

Analisis Perbandingan Naïve Bayes dan Neural Network dalam Klasifikasi Minat Masyarakat pada Kursus Komputer

Nabila Syah Fitria, Sudi Suryadi, Fitri Aini Nasution*

Faculty of Science and Technology, Information System, Universitas Labuhanbatu, Rantauprapat, Indonesia

Email: ¹nabillasyahfitri784@gmail.com, ²sudisuryadi28@gmail.com, ^{3,*}fitriaininasution689@gmail.com

Correspondence Author Email: fitriaininasution689@gmail.com

Submitted: 15/02/2025; Accepted: 05/03/2025; Published: 07/03/2025

Abstrak—Dalam era digital, pemanfaatan teknologi dalam dunia pendidikan semakin berkembang, terutama dalam meningkatkan literasi digital masyarakat melalui kursus komputer. Untuk menganalisis minat masyarakat terhadap kursus, diperlukan pendekatan berbasis data mining yang dapat mengolah data dalam jumlah besar dan mengidentifikasi pola tertentu. Naïve Bayes dan Neural Network merupakan dua metode klasifikasi yang banyak digunakan, di mana Naïve Bayes bekerja berdasarkan probabilitas independen antar fitur, sedangkan Neural Network menggunakan jaringan saraf tiruan untuk menangkap pola yang lebih kompleks. Penelitian ini bertujuan untuk membandingkan kedua metode dalam mengklasifikasikan minat masyarakat terhadap LKP Ibay Komputer dan mengevaluasi akurasi masing-masing model. Hasil klasifikasi menunjukkan bahwa kedua metode menghasilkan prediksi yang sama, yaitu 53 data dikategorikan sebagai berminat dan 20 data sebagai tidak berminat. Akurasi model mencapai 100%, yang mengindikasikan performa klasifikasi yang sangat tinggi. Meskipun hasil ini tampak ideal, akurasi sempurna seperti ini sering kali memunculkan pertanyaan terkait validitas dan ketahanan model pada skenario dunia nyata. Faktor-faktor seperti ukuran dataset yang relatif kecil, pola data yang terlalu terstruktur, atau kurangnya variasi dalam data latih dapat menjadi penyebab hasil yang tampak terlalu baik. Oleh karena itu, penting untuk melakukan evaluasi tambahan seperti validasi silang atau pengujian pada dataset berbeda untuk memastikan model tidak mengalami overfitting dan tetap andal dalam prediksi yang lebih luas. Dengan hasil ini, dapat disimpulkan bahwa baik Naïve Bayes maupun Neural Network memiliki performa optimal dalam klasifikasi minat masyarakat terhadap kursus komputer, namun pemilihan metode dapat disesuaikan dengan kebutuhan, di mana Naïve Bayes unggul dalam efisiensi komputasi, sedangkan Neural Network lebih adaptif terhadap data yang lebih kompleks.

Kata Kunci: Data Mining; Naïve Bayes; Neural Network; Klasifikasi; Evaluasi Model

Abstract—In the digital era, the use of technology in education is growing, especially in improving people's digital literacy through computer courses. To analyze people's interest in courses, a data mining-based approach is needed that can process large amounts of data and identify certain patterns. Naïve Bayes and Neural Network are two widely used classification methods, where Naïve Bayes works based on independent probabilities between features, while Neural Network uses artificial neural networks to capture more complex patterns. This study aims to compare the two methods in classifying people's interest in LKP Ibay Komputer and evaluate the accuracy of each model. The classification results show that both methods produce the same predictions, namely 53 data are categorized as interested and 20 data as not interested. The model accuracy reaches 100%, indicating very high classification performance. Although these results seem ideal, perfect accuracy like this often raises questions regarding the validity and robustness of the model in real-world scenarios. Factors such as relatively small dataset sizes, overly structured data patterns, or lack of variation in training data can cause results that appear too good. Therefore, it is important to conduct additional evaluations such as cross-validation or testing on different datasets to ensure that the model does not experience overfitting and remains reliable in broader predictions. With these results, it can be concluded that both Naïve Bayes and Neural Networks have optimal performance in classifying people's interest in computer courses, but the choice of method can be adjusted according to needs, where Naïve Bayes excels in computational efficiency, while Neural Networks are more adaptive to more complex data.

Keywords: Data Mining; Naïve Bayes; Neural Network; Classification; Model Evaluation

1. PENDAHULUAN

Perkembangan teknologi semakin pesat dan telah merambah hampir semua aspek kehidupan, termasuk dunia kerja. Saat ini, berbagai pekerjaan yang sebelumnya dilakukan secara manual telah beralih ke sistem otomatis dan digital, meningkatkan efisiensi serta produktivitas. Kecerdasan buatan, otomatisasi, dan komputasi awan semakin banyak digunakan dalam berbagai sektor, mulai dari industri manufaktur hingga layanan pelanggan. Transformasi ini tidak hanya mempermudah pekerjaan manusia, tetapi juga menciptakan tantangan baru terkait adaptasi keterampilan dan keamanan data di era digital. Untuk saat ini, pengetahuan tentang teknologi sangat diperlukan karena hampir semua aspek kehidupan telah bertransformasi ke arah digital. Kemampuan dalam memahami dan menggunakan teknologi, baik dalam pekerjaan maupun kehidupan sehari-hari, menjadi keterampilan dasar yang wajib dimiliki. Perkembangan pesat di bidang kecerdasan buatan, komputasi awan, dan otomatisasi menuntut setiap individu untuk terus belajar dan beradaptasi agar tidak tertinggal. Selain itu, pemahaman teknologi juga membantu dalam meningkatkan efisiensi, memperluas peluang kerja, serta menjaga keamanan data di era digital yang semakin kompleks.

Belajar dan memahami teknologi sangat penting di era digital ini karena hampir semua aspek kehidupan bergantung pada kemajuan teknologi. Dengan menguasai teknologi, seseorang dapat meningkatkan produktivitas, efisiensi, dan daya saing di dunia kerja. Selain itu, pemahaman yang baik tentang teknologi juga membantu dalam mengatasi tantangan seperti keamanan data dan adaptasi terhadap perubahan industri. Oleh karena itu, terus belajar dan mengikuti perkembangan teknologi menjadi kunci untuk menghadapi masa depan yang semakin digital. Tanpa adanya belajar, sulit bagi seseorang untuk memahami perkembangan zaman yang terus berubah dengan cepat. Kemajuan teknologi, perubahan tren, dan inovasi di berbagai bidang menuntut individu untuk terus memperbarui

pengetahuan dan keterampilannya. Jika tidak, seseorang akan tertinggal dan kesulitan beradaptasi dengan tuntutan kehidupan modern. Oleh karena itu, belajar menjadi kunci utama agar dapat mengikuti perkembangan zaman dan menghadapi tantangan di masa depan dengan lebih siap.

Oleh karena itu, hadirnya kursus komputer memberikan solusi kepada masyarakat agar dapat memahami dan menguasai keterampilan teknologi yang semakin dibutuhkan di era digital. Dalam dunia yang semakin maju, kemampuan mengoperasikan komputer tidak lagi menjadi pilihan, tetapi kebutuhan yang mendukung berbagai bidang pekerjaan dan kehidupan sehari-hari. Kursus komputer memungkinkan individu untuk belajar mulai dari dasar-dasar penggunaan perangkat lunak, desain grafis, pemrograman, hingga analisis data yang kini menjadi keahlian penting di banyak industri. Selain meningkatkan peluang kerja, pelatihan ini juga membantu masyarakat dalam beradaptasi dengan perkembangan teknologi yang terus berubah. Dengan demikian, kursus komputer berperan penting dalam menciptakan sumber daya manusia yang lebih siap menghadapi tantangan era digital serta meningkatkan daya saing di pasar kerja yang semakin kompetitif.

Tetapi, adanya kursus komputer terkadang masih menjadi masalah bagi masyarakat, terutama karena beberapa faktor yang mempengaruhi kualitas pembelajaran. Salah satu kendala utama adalah ketidaksesuaian antara fasilitas yang diberikan dengan biaya kursus yang harus dibayar. Banyak tempat kursus yang menawarkan harga tinggi, namun tidak diimbangi dengan fasilitas yang memadai, seperti perangkat komputer yang sudah usang, jaringan internet yang lambat, atau kurangnya modul pembelajaran yang berkualitas. Selain itu, kompetensi pengajar juga sering menjadi kendala. Tidak semua instruktur memiliki pengalaman dan keahlian yang cukup dalam mengajar teknologi secara efektif, sehingga peserta kursus kesulitan memahami materi dengan baik. Kurangnya metode pengajaran yang interaktif dan minimnya dukungan bagi peserta dalam mempraktikkan langsung keterampilan yang dipelajari juga menjadi hambatan dalam proses pembelajaran.

Dalam perkembangan era digital saat ini, keterampilan komputer menjadi kebutuhan yang sangat penting bagi masyarakat. Banyak orang ingin meningkatkan kemampuannya melalui kursus komputer, namun kenyataan di lapangan menunjukkan bahwa kualitas kursus yang tersedia belum sepenuhnya memenuhi harapan. Pada penelitian [1] bahwasanya Pentingnya komputer dalam perkembangan teknologi informasi tidak dapat diabaikan, karena kemampuannya dalam memfasilitasi komunikasi, pencarian informasi, serta mendukung pendidikan melalui jaringan komputer dan internet yang memungkinkan akses tanpa batas terhadap sumber belajar berbasis multimedia, sehingga dunia pendidikan dan industri harus mampu mengikuti perkembangan IPTEK dengan meningkatkan pemahaman serta keterampilan dalam pengelolaan dan pemanfaatan teknologi komputer secara optimal. Beberapa tempat kursus masih menghadapi kendala dalam menyediakan fasilitas yang sesuai dengan biaya yang ditawarkan, sehingga peserta tidak mendapatkan pengalaman belajar yang optimal.

