kinestetik
Laki-laki	Tidak	Tidak	Tidak	....	Tidak	Tidak	Tidak	Auditori
Laki-laki	Tidak	Tidak	Tidak	....	Ya	Ya	Ya	Kinestetik
Laki-laki	Ya	Ya	Ya	....	Tidak	Tidak	Tidak	Visual
Perempuan	Ya	Ya	Ya	....	Tidak	Tidak	Tidak	Visual

.....	.....	.....	.....	.....	Tidak	Tidak	Tidak	Visual
Perempuan	Tidak	Tidak	Tidak	.....	Tidak	Tidak	Tidak	Visual
Laki-laki	Ya	Ya	Ya	.....	Tidak	Tidak	Tidak	Auditori
Perempuan	Tidak	Tidak	Tidak	.....	Tidak	Tidak	Tidak	Visual
Perempuan	Tidak	Tidak	Tidak	.....	Tidak	Tidak	Tidak	Auditori
Laki-laki	Tidak	Tidak	Tidak	.....	Ya	Ya	Ya	Kinestetik

### 3.2 Hasil Transformasi Data

Proses Transformasi Data dilakukan dengan mengubah atribut menjadi format yang dapat dipahami oleh aplikasi. Untuk menerima label numerik, atribut merek, jenis, dan bahan mengubah nilai atribut dari diskrit menjadi numerik (label) pada program. Sebelum dimasukkan ke pemrograman *python*, data yang semula berbentuk *Microsoft Excel* diubah terlebih dahulu ke bentuk *csv* supaya bisa dipahami oleh pemrograman *python* di *google collab*. Jika dibuka melalui *notepad*, data yang sudah diubah akan berbentuk seperti gambar 3 dibawah ini :



Gambar 3. Data penelitian bentuk CSV

Dari data csv tersebut ditransform ke pemrograman *python*. Dibawah ini gambar 4 adalah *source code* tranformasi data csv ke pemrograman *python*:

```

data['Jenis Kelamin'] = enc.fit_transform(data['Jenis Kelamin'].values)
data['A1'] = enc.fit_transform(data['A1'].values)
data['A2'] = enc.fit_transform(data['A2'].values)
data['A3'] = enc.fit_transform(data['A3'].values)
data['A4'] = enc.fit_transform(data['A4'].values)
data['A5'] = enc.fit_transform(data['A5'].values)
data['A6'] = enc.fit_transform(data['A6'].values)
data['A7'] = enc.fit_transform(data['A7'].values)
data['B1'] = enc.fit_transform(data['B1'].values)
data['B2'] = enc.fit_transform(data['B2'].values)
data['B3'] = enc.fit_transform(data['B3'].values)
data['B4'] = enc.fit_transform(data['B4'].values)
data['B5'] = enc.fit_transform(data['B5'].values)
data['B6'] = enc.fit_transform(data['B6'].values)
data['B7'] = enc.fit_transform(data['B7'].values)
data['C1'] = enc.fit_transform(data['C1'].values)
data['C2'] = enc.fit_transform(data['C2'].values)
data['C3'] = enc.fit_transform(data['C3'].values)
data['N5'] = enc.fit_transform(data['N5'].values)
data['N6'] = enc.fit_transform(data['N6'].values)
data['O1'] = enc.fit_transform(data['O1'].values)
data['O2'] = enc.fit_transform(data['O2'].values)
data['O3'] = enc.fit_transform(data['O3'].values)
data['O4'] = enc.fit_transform(data['O4'].values)
data['O5'] = enc.fit_transform(data['O5'].values)
data['O6'] = enc.fit_transform(data['O6'].values)
data['O7'] = enc.fit_transform(data['O7'].values)
data['Gaya Belajar'] = enc.fit_transform(data['Gaya Belajar'].values)
    
```

Gambar 4. Data penelitian bentuk CSV

Dari hasil *coding* transformasi diatas maka akan muncul seperti gambar 5 dibawah ini :

**Gambar 5.** Hasil transformasi data

### 3.3 Deteksi Gaya Belajar Siswa dengan Algoritma *Decision Tree C4.5*

#### 3.3.1 Algoritma *Decision Tree*(DT) C.45

Metode C4.5 dapat digunakan untuk membuat pohon keputusan berdasarkan tabel diatas dengan terlebih dahulu menemukan *node root*. Rumus berikut harus digunakan untuk menentukan *entropy* sebelum menghitung *gain*.

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

Keterangan :

S = kasus yang ditentukan

A = Atribut

n = jumlah partisi S

pi= perbandingan Si dengan S

Setelah mendapatkan nilai *entropy* untuk setiap atribut, gunakan rumus di bawah ini untuk menentukan *gain ratio*:

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

Keterangan :

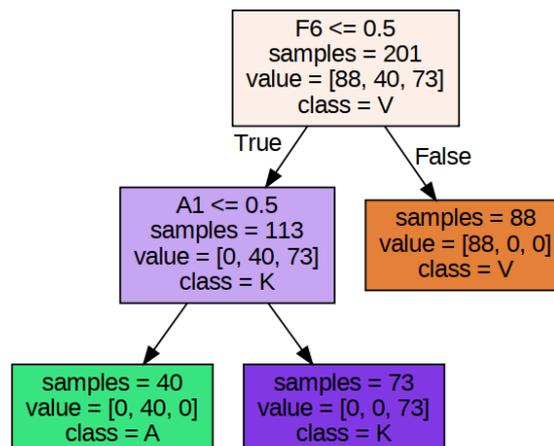
S = Kasus yang ditentukan

A = Atribut

n = Jumlah partisi S

pi= Perbandingan Si dengan S

Deteksi gaya belajar siswa dilakukan dalam penelitian ini dilakukan dengan menggunakan Pemrograman *Python* di *google collab*. Hasil dari proses deteksi gaya belajar siswa menghasilkan bentuk pola berupa pohon keputusan sebagaimana gambar 6 berikut:



**Gambar 6.** Pohon keputusan deteksi gaya belajar siswa dengan *Decision Tree C4.5*

Pada gambar diatas menjelaskan bahwa dari hasil pohon keputusan tersebut dapat disimpulkan, yang menjadi akar atau *node* 1 adalah atribut “mendengar merupakan cara belajar berbeda”, item “Saya berbicara pada diri saya sendiri ketika saya membaca penjelasan *online*” dengan kode item F6. Berdasarkan pilihan yang dibuat dengan menggunakan model klasifikasi, pohon keputusan telah menghasilkan peringkat nilai. Pada kasus ini  $F6 \leq 0,5$  merupakan sebuah kelas gaya belajar visual membuat sebuah keputusan bahwa setiap klasifikasi dari jumlah keseluruhan 252 siswa yaitu sebanyak 201 dihasilkan gaya belajar visual sebanyak 88 orang, gaya belajar auditori sebanyak 40 orang dan gaya belajar kinestetik sebanyak 73 orang.

Selanjutnya dari sampel = 201 tersebut dilakukan pemisahan yaitu jika bernilai *true* atau benar maka *node* kedua adalah atribut “menggunakan visual untuk belajar”, item “Saya selalu membaca instruksi sebelum melakukan latihan atau tugas *online*” dengan kode item A1. Pada *node* 2,  $A1 \leq 0,5$  ini masuk dalam kelas gaya belajar kinestetik dengan *value* = [0, 40, 73] memiliki arti bahwa dari 113 sampel diperoleh gaya belajar visual 0, gaya belajar auditori 40 dan gaya belajar kinestetik sebanyak 73 orang. Jika bernilai *false* maka sampelnya adalah 88 orang masuk dalam kelas visual dengan *value* = [88, 0, 0] memiliki arti gaya belajar visual sebanyak 88 namun tidak memiliki gaya belajar auditori dan kinestetik.

Dari nilai *true* tersebut itu dilakukan pemisahan lagi dengan rincian sebagai berikut :

- Pada *samples* 40 berada di kelas auditori dengan *value* = [0, 40, 0] memiliki arti gaya belajarnya hanya auditori saja sebanyak 40 orang dan tidak memiliki gaya belajar visual dan kinestetik
- Pada *samples* 73 berada di kelas kinestetik dengan *value* = [0, 0, 73] memiliki arti gaya belajarnya hanya kinestetik saja sebanyak 73 orang dan tidak memiliki gaya belajar visual dan auditori.