Oleh karena itu, perlunya dibuat sebuah penelitian tentang minat masyarakat pada kursus komputer, guna untuk mengetahui faktor-faktor yang mempengaruhi minat mereka dalam memilih dan mengikuti kursus komputer. Penelitian ini akan membantu mengidentifikasi aspek-aspek yang dianggap penting oleh masyarakat, seperti kualitas pengajaran, fasilitas yang disediakan, harga kursus, serta relevansi materi yang diajarkan dengan kebutuhan pasar kerja. Dengan memahami faktor-faktor tersebut, penyedia kursus dapat mengembangkan program yang lebih sesuai dengan harapan peserta dan meningkatkan daya tarik kursus komputer bagi masyarakat. Selain itu, hasil penelitian ini dapat memberikan wawasan bagi pengambil kebijakan, agar dapat merancang strategi pendidikan yang lebih efektif dalam meningkatkan keterampilan teknologi bagi masyarakat, terutama di era digital yang berkembang pesat.

Penelitian ini akan dilakukan pada machine learning dengan menggunakan metode Naive Bayes dan metode Neural Network, keduanya diterapkan dalam model klasifikasi untuk menganalisis minat masyarakat terhadap kursus komputer. Pada penelitian yang dilakukan oleh Sari et al., di tahun 2023 metode neural network [2] terbukti efektif dalam klasifikasi data pada penelitian, khususnya dalam menganalisis minat masyarakat terhadap Pertamina dengan proses pengolahan data yang sistematis dan akurat. Metode Naive Bayes, yang berbasis probabilitas, akan digunakan untuk mengklasifikasikan data berdasarkan distribusi fitur yang ada, sementara Neural Network, yang lebih kompleks, akan diterapkan untuk menangkap pola yang lebih mendalam dalam data, dengan mempertimbangkan hubungan non-linear antar fitur. Pada penelitian yang dilakukan oleh Siregar et al., di tahun 2023 metode Naive Bayes [3] terbukti efektif untuk klasifikasi data dengan menghitung probabilitas dari berbagai kombinasi atribut, sehingga dapat digunakan dalam menentukan kelayakan penerima bantuan sosial. Kedua metode ini diharapkan dapat memberikan hasil yang lebih komprehensif dan akurat dalam mengidentifikasi faktor-faktor yang mempengaruhi keputusan masyarakat dalam memilih kursus komputer. Melalui model klasifikasi ini, diharapkan dapat ditemukan pola yang jelas yang dapat membantu penyedia kursus untuk merancang program yang lebih menarik dan relevan dengan kebutuhan peserta. Pada penelitian yang dilakukan oleh Wahyudi et al., di tahun 2022 bahwasanya metode Naive Bayes [4] dalam klasifikasi ini digunakan untuk memprediksi kemungkinan kelakuan kartu internet baru di pasaran berdasarkan pola data historis penjualan yang telah teridentifikasi. Penelitian juga dilakukan pada penelitian Zarti et al., di tahun 2023 bahwasanya metode Naive Bayes [5] dalam klasifikasi pada penelitian ini digunakan untuk memprediksi partisipasi masyarakat pada Pemilu 2024 dengan menganalisis pola data historis dari Pemilu 2014 dan 2020, menghasilkan prediksi yang menunjukkan potensi peningkatan partisipasi. Untuk metode neural network terdapat pada penelitian yang dilakukan oleh Diana Dewi et al., di tahun 2023 bahwasanya Metode Artificial Neural Network (ANN) [6] digunakan dalam klasifikasi untuk mengenali pola kompleks antar atribut kesehatan, sehingga

mampu memprediksi dengan akurasi tinggi apakah seseorang menderita diabetes atau tidak. Pada penelitian ini untuk fokus analisis pada minat masyarakat terhadap kursus komputer yang belum banyak dibahas dalam penelitian sebelumnya. Penelitian terdahulu lebih banyak mengkaji prediksi minat terhadap produk bahan bakar, kelayakan bantuan sosial, partisipasi Pemilu, serta diagnosis penyakit, sementara penelitian ini secara khusus membandingkan performa Naïve Bayes dan Neural Network dalam klasifikasi minat masyarakat terhadap kursus komputer. Dengan demikian, penelitian ini menawarkan kontribusi baru dengan mengaplikasikan model klasifikasi pada sektor pendidikan, yang diharapkan dapat menjadi dasar pengembangan strategi promosi yang lebih efektif dan sesuai dengan kebutuhan pasar.

Metode Naive Bayes dan Neural Network dianggap tepat untuk analisis minat masyarakat karena keduanya memiliki keunggulan yang saling melengkapi dalam proses klasifikasi data. Naive Bayes, dengan pendekatannya yang sederhana dan berbasis probabilitas, sangat efektif untuk mengolah data dalam jumlah besar dengan cepat serta memberikan hasil yang mudah diinterpretasikan, terutama ketika fitur-fitur yang digunakan bersifat independen. Sementara itu, Neural Network mampu menangkap pola yang lebih kompleks dan hubungan non-linear antar fitur, sehingga lebih akurat dalam mengidentifikasi tren atau kecenderungan yang tersembunyi dalam data. Dengan mengombinasikan kedua metode ini, proses analisis menjadi lebih komprehensif: Naive Bayes memberikan dasar yang kuat untuk klasifikasi awal yang cepat dan efisien, sedangkan Neural Network memperdalam hasil analisis dengan menggali pola yang mungkin terlewatkan oleh metode probabilitas sederhana. Pendekatan gabungan ini tidak hanya meningkatkan akurasi, tetapi juga memberikan wawasan yang lebih kaya untuk memahami faktor-faktor yang mempengaruhi minat masyarakat terhadap kursus komputer, sehingga penyedia kursus dapat merancang strategi yang lebih tepat sasaran.

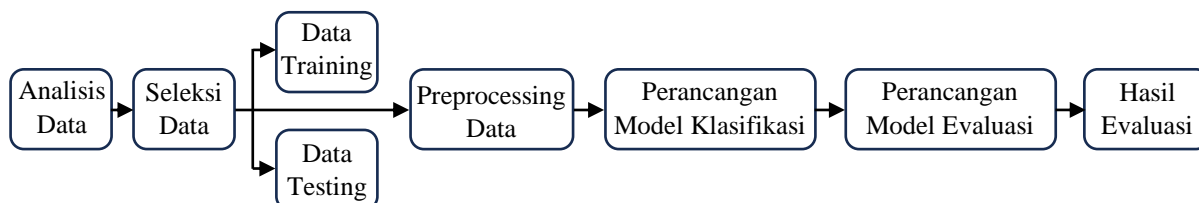
2. METODOLOGI PENELITIAN

2.1 Kajian Pustaka

Dalam penelitian ini, digunakan dua metode machine learning yang populer dalam klasifikasi, yaitu Naive Bayes dan Neural Network [7] [8] [9]. Metode Naive Bayes adalah algoritma klasifikasi berbasis probabilitas yang mengasumsikan bahwa setiap fitur dalam dataset bersifat independen, meskipun dalam praktiknya hal ini seringkali tidak terjadi [8] [10]. Meskipun demikian, Naive Bayes terbukti efektif dalam berbagai aplikasi klasifikasi, terutama pada dataset dengan dimensi besar dan variabel yang saling bergantung [11] [12] [13]. Sementara itu, Neural Network merupakan pendekatan yang lebih kompleks, terinspirasi dari cara kerja otak manusia, yang mampu menangkap pola dan hubungan non-linear antar fitur melalui lapisan-lapisan neuron [14] [15]. Neural Network sering digunakan untuk tugas-tugas yang membutuhkan model yang dapat mengatasi data yang sangat besar dan kompleks, dengan kemampuan untuk memodelkan hubungan yang lebih rumit dalam data [16]. Kedua metode ini dipilih karena keduanya memiliki keunggulan dalam menangani berbagai jenis data dan memberikan hasil yang optimal dalam tugas klasifikasi [17] [18].

2.2 Tahapan Penelitian

Penelitian ini bertujuan untuk mengklasifikasikan minat masyarakat terhadap kursus komputer di LKP Ibay Komputer, dengan fokus pada faktor-faktor yang mempengaruhi keputusan mereka. Beberapa tahapan yang akan dilakukan meliputi pengumpulan data dari masyarakat terkait demografi dan preferensi mereka terhadap kursus komputer, kemudian data tersebut akan dianalisis menggunakan metode machine learning, yaitu Naive Bayes dan Neural Network, untuk membangun model klasifikasi. Setelah model selesai, evaluasi akurasi akan dilakukan untuk mengukur efektivitas model dalam memprediksi minat masyarakat, yang diharapkan dapat memberikan wawasan bagi LKP Ibay Komputer dalam merancang program kursus yang lebih relevan.



Gambar 1. Tahapan Penelitian

Pada gambar 1. Merupakan tahapan penelitian yang digunakan untuk penelitian ini, berikut penjelasan dari tahapan diatas yaitu sebagai berikut.

- Analisis Data, Pada tahap ini, dilakukan identifikasi dan pemilihan atribut atau variabel yang akan digunakan dalam penelitian. Pemilihan variabel ini bertujuan untuk memastikan bahwa data yang digunakan relevan dan memiliki pengaruh terhadap klasifikasi minat masyarakat pada kursus komputer.

- b. Seleksi Data, Data yang digunakan dalam penelitian ini dikumpulkan dari berbagai sumber dan kemudian dipisahkan menjadi dua set, yaitu data training dan data testing. Data training digunakan untuk membangun model klasifikasi, sementara data testing digunakan untuk menguji kinerja model yang telah dibuat.
- c. Preprocessing Data, preprocessing data menjadi langkah krusial yang memastikan kualitas data sebelum digunakan dalam model klasifikasi. Pada penelitian ini, preprocessing mencakup beberapa teknik penting, seperti normalisasi data untuk menyamakan skala nilai sehingga algoritma klasifikasi dapat bekerja lebih optimal, pengisian nilai yang hilang (missing value imputation) agar tidak terjadi bias atau kehilangan informasi penting, serta encoding data kategorikal yang mengubah data non-numerik menjadi format angka yang dapat diproses oleh model. Proses ini bertujuan untuk menghilangkan noise, meningkatkan konsistensi, dan memastikan bahwa data yang diolah memiliki kualitas yang baik, sehingga hasil klasifikasi dengan metode Naive Bayes dan Neural Network menjadi lebih akurat dan dapat diandalkan dalam menganalisis minat masyarakat terhadap kursus komputer.
- d. Perancangan Model Klasifikasi, Model klasifikasi dibuat dengan menerapkan metode Naive Bayes dan Neural Network untuk mengolah data yang telah diproses. Proses klasifikasi ini bertujuan untuk mengelompokkan data berdasarkan tingkat minat masyarakat terhadap kursus komputer.
- e. Perancangan Model Evaluasi, Tahap ini bertujuan untuk menentukan metode evaluasi yang akan digunakan dalam mengukur kinerja model klasifikasi. Evaluasi dilakukan untuk menguji efektivitas metode yang digunakan dalam mengklasifikasikan minat masyarakat secara akurat.
- f. Hasil Evaluasi, Hasil evaluasi berupa metrik performa seperti akurasi, presisi, dan recall yang menunjukkan sejauh mana model dapat mengklasifikasikan data dengan baik. Dari hasil ini, dapat diketahui metode mana yang lebih efektif dalam memprediksi minat masyarakat terhadap kursus komputer.

2.3 Confusion Matrix

Confusion matrix digunakan untuk mengevaluasi kinerja metode klasifikasi dengan membandingkan hasil prediksi model terhadap data aktual [19] [20]. Dari confusion matrix, dapat dihitung metrik evaluasi seperti akurasi, presisi, dan recall, yang masing-masing menunjukkan tingkat ketepatan, keakuratan dalam mengidentifikasi kelas positif, serta kemampuan model dalam menangkap seluruh data yang relevan [21]. Dengan menggunakan confusion matrix, dapat diketahui sejauh mana metode yang digunakan efektif dalam mengklasifikasikan data secara akurat [22].

Tabel 1. Confusion Matrix

Kelas Aktual	Kelas Prediksi		
	Kelas Benar	Benar True Positive (TP)	Salah False Positive (FP)
	Salah	False Negative (FN)	True Negative (TN)

Pada tabel 1. berisi tentang kelas aktual dan prediksi pada Confusion Matrix, untuk penjelasannya yaitu sebagai berikut.

- a. TP (True Positive) adalah jumlah data yang sebenarnya positif dan berhasil diklasifikasikan dengan benar sebagai positif.
- b. TN (True Negative) adalah jumlah data yang sebenarnya negatif dan berhasil diklasifikasikan dengan benar sebagai negatif.
- c. FN (False Negative) adalah jumlah data yang sebenarnya positif tetapi diklasifikasikan secara keliru sebagai negatif.
- d. FP (False Positive) adalah jumlah data yang sebenarnya negatif tetapi diklasifikasikan secara keliru sebagai positif.

$$Accuracy = \frac{TP+TN}{TP+TN+FN+FP} \times 100\% \quad (1)$$

$$Presisi = \frac{TP}{TP+FP} \times 100\% \quad (2)$$

$$Recall = \frac{TP}{TP+FN} \times 100\% \quad (3)$$

3. HASIL DAN PEMBAHASAN

Bagian ini menguraikan hasil penelitian yang telah dilakukan serta analisis terhadap berbagai eksperimen yang dijalankan. Penelitian ini berfokus pada klasifikasi minat masyarakat terhadap kursus komputer di LKP Ibay Komputer dengan menerapkan metode Naive Bayes dan Neural Network. Seiring dengan perkembangan teknologi dalam analisis data dan machine learning, penerapan metode ini menjadi solusi yang tepat untuk mengidentifikasi faktor-faktor yang memengaruhi minat masyarakat secara lebih akurat.

Model klasifikasi yang dikembangkan bertujuan untuk mengelompokkan tingkat minat masyarakat terhadap kursus komputer berdasarkan data yang telah dikumpulkan. Dataset penelitian ini terdiri dari informasi responden yang mencakup faktor demografis dan preferensi mereka terhadap kursus komputer. Oleh karena itu, diperlukan teknik pengolahan data yang efektif, seperti preprocessing, pemilihan fitur yang relevan, serta pemisahan data menjadi

training dan testing. Tahapan penelitian meliputi proses pengumpulan data, pembersihan data, pembangunan model klasifikasi, serta pengujian performa model menggunakan berbagai metrik evaluasi.

Berdasarkan hasil penelitian, model yang dikembangkan mampu mengklasifikasikan minat masyarakat dengan tingkat akurasi yang baik. Akurasi tinggi yang diperoleh dalam proses pelatihan serta konsistensi hasil pada data uji menunjukkan bahwa metode yang digunakan memiliki kinerja yang optimal. Selama pelatihan, model dievaluasi menggunakan confusion matrix untuk mengukur akurasi, presisi, recall, dan skor F1. Evaluasi ini bertujuan untuk memastikan bahwa model tidak hanya memberikan hasil yang baik pada data pelatihan, tetapi juga memiliki kemampuan generalisasi yang baik terhadap data baru.

3.1 Analisis Data

Pada tahapan analisis data, dilakukan proses penentuan variabel yang akan digunakan dalam penelitian ini. Penentuan variabel ini bertujuan untuk memastikan bahwa data yang digunakan memiliki relevansi dan signifikansi dalam membangun model klasifikasi. Dengan menentukan variabel yang tepat, penelitian dapat menghasilkan analisis yang lebih akurat dan mendukung proses pengambilan keputusan dalam mengidentifikasi faktor-faktor yang memengaruhi minat masyarakat terhadap kursus komputer.

Tabel 2. Data Atribut Penelitian

Variabel	Tipe	Keterangan
Jarak Kursus	Kategori	Jarak kursus yang dekat dengan lokasi tempat tinggal siswa menjadi salah satu faktor yang dapat memengaruhi minat mereka untuk mengikuti kursus komputer di LKP Ibay Komputer.
Kompetensi Pengajar	Kategori	Kompetensi pengajar yang profesional dan berpengalaman menjadi daya tarik utama bagi siswa untuk memilih kursus komputer sebagai tempat belajar.
Ketersediaan waktu	Kategori	Ketersediaan waktu siswa yang fleksibel menjadi salah satu alasan mereka mempertimbangkan untuk mendaftar di kursus komputer yang sesuai dengan jadwal mereka.
Fasilitas	Kategori	Fasilitas yang lengkap dan modern di kursus komputer dapat meningkatkan kenyamanan siswa selama proses pembelajaran berlangsung.
Harga Kursus	Kategori	Harga kursus yang terjangkau menjadi faktor penting dalam menentukan keputusan siswa untuk bergabung dengan LKP Ibay Komputer.

Pada tabel 2. menampilkan data variabel yang digunakan dalam penelitian ini. Keberadaan variabel tersebut membantu mempermudah analisis serta memungkinkan penelitian ini mengidentifikasi faktor-faktor yang memengaruhi minat masyarakat terhadap kursus komputer.

3.2 Seleksi Data

Pada tahapan seleksi data merupakan tahapan yang dilakukan untuk memilih dan menentukan data yang akan digunakan pada penelitian ini. Untuk data yang akan digunakan pada penelitian ini terdapat 2 data set yang digunakan pada penelitian ini, data set yang pertama adalah data training dan data set yang kedua adalah data testing.

3.2.1. Data Training

Data training merupakan sekumpulan data latih yang digunakan sebagai fondasi dalam membangun, melatih, dan mengoptimalkan model klasifikasi agar mampu mengidentifikasi pola serta karakteristik yang terdapat dalam dataset. Melalui proses pembelajaran yang dilakukan pada data training, model dapat menyesuaikan bobot dan parameter yang diperlukan untuk meningkatkan akurasi serta efektivitas dalam melakukan prediksi. Dengan demikian, model yang telah terlatih memiliki kemampuan yang lebih baik dalam mengklasifikasikan data uji secara akurat, sehingga menghasilkan hasil analisis yang lebih valid dan reliabel.