### 3.4 Deteksi Gaya Belajar Siswa dengan Algoritma Naive Bayes

Teorema Bayes adalah dasar dari *Naive Bayes Classifier*, sebuah teknik klasifikasi. Di bawah ini adalah rumus Algoritma *Naive Bayes* :

$$P(H|X)^{\wedge} = \frac{P(X|H)}{P(X)} x P(H) \quad (3)$$

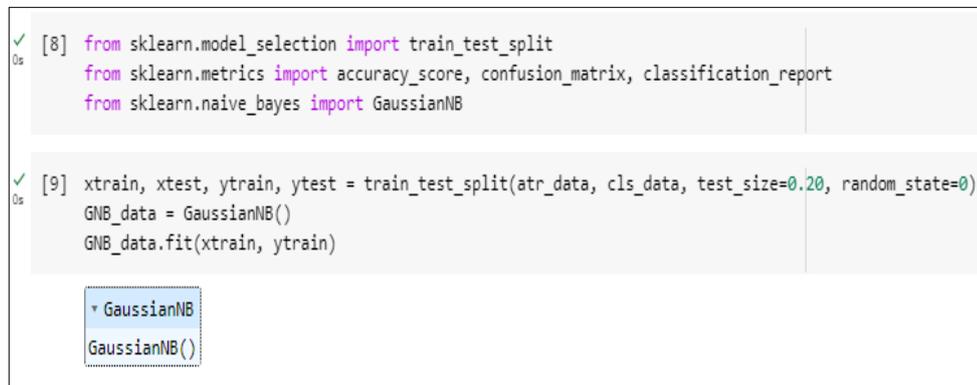
Keterangan :

- X : Data dengan *class* yang belum diketahui
- H : Kelas hipotesis data bersifat spesifik
- $P(H|X)^{\wedge}$  : Probabilitas hipotesis H berdasar kondisi X
- P(H) : Probabilitas hipotesis H (prior probabilitas)
- P(X|H) : Probabilitas X berdasarkan kondisi pada hipotesis H
- P(X) : Probabilitas X

*Source code* dibawah ini merupakan pengolahan data pemrograman *python* di *google collab* beserta pembagian data dalam gambar 7 menggunakan metode *Naive Bayes* :

```
[8] from sklearn.model_selection import train_test_split
    from sklearn.metrics import accuracy_score, confusion_matrix, classification_report
    from sklearn.naive_bayes import GaussianNB

[9] xtrain, xtest, ytrain, ytest = train_test_split(atr_data, cls_data, test_size=0.20, random_state=0)
    GNB_data = GaussianNB()
    GNB_data.fit(xtrain, ytrain)
```



Gambar 7. Source Code Naive Bayes

Untuk mendapatkan tingkat akurasi pada metode ini menggunakan *source code* dibawah ini :

```
akurasi = classification_report(ytest,y_pred)
akurasi = accuracy_score(ytest,y_pred)
print("Tingkat Akurasi : %dpersen" %(akurasi*100))
```

Tingkat Akurasi : 98persen

Gambar 8. Source Code Tingkat Akurasi Naive Bayes

Dari pengolahan data menggunakan metode *Naive Bayes* dihasilkan tingkat akurasi sebesar 98%.

### 3.5 Perbandingan Performa Algoritma Decision Tree C4.5 dan Naive Bayes

Membandingkan performa atau kinerja algoritma *Decision Tree*(DT) C.45 dan *Naive Bayes*(NB) untuk mendeteksi gaya belajar siswa dilakukan dengan menggunakan *tool google collab* dengan bahasa pemrograman *python*. Performa algoritma DT dan NB diukur dengan menggunakan lima(5) parameter yaitu (1)*Accuracy*, (2)*Precision*, (3)*Recall/Sensitivity*, (4)*f-measure* dan (5)*Time Consumse*. Jika parameter tersebut dilakukan perhitungan manual maka untuk menghitungnya seperti dibawah ini :

a. *Accuracy*

Untuk proses uji model yang dilakukan pada penelitian ini menggunakan *confusion matrix* 3 x 3 seperti dalam tabel 5. dibawah ini :