Tabel 3. Data Training Penelitian

Nama	Jarak kursus	Kompetensi Pengajar	Ketersediaan Waktu	Fasilitas	Harga Kursus	Kategori
Abdur Rohim	Dekat	Baik	Ada	Tidak Lengkap	Mahal	Minat
Ahmad Azhar Tambak	Dekat	Baik	Ada	Lengkap	Murah	Minat
Aiyub Al-Anshari Simbolon	Dekat	Baik	Tidak Ada	Tidak Lengkap	Murah	Minat
Amelia Pradikta Dewi Siregar	Dekat	Baik	Ada	Lengkap	Murah	Minat
Andre Praditya	Jauh	Kurang Baik	Tidak Ada	Lengkap	Mahal	Tidak Minat
Dewi Rospita Sari	Dekat	Baik	Ada	Lengkap	Murah	Minat
Dewi Suriawan	Dekat	Baik	Tidak Ada	Lengkap	Murah	Minat
Eliana Siregar	Dekat	Baik	Ada	Lengkap	Mahal	Minat
Fadhil Ahmad Fahrezi	Jauh	Baik	Tidak Ada	Tidak Lengkap	Mahal	Tidak Minat
Faiza Fairuz Idna	Dekat	Kurang Baik	Ada	Tidak Lengkap	Murah	Minat
Febri Novatmawati	Dekat	Baik	Ada	Lengkap	Murah	Minat
Fitri Hardiyanti	Dekat	Baik	Ada	Lengkap	Murah	Minat
Helmi Zulfa Ilyasa	Jauh	Baik	Tidak Ada	Tidak Lengkap	Mahal	Tidak Minat
Holizah	Jauh	Kurang Baik	Tidak Ada	Tidak Lengkap	Mahal	Tidak Minat

Ifrah Rizki	Jauh	Baik	Ada	Lengkap	Mahal	Minat
Ilal Bahri	Jauh	Baik	Ada	Lengkap	Murah	Minat
Ilyatul Izzah Ramber	Jauh	Kurang Baik	Ada	Tidak Lengkap	Mahal	Tidak Minat
Indah Maya Sari Sidabutar	Dekat	Baik	Ada	Lengkap	Murah	Minat
Juni	Jauh	Kurang Baik	Tidak Ada	Tidak Lengkap	Mahal	Tidak Minat
Jusri Khan	Dekat	Kurang Baik	Ada	Lengkap	Murah	Minat
Lusiana Dina Apriyani	Dekat	Baik	Ada	Lengkap	Murah	Minat
Maikel Rinaldi Sihombing	Jauh	Kurang Baik	Tidak Ada	Tidak Lengkap	Mahal	Tidak Minat
Maysarah	Jauh	Kurang Baik	Tidak Ada	Tidak Lengkap	Murah	Tidak Minat
Melien Bunga Valentine	Dekat	Baik	Ada	Lengkap	Murah	Minat
Nandalia Devina	Dekat	Baik	Ada	Lengkap	Murah	Minat
Ocha Suci Puspita Sari	Jauh	Baik	Tidak Ada	Tidak Lengkap	Mahal	Tidak Minat
Pitriani	Jauh	Kurang Baik	Ada	Lengkap	Mahal	Tidak Minat
Putri Yani Siregar	Jauh	Baik	Ada	Lengkap	Murah	Minat
Raihan Permana	Dekat	Baik	Tidak Ada	Lengkap	Murah	Minat
Reissa Remysaura	Jauh	Kurang Baik	Tidak Ada	Tidak Lengkap	Mahal	Tidak Minat
Salsabila Fahira Hasibuan	Jauh	Kurang Baik	Tidak Ada	Tidak Lengkap	Mahal	Tidak Minat
Salsahidah Atira Nasution	Jauh	Kurang Baik	Tidak Ada	Tidak Lengkap	Mahal	Tidak Minat
Syhifa Salsabila	Dekat	Baik	Ada	Lengkap	Murah	Minat
Valentina Hutapea	Dekat	Baik	Ada	Lengkap	Murah	Minat
Windi Amelia	Dekat	Kurang Baik	Tidak Ada	Tidak Lengkap	Mahal	Tidak Minat
Yeni Wati Sihombing	Dekat	Baik	Ada	Lengkap	Murah	Minat
Yusuf Arifin Rambe	Dekat	Kurang Baik	Tidak Ada	Lengkap	Murah	Minat

Pada tabel 3. Menyajikan data training yang berperan sebagai dasar dalam membangun dan mengoptimalkan model klasifikasi yang digunakan dalam penelitian ini. Data training, yang terdiri dari 37 sampel, digunakan untuk melatih model agar mampu mengenali pola, struktur, dan karakteristik yang terkandung dalam dataset. Dengan adanya data training, model dapat menyesuaikan parameter dan meningkatkan akurasi, sehingga mampu melakukan klasifikasi dengan lebih efektif ketika diterapkan pada data uji. Tahap ini menjadi krusial dalam memastikan bahwa model yang dikembangkan dapat menghasilkan prediksi yang valid dan reliabel.

3.2.2. Data Testing

Data testing digunakan sebagai sampel dalam penelitian ini untuk menguji performa model klasifikasi. Dengan adanya data sampel, proses klasifikasi dapat dilakukan secara lebih akurat, sehingga hasil penelitian dapat menggambarkan pola minat masyarakat terhadap kursus komputer dengan lebih jelas. Untuk data testing yang digunakan dapat dilihat pada tabel dibawah ini.

Tabel 4. Data Testing Penelitian

Nama	Jarak kursus	Kompetensi Pengajar	Ketersediaan Waktu	Fasilitas	Harga Kursus
Aldi Pratama	Dekat	Kurang Baik	Ada	Lengkap	Murah
Aldi Siregar	Dekat	Baik	Ada	Lengkap	Murah
Amelia Putri	Dekat	Baik	Ada	Lengkap	Murah
Andi Nasution	Jauh	Baik	Ada	Lengkap	Murah
Andi Saputra	Dekat	Baik	Ada	Lengkap	Murah
Andini Maharani	Dekat	Baik	Ada	Lengkap	Murah
Angga Perdana	Dekat	Baik	Ada	Lengkap	Murah
Anton Wirawan	Jauh	Kurang Baik	Tidak Ada	Tidak Lengkap	Murah
Anwar Putra	Dekat	Baik	Ada	Lengkap	Murah
Ardi Kurniawan	Dekat	Baik	Ada	Lengkap	Murah
Arman Fadil	Jauh	Baik	Ada	Lengkap	Murah
Bella Permata	Jauh	Baik	Tidak Ada	Tidak Lengkap	Mahal
Budi Setiawan	Dekat	Baik	Tidak Ada	Lengkap	Murah
Citra Anindya	Jauh	Kurang Baik	Tidak Ada	Tidak Lengkap	Mahal
Dewi Anjani	Dekat	Kurang Baik	Ada	Tidak Lengkap	Murah
Dian Nasution	Dekat	Baik	Tidak Ada	Tidak Lengkap	Murah
Dika Pratama	Jauh	Baik	Tidak Ada	Tidak Lengkap	Mahal
Dimas Rizki	Dekat	Baik	Tidak Ada	Tidak Lengkap	Murah
Dina Andini	Jauh	Kurang Baik	Tidak Ada	Lengkap	Mahal
Elvina Rahma	Jauh	Kurang Baik	Ada	Tidak Lengkap	Mahal
Fajar Kurnia	Dekat	Baik	Ada	Tidak Lengkap	Mahal
Farah Indriani	Jauh	Kurang Baik	Tidak Ada	Tidak Lengkap	Murah
Farhan Siregar	Dekat	Baik	Ada	Tidak Lengkap	Mahal
Fauzan Fikri	Dekat	Kurang Baik	Tidak Ada	Tidak Lengkap	Mahal
Fikri Hidayat	Dekat	Baik	Ada	Lengkap	Murah
Fikri Rambe	Dekat	Baik	Ada	Lengkap	Mahal
Gita Amanda	Jauh	Kurang Baik	Ada	Lengkap	Mahal
Hanif Alif	Jauh	Baik	Tidak Ada	Tidak Lengkap	Mahal
Hendra Jaya	Jauh	Baik	Tidak Ada	Tidak Lengkap	Mahal
Ida Lestari	Dekat	Kurang Baik	Tidak Ada	Tidak Lengkap	Mahal
Ikhwan Nasution	Jauh	Kurang Baik	Ada	Lengkap	Mahal
Indah Harahap	Dekat	Kurang Baik	Tidak Ada	Lengkap	Murah
Joko Santoso	Jauh	Kurang Baik	Tidak Ada	Lengkap	Mahal
Joko Santoso	Jauh	Baik	Tidak Ada	Tidak Lengkap	Mahal
Junaidi Hasibuan	Jauh	Baik	Ada	Lengkap	Murah

Karina Anjani	Dekat	Baik	Ada	Lengkap	Murah
Laila Siregar	Dekat	Baik	Ada	Lengkap	Murah
Lala Widya	Dekat	Baik	Ada	Lengkap	Murah
Lisa Anggraini	Jauh	Kurang Baik	Ada	Tidak Lengkap	Mahal
Lisa Febriani	Dekat	Baik	Ada	Lengkap	Murah
Mario Rizky	Dekat	Baik	Ada	Lengkap	Murah
Maya Nasution	Dekat	Kurang Baik	Ada	Tidak Lengkap	Murah
Maya Sari	Dekat	Baik	Ada	Lengkap	Mahal
Nadia Lestari	Dekat	Baik	Ada	Lengkap	Murah
Nia Rahma	Dekat	Baik	Ada	Lengkap	Murah
Nisa Ayu	Dekat	Baik	Ada	Lengkap	Murah
Nita Maharani	Dekat	Baik	Ada	Lengkap	Murah
Oki Saputra	Dekat	Baik	Ada	Lengkap	Murah
Putri Melani	Dekat	Baik	Ada	Lengkap	Murah
Rahmat Rambe	Jauh	Baik	Ada	Lengkap	Mahal
Raisa Aulia	Dekat	Baik	Ada	Lengkap	Murah
Rani Dewi	Dekat	Baik	Ada	Lengkap	Murah
Rendy Alamsyah	Dekat	Baik	Ada	Lengkap	Murah
Rina Putri	Jauh	Baik	Ada	Lengkap	Murah
Rio Bagus	Dekat	Baik	Ada	Lengkap	Murah
Rizki Harahap	Dekat	Kurang Baik	Ada	Lengkap	Murah
Rizky Ahmad	Jauh	Baik	Ada	Lengkap	Mahal
Rizwan Rambe	Dekat	Baik	Ada	Lengkap	Murah
Satria Bintang	Dekat	Baik	Ada	Lengkap	Murah
Sinta Hasibuan	Dekat	Baik	Tidak Ada	Lengkap	Murah
Siti Ayu	Dekat	Baik	Tidak Ada	Lengkap	Murah
Suci Maharani	Dekat	Baik	Ada	Lengkap	Murah
Taufik Harahap	Dekat	Baik	Tidak Ada	Lengkap	Murah
Tia Lestari	Dekat	Kurang Baik	Tidak Ada	Lengkap	Murah
Tiara Puspita	Dekat	Baik	Ada	Lengkap	Murah
Umar Fadli	Dekat	Baik	Ada	Lengkap	Murah
Vera Syafira	Dekat	Baik	Ada	Lengkap	Murah
Winda Amalia	Jauh	Kurang Baik	Tidak Ada	Tidak Lengkap	Murah
Wulan Fitri	Dekat	Baik	Ada	Lengkap	Murah
Yuli Fitriani	Jauh	Kurang Baik	Ada	Lengkap	Mahal
Yulia Amira	Jauh	Kurang Baik	Tidak Ada	Tidak Lengkap	Mahal
Yuliana Hasibuan	Dekat	Baik	Ada	Lengkap	Murah
Zakia Melinda	Jauh	Baik	Tidak Ada	Tidak Lengkap	Mahal

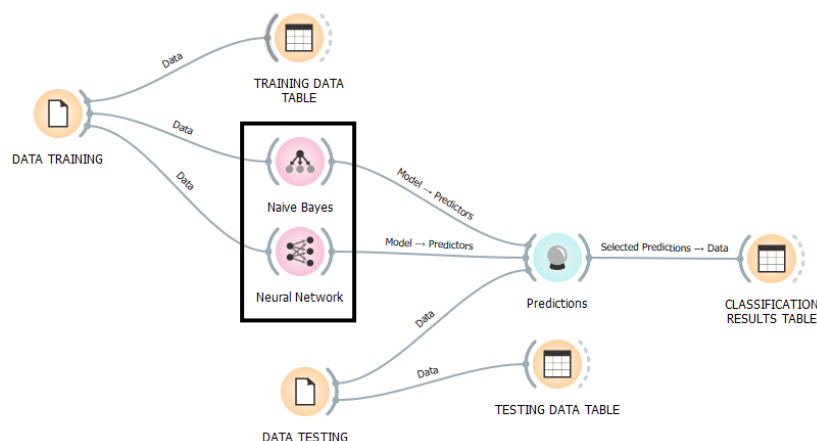
Pada tabel 4. Menyajikan data testing yang digunakan dalam penelitian ini, yang berfungsi sebagai sampel untuk menguji kinerja model yang diterapkan. Data testing ini dipilih secara sistematis untuk memastikan bahwa model dapat menggeneralisasi dengan baik terhadap data baru. Dalam penelitian ini, sebanyak 73 data digunakan sebagai data uji, yang bertujuan untuk mengevaluasi akurasi dan efektivitas model dalam melakukan klasifikasi atau prediksi. Penggunaan data testing yang representatif sangat penting agar hasil penelitian dapat mencerminkan performa model secara objektif dan reliabel.

3.3 Preprocessing Data

Tahap preprocessing data merupakan langkah awal yang esensial dalam proses analisis, yang bertujuan untuk menyaring, menyeleksi, dan mempersiapkan data agar dapat digunakan secara optimal dalam pemodelan. Proses ini mencakup berbagai teknik, seperti pembersihan data dari nilai yang tidak valid, standarisasi format, serta eliminasi atribut yang kurang relevan, sehingga data yang digunakan lebih akurat dan terstruktur. Dengan melakukan preprocessing yang sistematis, kualitas data dapat ditingkatkan, yang pada akhirnya berkontribusi terhadap keakuratan dan efektivitas model dalam menghasilkan hasil analisis yang valid dan dapat diandalkan.

3.4 Perancangan Model Klasifikasi

Tahap selanjutnya dalam penelitian ini adalah perancangan model klasifikasi data, yang bertujuan untuk mengoptimalkan proses pengelompokan dan identifikasi pola dalam dataset. Model ini dirancang secara sistematis menggunakan aplikasi Orange, sebuah perangkat lunak data mining berbasis visual yang memungkinkan analisis data secara interaktif. Dengan memanfaatkan fitur dan algoritma yang tersedia dalam Orange, model klasifikasi dikembangkan untuk memastikan bahwa proses pengujian dan validasi data dapat dilakukan secara efektif dan akurat. Perancangan model ini menjadi langkah krusial dalam menentukan sejauh mana model dapat mengklasifikasikan data dengan tingkat keakuratan yang tinggi.



Gambar 2. Model Klasifikasi

Pada gambar 2. menampilkan model klasifikasi yang dirancang sebagai bagian dari proses analisis dalam penelitian ini, di mana model tersebut dikembangkan untuk mengelompokkan data secara sistematis berdasarkan pola dan karakteristik tertentu. Metode klasifikasi yang digunakan dalam penelitian ini dapat diidentifikasi melalui widget yang ditampilkan dalam kotak hitam, yang berfungsi sebagai indikator algoritma dan teknik yang diterapkan dalam proses pemodelan. Keberadaan model ini menjadi elemen krusial dalam memastikan bahwa data dapat diklasifikasikan dengan akurasi tinggi, sehingga mendukung validitas hasil penelitian secara keseluruhan.

3.5 Hasil Klasifikasi Data

Hasil klasifikasi data merupakan output dari model yang telah dirancang, di mana hasil tersebut menunjukkan bagaimana data dikelompokkan berdasarkan pola yang telah dipelajari. Hasil klasifikasi ini disajikan dalam bentuk tabel untuk mempermudah analisis dan evaluasi performa model. Untuk hasil klasifikasinya dapat dilihat pada tabel dibawah ini.

Tabel 5. Hasil Klasifikasi

Nama	Jarak kursus	Kompetensi Pengajar	Ketersediaan Waktu	Fasilitas	Harga Kursus	Naive Bayes	Neural Network
Aldi Pratama	Dekat	Kurang Baik	Ada	Lengkap	Murah	Minat	Minat
Aldi Siregar	Dekat	Baik	Ada	Lengkap	Murah	Minat	Minat
Amelia Putri	Dekat	Baik	Ada	Lengkap	Murah	Minat	Minat
Andi Nasution	Jauh	Baik	Ada	Lengkap	Murah	Minat	Minat
Andi Saputra	Dekat	Baik	Ada	Lengkap	Murah	Minat	Minat
Andini Maharani	Dekat	Baik	Ada	Lengkap	Murah	Minat	Minat
Angga Perdana	Dekat	Baik	Ada	Lengkap	Murah	Minat	Minat
Anton Wirawan	Jauh	Kurang Baik	Tidak Ada	Tidak Lengkap	Murah	Tidak Minat	Tidak Minat
Anwar Putra	Dekat	Baik	Ada	Lengkap	Murah	Minat	Minat
Ardi Kurniawan	Dekat	Baik	Ada	Lengkap	Murah	Minat	Minat
Arman Fadil	Jauh	Baik	Ada	Lengkap	Murah	Minat	Minat
Bella Permata	Jauh	Baik	Tidak Ada	Tidak Lengkap	Mahal	Tidak Minat	Tidak Minat
Budi Setiawan	Dekat	Baik	Tidak Ada	Lengkap	Murah	Minat	Minat
Citra Anindya	Jauh	Kurang Baik	Tidak Ada	Tidak Lengkap	Mahal	Tidak Minat	Tidak Minat
Dewi Anjani	Dekat	Kurang Baik	Ada	Tidak Lengkap	Murah	Minat	Minat
Dian Nasution	Dekat	Baik	Tidak Ada	Tidak Lengkap	Murah	Minat	Minat
Dika Pratama	Jauh	Baik	Tidak Ada	Tidak Lengkap	Mahal	Tidak Minat	Tidak Minat
Dimas Rizki	Dekat	Baik	Tidak Ada	Tidak Lengkap	Murah	Minat	Minat
Dina Andini	Jauh	Kurang Baik	Tidak Ada	Lengkap	Mahal	Tidak Minat	Tidak Minat
Elvina Rahma	Jauh	Kurang Baik	Ada	Tidak Lengkap	Mahal	Tidak Minat	Tidak Minat
Fajar Kurnia	Dekat	Baik	Ada	Tidak Lengkap	Mahal	Minat	Minat
Farah Indriani	Jauh	Kurang Baik	Tidak Ada	Tidak Lengkap	Murah	Tidak Minat	Tidak Minat
Farhan Siregar	Dekat	Baik	Ada	Tidak Lengkap	Mahal	Minat	Minat
Fauzan Fikri	Dekat	Kurang Baik	Tidak Ada	Tidak Lengkap	Mahal	Tidak Minat	Tidak Minat
Fikri Hidayat	Dekat	Baik	Ada	Lengkap	Murah	Minat	Minat
Fikri Rambe	Dekat	Baik	Ada	Lengkap	Mahal	Minat	Minat
Gita Amanda	Jauh	Kurang Baik	Ada	Lengkap	Mahal	Tidak Minat	Tidak Minat
Hanif Alif	Jauh	Baik	Tidak Ada	Tidak Lengkap	Mahal	Tidak Minat	Tidak Minat
Hendra Jaya	Jauh	Baik	Tidak Ada	Tidak Lengkap	Mahal	Tidak Minat	Tidak Minat
Ida Lestari	Dekat	Kurang Baik	Tidak Ada	Tidak Lengkap	Mahal	Tidak Minat	Tidak Minat
Ikhwan Nasution	Jauh	Kurang Baik	Ada	Lengkap	Mahal	Tidak Minat	Tidak Minat
Indah Harahap	Dekat	Kurang Baik	Tidak Ada	Lengkap	Murah	Minat	Minat
Joko Santoso	Jauh	Kurang Baik	Tidak Ada	Lengkap	Mahal	Tidak Minat	Tidak Minat
Joko Santoso	Jauh	Baik	Tidak Ada	Tidak Lengkap	Mahal	Tidak Minat	Tidak Minat
Junaidi Hasibuan	Jauh	Baik	Ada	Lengkap	Murah	Minat	Minat
Karina Anjani	Dekat	Baik	Ada	Lengkap	Murah	Minat	Minat