**Tabel 5.** *Confusion matrix* 3 x 3

<i>Actual</i>	<i>Predict class</i>		
	<i>Class A</i>	<i>Class B</i>	<i>Class C</i>
<i>Class A</i>	AA	AB	AC
<i>Class B</i>	BA	BB	BC
<i>Class C</i>	CA	CB	CC

$$Accuracy = \frac{AA + BB + CC}{AA + AB + AC + BA + BB + BC + CA + CB + CC} \times 100\% \quad (4)$$

b. *Precision*

$$Precision = \frac{TP}{TP + FP} \quad (5)$$

c. *Recall/Sensitivity*

$$Recall = \frac{TP}{TP + FN} \quad (6)$$

d. *f-measure*

$$F - Measure = 2 \times \left( \frac{Precision \times Recall}{Precision + Recall} \right) \quad (7)$$

Informasi:

TP : Jumlah record positif yang diklasifikasikan positif

FP : jumlah record negatif yang diklasifikasikan positif

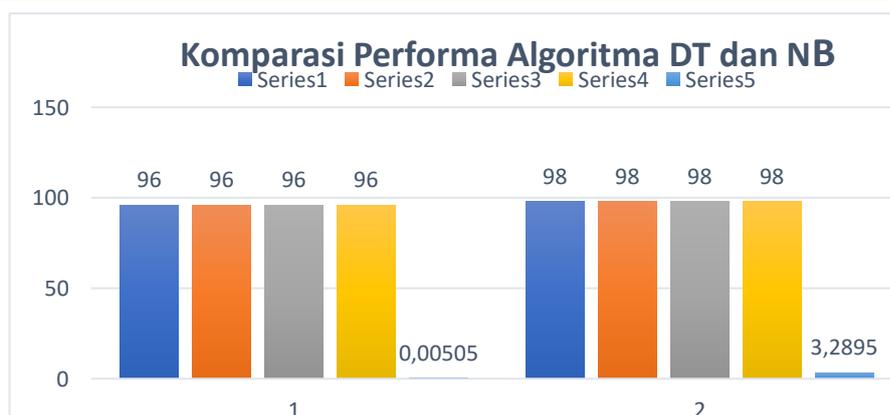
FN : jumlah record positif yang diklasifikasikan negatif

TN : jumlah record negatif yang diklasifikasikan negatif

Berdasarkan perhitungan performa algoritma DT dan NB menggunakan pemrograman *python* di *google collab* maka didapatkan hasil tabel 6 sebagai berikut:

**Tabel 6.** Perbandingan performa algoritma *Decision Tree* C4.5 dan *Naive Bayes*

Performa Algoritma	<i>Decision Tree</i> C4.5	<i>Naive Bayes</i>	Satuan
<i>Accuracy</i>	96	98	%
<i>Precision</i>	96	98	%
<i>Recall/Sensitivity</i>	96	98	%
<i>f-measure</i>	96	98	%
<i>Time consumes</i>	0,00505	3,2895	Detik



**Gambar 9.** Grafik komparasi performa algoritma *Decision Tree* C4.5 dan *Naive Bayes*

Grafik dalam gambar 9 diatas yang membandingkan kinerja dua algoritma, *Decision Tree C4.5* dan *Naive Bayes*, dalam hal berbagai pengukuran atau metrik parameter. Tabel diatas menunjukkan nilai pengukuran ini untuk kedua algoritma, bersama dengan unit pengukurannya. Metrik pertama adalah akurasi, yang mengukur persentase *instance* yang diklasifikasikan dengan benar dari semua *instance*. Tabel menunjukkan bahwa kedua algoritma memiliki akurasi tinggi, dengan *Naive Bayes* berkinerja sedikit lebih baik daripada *Decision Tree C4.5* (98% vs 96%). Metrik kedua adalah presisi, yang mengukur persentase contoh positif yang diklasifikasikan dengan benar dari semua contoh yang diklasifikasikan sebagai positif. Sekali lagi, kedua algoritma memiliki presisi tinggi, dengan *Naive Bayes* berkinerja sedikit lebih baik daripada *Decision Tree C4.5* (98% vs 96%). Metrik ketiga adalah *recall* atau sensitivitas, yang mengukur persentase contoh positif yang diklasifikasikan dengan benar dari semua contoh positif aktual. Kedua algoritma memiliki nilai *recall* yang berbeda (96% untuk *Decision Tree C4.5* dan 98% untuk *Naive Bayes*). Metrik keempat adalah *f-measure* merupakan metrik evaluasi model yang menggabungkan presisi dan *recall* untuk menentukan kinerja model secara keseluruhan, dalam hal *f-measure* (96% untuk *Decision Tree C4.5* dan 98% untuk *Naive Bayes*). Metrik terakhir adalah waktu yang dikonsumsi oleh setiap algoritma untuk melakukan tugas klasifikasi. *Decision Tree C4.5* jauh lebih cepat daripada *Naive Bayes*, *Decision Tree C4.5* hanya membutuhkan 0,00505 detik dibandingkan dengan 3,2895 detik untuk *Naive Bayes*. Secara keseluruhan, tabel ini memberikan perbandingan yang jelas dan informatif dari kinerja kedua algoritma, memungkinkan pembaca untuk dengan cepat menilai kekuatan dan kelemahan mereka.