Laila Siregar	Dekat	Baik	Ada	Lengkap	Murah	Minat	Minat
Lala Widyia	Dekat	Baik	Ada	Lengkap	Murah	Minat	Minat
Lisa Anggraini	Jauh	Kurang Baik	Ada	Tidak Lengkap	Mahal	Tidak Minat	Tidak Minat
Lisa Febriani	Dekat	Baik	Ada	Lengkap	Murah	Minat	Minat
Mario Rizky	Dekat	Baik	Ada	Lengkap	Murah	Minat	Minat
Maya Nasution	Dekat	Kurang Baik	Ada	Tidak Lengkap	Murah	Minat	Minat
Maya Sari	Dekat	Baik	Ada	Lengkap	Mahal	Minat	Minat
Nadia Lestari	Dekat	Baik	Ada	Lengkap	Murah	Minat	Minat
Nia Rahma	Dekat	Baik	Ada	Lengkap	Murah	Minat	Minat
Nisa Ayu	Dekat	Baik	Ada	Lengkap	Murah	Minat	Minat
Nita Maharani	Dekat	Baik	Ada	Lengkap	Murah	Minat	Minat
Oki Saputra	Dekat	Baik	Ada	Lengkap	Murah	Minat	Minat
Putri Melani	Dekat	Baik	Ada	Lengkap	Murah	Minat	Minat
Rahmat Rambe	Jauh	Baik	Ada	Lengkap	Mahal	Minat	Minat
Raisa Aulia	Dekat	Baik	Ada	Lengkap	Murah	Minat	Minat
Rani Dewi	Dekat	Baik	Ada	Lengkap	Murah	Minat	Minat
Rendy Alamsyah	Dekat	Baik	Ada	Lengkap	Murah	Minat	Minat
Rina Putri	Jauh	Baik	Ada	Lengkap	Murah	Minat	Minat
Rio Bagus	Dekat	Baik	Ada	Lengkap	Murah	Minat	Minat
Rizki Harahap	Dekat	Kurang Baik	Ada	Lengkap	Murah	Minat	Minat
Rizky Ahmad	Jauh	Baik	Ada	Lengkap	Mahal	Minat	Minat
Rizwan Rambe	Dekat	Baik	Ada	Lengkap	Murah	Minat	Minat
Satria Bintang	Dekat	Baik	Ada	Lengkap	Murah	Minat	Minat
Sinta Hasibuan	Dekat	Baik	Tidak Ada	Lengkap	Murah	Minat	Minat
Siti Ayu	Dekat	Baik	Tidak Ada	Lengkap	Murah	Minat	Minat
Suci Maharani	Dekat	Baik	Ada	Lengkap	Murah	Minat	Minat
Taufik Harahap	Dekat	Baik	Tidak Ada	Lengkap	Murah	Minat	Minat
Tia Lestari	Dekat	Kurang Baik	Tidak Ada	Lengkap	Murah	Minat	Minat
Tiara Puspita	Dekat	Baik	Ada	Lengkap	Murah	Minat	Minat
Umar Fadli	Dekat	Baik	Ada	Lengkap	Murah	Minat	Minat
Vera Syafira	Dekat	Baik	Ada	Lengkap	Murah	Minat	Minat
Winda Amalia	Jauh	Kurang Baik	Tidak Ada	Tidak Lengkap	Murah	Tidak Minat	Tidak Minat
Wulan Fitri	Dekat	Baik	Ada	Lengkap	Murah	Minat	Minat
Yuli Fitriani	Jauh	Kurang Baik	Ada	Lengkap	Mahal	Tidak Minat	Tidak Minat
Yulia Amira	Jauh	Kurang Baik	Tidak Ada	Tidak Lengkap	Mahal	Tidak Minat	Tidak Minat
Yuliana Hasibuan	Dekat	Baik	Ada	Lengkap	Murah	Minat	Minat
Zakia Melinda	Jauh	Baik	Tidak Ada	Tidak Lengkap	Mahal	Tidak Minat	Tidak Minat

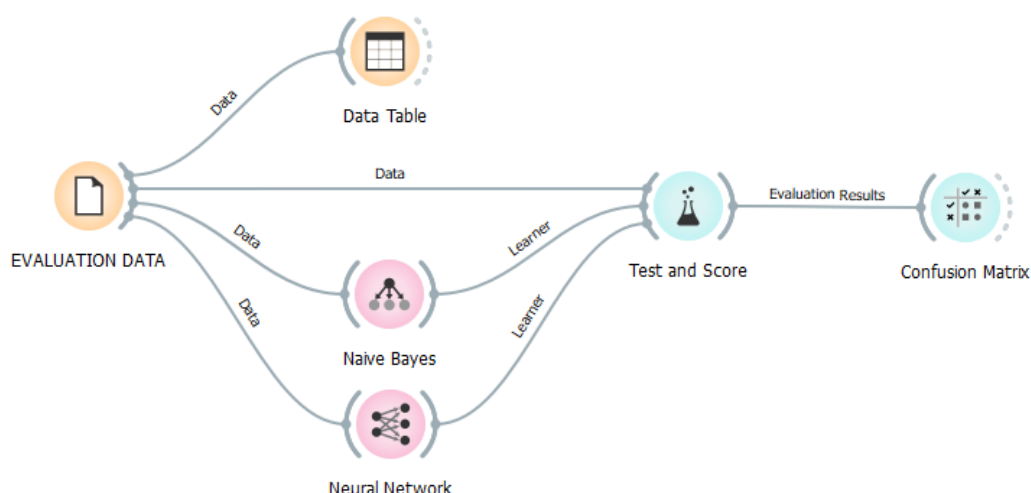
Pada Tabel 5. menyajikan hasil klasifikasi yang diperoleh dari model yang dirancang menggunakan aplikasi Orange, dengan penerapan dua metode utama, yaitu Naïve Bayes dan Neural Network. Model ini dikembangkan untuk mengelompokkan data berdasarkan pola yang telah dipelajari guna menganalisis minat masyarakat terhadap LKP Ibay Komputer. Hasil klasifikasi ini memberikan gambaran mengenai distribusi minat masyarakat berdasarkan prediksi model yang telah diuji.

Hasil klasifikasi yang diperoleh menunjukkan bahwa metode Naïve Bayes mengelompokkan 53 data sebagai masyarakat yang berminat pada LKP Ibay Komputer, sementara 20 data lainnya diklasifikasikan sebagai tidak berminat. Sementara itu, metode Neural Network menghasilkan pola klasifikasi yang serupa, dengan jumlah prediksi minat dan ketidakminatan yang hampir sama.

Meskipun kedua metode menghasilkan distribusi klasifikasi yang mirip, Neural Network memiliki keunggulan dalam menangkap pola yang lebih kompleks dibandingkan Naïve Bayes, yang lebih bergantung pada probabilitas fitur individual. Hal ini menunjukkan bahwa Neural Network dapat memberikan hasil yang lebih adaptif terhadap variasi data, sementara Naïve Bayes tetap menjadi pilihan yang andal untuk klasifikasi berbasis probabilitas dengan kecepatan komputasi yang lebih tinggi.

3.6 Perancangan Model Evaluasi

Tahap evaluasi model merupakan proses krusial dalam mengukur sejauh mana metode yang digunakan mampu mengklasifikasikan data dengan akurat dan andal. Evaluasi ini bertujuan untuk menguji performa model dalam mengidentifikasi pola serta memastikan bahwa hasil prediksi memiliki tingkat akurasi yang tinggi. Berbagai metrik, seperti akurasi, presisi, recall, dan confusion matrix, digunakan untuk menilai efektivitas metode yang diterapkan. Dengan adanya tahap evaluasi, kelebihan dan kekurangan dari setiap metode dapat dianalisis secara mendalam, sehingga memungkinkan optimasi lebih lanjut guna meningkatkan kualitas prediksi dan kinerja model secara keseluruhan.



Gambar 3. Model Evaluasi Metode

Pada Gambar 3. Merupakan model yang dirancang untuk melakukan evaluasi metode yang digunakan pada penelitian ini. Untuk metode yang digunakan yaitu metode Naive Bayes dan metode Neural Network. Model yang dirancang dalam penelitian ini dievaluasi menggunakan widget Test and Score serta widget Confusion Matrix untuk mengukur performa metode yang diterapkan. Test and Score digunakan untuk menghitung berbagai metrik evaluasi, seperti akurasi, presisi, recall, dan F1-score, guna menilai seberapa baik model dalam melakukan klasifikasi. Sementara itu, Confusion Matrix berperan dalam menganalisis distribusi kesalahan prediksi dengan membandingkan hasil klasifikasi model terhadap data aktual. Kombinasi kedua metode evaluasi ini memungkinkan penilaian yang lebih komprehensif, sehingga efektivitas model dapat diukur secara objektif dan akurat.