#### 4. KESIMPULAN

Penelitian ini bertujuan dua(2) hal yaitu (1) untuk mendeteksi gaya belajar siswa sekolah menengah atas swasta(SMAS) dengan algoritma *Decision Tree(DT) C4.5* dan *Naive Bayes(NB)* dan (2) melakukan komparasi performa algoritma DT dan NB untuk prediksi gaya belajar siswa pada konteks VBLE. Berdasarkan hasil penelitian yang telah dilakukan maka didapatkan kesimpulan bahwa untuk tujuan pertama algoritma DT dan NB sangat baik dalam memprediksi gaya belajar siswa sekolah menengah swasta(SMAS) terlihat dari akurasi yang sangat tinggi yaitu DT(96%) dan NB(98%). Untuk tujuan kedua didapatkan kesimpulan bahwa algoritma NB sedikit lebih baik daripada algoritma DT dengan hasil pengukuran parameter *Accuracy*, *Precision*, *Recall/sensitivity* dan *f-measure* untuk NB(98%) sedangkan untuk DT sedikit lebih rendah(96%). Akan tetapi kebalikannya untuk parameter *Time consumes* DT(0,00505 detik) jauh lebih cepat daripada NB(3,2895 detik). Dengan demikian kesimpulannya algoritma NB sedikit lebih baik daripada DT meskipun NB memerlukan waktu penyelesaiannya yang lebih lama dari DT.