3.7 Hasil Evaluasi

Evaluasi yang dilakukan pada penelitian ini akan memberikan hasil dari 2 widget yaitu widget test and score dan widget confusion matrix.

3.7.1. Hasil Evaluasi Test and Score

Widget Test and Score merupakan alat evaluasi dalam Orange yang digunakan untuk mengukur performa model klasifikasi berdasarkan berbagai metrik evaluasi. Widget ini melakukan pengujian model dengan metode validasi, seperti Cross-Validation atau Train-Test Split, untuk menilai keakuratan prediksi model terhadap data uji. Metrik yang dihasilkan meliputi akurasi, presisi, recall, F1-score, dan AUC (Area Under Curve), yang memberikan gambaran tentang seberapa baik model dalam mengklasifikasikan data. Dengan menggunakan Test and Score, peneliti dapat membandingkan efektivitas berbagai metode klasifikasi dan mengidentifikasi model dengan performa terbaik.

Tabel 6. Hasil Evaluasi Widget Test and Score

Metode	AUC	CA	F1	Prec	Recall	MCC
Naïve Bayes	1.000	1.000	1.000	1.000	1.000	1.000
Neural Network	1.000	0.973	0.973	0.975	0.973	0.935

Pada Tabel 6. menyajikan hasil evaluasi yang diperoleh dari widget Test and Score menunjukkan bahwa baik metode Naïve Bayes maupun Neural Network menghasilkan performa yang sama tinggi, dengan nilai evaluasi yang mencapai 1.000 atau 100% pada semua metrik, termasuk AUC (Area Under Curve), Classification Accuracy (CA), F1-score, Precision, Recall, dan Matthews Correlation Coefficient (MCC). Hasil ini mengindikasikan bahwa kedua metode mampu mengklasifikasikan data dengan akurasi sempurna, tanpa kesalahan prediksi. Meskipun demikian, penting untuk menganalisis lebih dalam mengapa model-model ini mencapai hasil sempurna. Salah satu kemungkinan penyebabnya adalah ukuran dataset yang relatif kecil, yang sering kali membuat model lebih mudah untuk mengidentifikasi pola-pola sederhana tanpa tantangan signifikan. Selain itu, jika dataset memiliki distribusi yang seimbang dan tidak mengandung noise atau outlier yang signifikan, maka performa tinggi ini bisa jadi mencerminkan kesesuaian model yang sangat baik dengan data uji. Namun, penting juga untuk mempertimbangkan potensi overfitting, terutama jika model tampak terlalu "terlatih" pada data yang terbatas. Hasil Confusion Matrix yang menunjukkan performa ideal tanpa kesalahan klasifikasi bisa menjadi indikasi bahwa model terlalu sensitif terhadap pola dalam dataset tersebut, yang berisiko mengurangi kemampuan generalisasi ketika diterapkan pada data baru di luar sampel penelitian ini. Oleh karena itu, evaluasi tambahan dengan teknik validasi silang (cross-validation) atau pengujian dengan dataset yang lebih besar dan beragam akan sangat disarankan untuk memastikan bahwa performa tinggi tersebut bukan sekadar hasil dari overfitting atau kekhususan dataset yang digunakan.

3.7.2. Hasil Evaluasi Confusion Matrix

Hasil evaluasi menggunakan Confusion Matrix menghasilkan dua matriks yang mewakili masing-masing metode, yaitu Naïve Bayes dan Neural Network. Hal ini disebabkan oleh penggunaan dua pendekatan klasifikasi dalam penelitian ini, yang memungkinkan perbandingan langsung terhadap performa keduanya dalam mengklasifikasikan data.

Tabel 7. Hasil Confusion Matrix Metode Naïve Bayes

		Predicted		
		Interest	Not Satisfied	Σ
Actual	Interest	53	0	53
	Not Interested	0	20	20
Σ		53	20	73

Pada Tabel 7. menunjukkan hasil Confusion Matrix yang diperoleh dari evaluasi model 2522enelitian2522 metode Naïve Bayes. Berdasarkan hasil tersebut, nilai True Positive (TP) adalah 53, yang menunjukkan jumlah siswa yang benar-benar berminat untuk Kursus Komputer, sementara True Negative (TN) adalah 20, yang menunjukkan jumlah siswa yang tidak berminat. Tidak ada prediksi yang salah, dengan False Positive (FP) dan False Negative (FN) masing-masing bernilai 0. Dari hasil ini, akan dihitung untuk nilai akurasi, presisi dan recall. Untuk perhitungannya yaitu sebagai berikut.

$$Accuracy = \frac{53+20}{53+20+0+0} + 100\% \quad \text{Then the Accuracy value} = 100\%$$

$$Precision = \frac{53}{53+0} + 100\% \quad \text{Then the Precision value} = 100\%$$

$$Recall = \frac{53}{53+0} + 100\% \quad \text{Then the Recall value} = 100\%$$

Hasil evaluasi menggunakan Confusion Matrix pada metode Naïve Bayes menunjukkan performa klasifikasi yang sempurna, dengan nilai akurasi, presisi, dan recall mencapai 100%. Dari total 73 data uji, sebanyak 53 data diklasifikasikan dengan benar sebagai berminat (True Positive), dan 20 data diklasifikasikan dengan benar sebagai tidak berminat (True Negative). Tidak ditemukan kesalahan klasifikasi, yang ditunjukkan dengan False Positive (FP) = 0 dan False Negative (FN) = 0. Dengan hasil ini, metode Naïve Bayes terbukti sangat efektif dalam mengklasifikasikan minat masyarakat terhadap kursus komputer, karena seluruh prediksi yang dihasilkan sesuai dengan data aktual.

Tabel 8. Hasil Confusion Matrix Neural Network

		Predicted		
		Interest	Not Satisfied	Σ
Actual	Interest	51	2	53
	Not Interested	0	20	20
Σ		51	22	73

Pada Tabel 8. menunjukkan hasil Confusion Matrix yang diperoleh dari evaluasi model klasifikasi menggunakan metode Neural Network. Berdasarkan hasil tersebut, nilai True Positive (TP) tercatat sebanyak 51, True Negative (TN) sebanyak 20, False Positive (FP) sebanyak 2, dan False Negative (FN) sebesar 0. Dengan data ini, dapat dihitung nilai akurasi, presisi, dan recall, yang memberikan gambaran mendalam tentang kinerja model dalam mengklasifikasikan minat siswa dan siswi terhadap Kursus Komputer. Nilai-nilai ini sangat penting untuk mengevaluasi sejauh mana model dapat melakukan klasifikasi dengan benar, serta mengidentifikasi tingkat kesalahan yang terjadi dalam proses prediksi.

$$Accuracy = \frac{51+20}{51+20+2+0} + 100\% \quad \text{Then the Accuracy value} = 97\%$$

$$Precision = \frac{51}{51+2} + 100\% \quad \text{Then the Precision value} = 98\%$$

$$Recall = \frac{51}{51+0} + 100\% \quad \text{Then the Recall value} = 100\%$$

Hasil evaluasi Confusion Matrix pada metode Neural Network menunjukkan tingkat akurasi yang sangat baik, yaitu 97%, meskipun masih terdapat sedikit kesalahan klasifikasi. Dari total 73 data uji, sebanyak 51 data diklasifikasikan dengan benar sebagai berminat (True Positive), sementara 20 data diklasifikasikan dengan benar sebagai tidak berminat (True Negative). Namun, terdapat 2 data yang salah diklasifikasikan sebagai berminat padahal

sebenarnya tidak berminat (False Positive), sedangkan kesalahan klasifikasi False Negative tidak ditemukan. Dengan presisi sebesar 98%, model ini menunjukkan kemampuannya yang tinggi dalam memprediksi minat dengan sedikit kesalahan positif. Sementara itu, nilai recall mencapai 100%, yang berarti semua individu yang berminat terhadap kursus komputer berhasil terdeteksi dengan benar. Hasil ini mengindikasikan bahwa Neural Network memiliki performa yang mendekati sempurna, meskipun tidak seakurat metode Naïve Bayes, yang memiliki tingkat kesalahan nol.