#### REFERENCES

- [1] N. D. S. Chetty, L. Handayani, N. A. Binti Sahabudin, Z. Ali, N. Hamzah, and S. Kasim, "Learning styles and teaching styles determine students' academic performances," *IJERE*, vol. 8, no. 4, p. 610, Dec. 2019, doi: 10.11591/ijere.v8i4.20345.
- [2] S. T. Sianturi and U. L. Yuhana, "Student Behaviour Analysis To Detect Learning Styles Using Decision Tree, Naive Bayes, And K-Nearest Neighbor Method In Moodle Learning Management System," *JTS*, vol. 33, no. 2, p. 94, Aug. 2022, doi: 10.12962/j20882033.v33i2.13665.
- [3] Z. Mehenaoui, Y. Lafifi, and L. Zemmouri, "Learning Behavior Analysis to Identify Learner's Learning Style based on Machine Learning Techniques," *jucs*, vol. 28, no. 11, pp. 1193–1220, Nov. 2022, doi: 10.3897/jucs.81518.
- [4] L. X. Li and S. S. Abdul Rahman, "Students' learning style detection using tree augmented naive Bayes," *R. Soc. open sci.*, vol. 5, no. 7, p. 172108, Jul. 2018, doi: 10.1098/rsos.172108.
- [5] A. et all. Sopian, "A Study on Learning Styles among Arabic Language Students at Universiti Teknologi Mara, Malacca Campus," *IOSR-JHSS*, vol. 15, no. 4, pp. 38–40, 2013, doi: 10.9790/0837-1543840.
- [6] A. P. Gilakjani, "A Match or Mismatch Between Learning Styles of the Learners and Teaching Styles of the Teachers," *IJMCECS*, vol. 4, no. 11, pp. 51–60, Dec. 2012, doi: 10.5815/ijmecs.2012.11.05.
- [7] D. Hernández-Torrano, S. Ali, and C.-K. Chan, "First year medical students' learning style preferences and their correlation with performance in different subjects within the medical course," *BMC Med Educ*, vol. 17, no. 1, p. 131, Dec. 2017, doi: 10.1186/s12909-017-0965-5.
- [8] A.-M. L. Villajuan, "Relationship between Learning Styles & Academic Achievement in Mathematics of Grade 8 Students," *IJELS*, vol. 4, no. 4, pp. 1052–1055, 2019, doi: 10.22161/ijels.4419.
- [9] R. Nizami, M. Z. Latif, and G. Wajid, "Preferred Learning Styles of Medical and Physiotherapy Students," *Annals KEMU*, vol. 23, no. 1, Mar. 2017, doi: 10.21649/akemu.v23i1.1511.
- [10] B. L. Garton, A. L. Ball, and J. E. Dyer, "The Academic Performance And Retention Of College Of Agriculture Students," *JAE*, vol. 43, no. 1, pp. 46–56, Mar. 2002, doi: 10.5032/jae.2002.01046.
- [11] F. Bell, "Connectivism: Its place in theory-informed research and innovation in technology-enabled learning," *IRRODL*, vol. 12, no. 3, p. 98, Mar. 2011, doi: 10.19173/irrodl.v12i3.902.
- [12] A. Sypsas, E. Paxinou, and D. Kalles, "Reviewing inquiry-based learning approaches in virtual laboratory environment for science education," *ICODL*, vol. 10, no. 2A, p. 74, Feb. 2020, doi: 10.12681/icodl.2288.
- [13] J. Capacho, "Representative Model of the Learning Process in Virtual Spaces Supported by ICT," *Turkish Online Journal of Distance Education*, vol. 16, no. 1, Mar. 2015, doi: 10.17718/tojde.82918.
- [14] F. Moazami, E. Bahrapour, M. R. Azar, F. Jahedi, and M. Moattari, "Comparing two methods of education (virtual versus traditional) on learning of Iranian dental students: a post-test only design study," *BMC Med Educ*, vol. 14, no. 1, p. 45, Dec. 2014, doi: 10.1186/1472-6920-14-45.
- [15] M. S. Hasibuan and R. A. Aziz, "Detection of learning styles with prior knowledge data using the SVM, K-NN and Naive Bayes algorithms," *JINFOTEL*, vol. 14, no. 3, pp. 209–213, Aug. 2022, doi: 10.20895/infotel.v14i3.788.

- [16] A. M. Altamimi, M. Azzeh, and M. Albashayreh, "Predicting students' learning styles using regression techniques," *IJECS*, vol. 25, no. 2, p. 1177, Feb. 2022, doi: 10.11591/ijeecs.v25.i2.pp1177-1185.
- [17] F. Rasheed and A. Wahid, "Learning style detection in E-learning systems using machine learning techniques," *Expert Systems with Applications*, vol. 174, p. 114774, Jul. 2021, doi: 10.1016/j.eswa.2021.114774.
- [18] P. Ramírez-Correa, J. Alfaro-Pérez, and M. Gallardo, "Identifying Engineering Undergraduates' Learning Style Profiles Using Machine Learning Techniques," *Applied Sciences*, vol. 11, no. 22, p. 10505, Nov. 2021, doi: 10.3390/app112210505.
- [19] M. Sani Ibrahim, "LEARNING STYLE DETECTION USING K-MEANS CLUSTERING," *FJS*, vol. 4, no. 3, pp. 375–381, Sep. 2020, doi: 10.33003/fjs-2020-0403-351.
- [20] B. Hmedna, A. El Mezouary, and O. Baz, "A predictive model for the identification of learning styles in MOOC environments," *Cluster Comput.*, vol. 23, no. 2, pp. 1303–1328, Jun. 2020, doi: 10.1007/s10586-019-02992-4.
- [21] F. A. Khan, A. Akbar, M. Altaf, S. A. K. Tanoli, and A. Ahmad, "Automatic Student Modelling for Detection of Learning Styles and Affective States in Web Based Learning Management Systems," *IEEE Access*, vol. 7, pp. 128242–128262, 2019, doi: 10.1109/ACCESS.2019.2937178.