4. KESIMPULAN

Hasil evaluasi yang diperoleh dari widget Test and Score menunjukkan bahwa baik metode Naïve Bayes maupun Neural Network menghasilkan performa yang sama tinggi, dengan nilai evaluasi yang mencapai 1.000 atau 100% pada semua metrik, termasuk AUC (Area Under Curve), Classification Accuracy (CA), F1-score, Precision, Recall, dan Matthews Correlation Coefficient (MCC). Hasil ini mengindikasikan bahwa kedua metode mampu mengklasifikasikan data dengan akurasi sempurna, tanpa kesalahan prediksi. Meskipun demikian, penting untuk menganalisis lebih dalam mengapa model-model ini mencapai hasil sempurna. Salah satu kemungkinan penyebabnya adalah ukuran dataset yang relatif kecil, yang sering kali membuat model lebih mudah untuk mengidentifikasi pola-pola sederhana tanpa tantangan signifikan. Selain itu, jika dataset memiliki distribusi yang seimbang dan tidak mengandung noise atau outlier yang signifikan, maka performa tinggi ini bisa jadi mencerminkan kesesuaian model yang sangat baik dengan data uji. Hasil klasifikasi dari penelitian ini menunjukkan bahwa metode Naïve Bayes dan Neural Network memberikan hasil yang identik, dengan 53 data diklasifikasikan sebagai masyarakat yang berminat terhadap LKP Ibay Komputer dan 20 data diklasifikasikan sebagai tidak berminat. Berdasarkan Confusion Matrix, kedua metode menghasilkan True Positive (TP) sebesar 53, True Negative (TN) sebesar 20, False Positive (FP) sebesar 0, dan False Negative (FN) sebesar 0, yang berarti tidak ada kesalahan dalam prediksi. Dengan kata lain, baik Naïve Bayes maupun Neural Network mampu mengklasifikasikan data dengan sempurna tanpa kesalahan, sehingga akurasi yang diperoleh mencapai 100%. Hasil ini menunjukkan bahwa dataset yang digunakan memiliki pola klasifikasi yang cukup jelas, sehingga kedua metode dapat mengenali pola dengan sangat baik tanpa mengalami kendala dalam prediksi. Namun, penting juga untuk mempertimbangkan potensi overfitting, terutama jika model tampak terlalu "terlatih" pada data yang terbatas. Hasil Confusion Matrix yang menunjukkan performa ideal tanpa kesalahan klasifikasi bisa menjadi indikasi bahwa model terlalu sensitif terhadap pola dalam dataset tersebut, yang berisiko mengurangi kemampuan generalisasi ketika diterapkan pada data baru di luar sampel penelitian ini. Oleh karena itu, evaluasi tambahan dengan teknik validasi silang (cross-validation) atau pengujian dengan dataset yang lebih besar dan beragam akan sangat disarankan untuk memastikan bahwa performa tinggi tersebut bukan sekadar hasil dari overfitting atau kekhususan dataset yang digunakan.

REFERENCES

- [1] Putri Ayu Negara, Yuni Yulistianti, Elsit Julia Pratiwi, and Yusup Saeful bayan, "Pentingnya Pemanfaatan Teknologi Informasi Dalam Bidang Pendidikan Di Desa Mekar Asih," *J. Abdi Nusa*, vol. 3, no. 3, pp. 208–213, 2023, doi: 10.52005/abdinusa.v3i3.197.
- [2] M. Sari, G. J. Yanris, and M. N. S. Hasibuan, "Analysis of the Neural Network Method to Determine Interest in Buying Pertamina Fuel," *Sinkron*, vol. 8, no. 2, pp. 1031–1039, 2023, doi: 10.33395/sinkron.v8i2.12292.
- [3] A. P. Siregar, D. Irmayani, and M. N. Sari, "Analysis of the Naïve Bayes Method for Determining Social Assistance Eligibility Public," *Sinkron*, vol. 8, no. 2, pp. 805–817, 2023, doi: 10.33395/sinkron.v8i2.12259.
- [4] A. Wahyudi, S. Ovelia Tampubolon, N. Afrilia Putri, A. Ghassa, E. Rasywir, and D. Kisbianty, "Penerapan Data Mining Algoritma Naive Bayes Clasifier Untuk Mengetahui Minat Beli Pelanggan Terhadap INDIHOME," *J. Inform. Dan Rekayasa Komputer (JAKAKOM)*, vol. 2, no. 2, pp. 240–247, 2022, doi: 10.33998/jakakom.2022.2.2.111.
- [5] M. N. Zarti, E. Sahputra, A. Sonita, and Y. Apridiansyah, "Application Of Data Mining Using The Naïve Bayes Classification Method To Predict Public Interest Participation In The 2024 Elections," *J. Komputer, Inf. dan Teknol.*, vol. 3, no. 1, pp. 105–114, 2023, doi: 10.53697/jkomitek.v3i1.1192.
- [6] D. Diana Dewi, N. Qisthi, S. S. S. Lestari, and Z. H. S. Putri, "Perbandingan Metode Neural Network Dan Support Vector Machine Dalam Klasifikasi Diagnosa Penyakit Diabetes," *Cerdika J. Ilm. Indones.*, vol. 3, no. 09, pp. 828–839, 2023, doi: 10.59141/cerdika.v3i09.662.
- [7] E. Suherman, D. Hindarto, A. Makmur, and H. Santoso, "Comparison of Convolutional Neural Network and Artificial Neural Network for Rice Detection," *Sinkron*, vol. 8, no. 1, pp. 247–255, 2023, doi: 10.33395/sinkron.v8i1.11944.
- [8] R. F. Nasution, M. H. Dar, and F. A. Nasution, "Implementation of the Naïve Bayes Method to Determine Student Interest in Gaming Laptops," *Sinkron*, vol. 8, no. 3, pp. 1709–1723, 2023, doi: 10.33395/sinkron.v8i3.12562.
- [9] M. E. Apriyani, R. A. Maskuri, M. H. Ratsanjani, A. N. Pramudhita, and R. Rawansyah, "Digital Forensic Investigates Sexual Harassment on Telegram using Naïve Bayes," *Sinkron*, vol. 8, no. 3, pp. 1409–1417, 2023, doi: 10.33395/sinkron.v8i3.12514.
- [10] J. P. Tanjung, F. C. Tampubolon, A. W. Panggabean, and M. A. A. Nandrawan, "Customer Classification Using Naive Bayes Classifier With Genetic Algorithm Feature Selection," *Sinkron*, vol. 8, no. 1, pp. 584–589, 2023, doi: 10.33395/sinkron.v8i1.12182.
- [11] A. Saleh, N. Dharshinni, D. Perangin-Angin, F. Azmi, and M. I. Sarif, "Implementation of Recommendation Systems in Determining Learning Strategies Using the Naïve Bayes Classifier Algorithm," *Sinkron*, vol. 8, no. 1, pp. 256–267, 2023, doi: 10.33395/sinkron.v8i1.11954.
- [12] R. Rahman and F. Fauzi Abdulloh, "Performance of Various Naïve Bayes Using GridSearch Approach In Phishing Email



- Dataset,” *Sinkron*, vol. 8, no. 4, pp. 2336–2344, 2023, doi: 10.33395/sinkron.v8i4.12958.
- [13] H. Silaban, S. Putri, and N. Evta, “Analysis Indonesia ’ s Export Value Forecasting to G20 Countries Using Long Short-Term Memory Neural Network Method,” *Sinkron*, vol. 8, no. 3, pp. 2008–2017, 2023.
- [14] V. Lestari, H. Mawengkang, and Z. Situmorang, “Artificial Neural Network Backpropagation Method to Predict Tuberculosis Cases,” *Sinkron*, vol. 8, no. 1, pp. 35–47, 2023, doi: 10.33395/sinkron.v8i1.11998.
- [15] S. Willian, T. H. Rochadiani, and T. Sofian, “Design of Batak Toba Script Recognition System Using Convolutional Neural Network Algorithm,” *Sinkron*, vol. 8, no. 3, pp. 1609–1618, 2023, doi: 10.33395/sinkron.v8i3.12617.
- [16] B. Anwar, N. Jalinus, and R. Abdullah, “Weather Forecast In Medan City With Hopfield Artificial Neural Network Algorithm,” *Sinkron*, vol. 8, no. 1, pp. 398–404, 2023, doi: 10.33395/sinkron.v8i1.12048.
- [17] A. B. P. B. Wedha, B. Rahman, D. Hindarto, and B. Y. Wedha, “Drowsy Detection in the Eye Area using the Convolutional Neural Network,” *Sinkron*, vol. 8, no. 2, pp. 1097–1107, 2023, doi: 10.33395/sinkron.v8i2.12386.
- [18] A. Alam, D. A. F. Alana, and C. Juliane, “Comparison Of The C.45 And Naive Bayes Algorithms To Predict Diabetes,” *Sinkron*, vol. 8, no. 4, pp. 2641–2650, 2023, doi: 10.33395/sinkron.v8i4.12998.
- [19] F. F. Hasibuan, M. H. Dar, and G. J. Yanris, “Implementation of the Naïve Bayes Method to determine the Level of Consumer Satisfaction,” *Sinkron*, vol. 8, no. 2, pp. 1000–1011, 2023, doi: 10.33395/sinkron.v8i2.12349.
- [20] H. A. Pratama, G. J. Yanris, M. Nirmala, and S. Hasibuan, “Implementation of Data Mining for Data Classification of Visitor Satisfaction Levels,” *Sinkron*, vol. 8, no. 3, pp. 1832–1851, 2023.
- [21] P. Violita, G. J. Yanris, and M. N. S. Hasibuan, “Analysis of Visitor Satisfaction Levels Using the K-Nearest Neighbor Method,” *Sinkron*, vol. 8, no. 2, pp. 898–914, 2023, doi: 10.33395/sinkron.v8i2.12257.
- [22] S. Maizura, V. Sihombing, and M. H. Dar, “Analysis of the Decision Tree Method for Determining Interest in Prospective Student College,” *Sinkron*, vol. 8, no. 2, pp. 956–979, 2023, doi: 10.33395/sinkron.v8i2.12258.
- [23] D. J. Triani, M. H. Dar, and G. J. Yanris, “Analysis of Public Purchase Interest in Yamaha Motorcycles Using the K-Nearest Neighbor Method,” *Sinkron*, vol. 8, no. 3, pp. 1238–1254, 2023, doi: 10.33395/sinkron.v8i3.12433.
- [24] A. Mawaddah, M. H. Dar, and G. J. Yanris, “Analysis of the SVM Method to Determine the Level of Online Shopping Satisfaction in the Community,” *Sinkron*, vol. 8, no. 2, pp. 838–855, 2023, doi: 10.33395/sinkron.v8i2.12261.
- [25] S. A. Hasibuan, V. Sihombing, and F. A. Nasution, “Analysis of Community Satisfaction Levels using the Neural Network Method in Data Mining,” *Sinkron*, vol. 8, no. 3, pp. 1724–1735, 2023, doi: 10.33395/sinkron.v8i3.12634.