



Studi Komperatif Penggunaan Algoritma *Decision Tree* dan *K-NN (K-Nearest Neighbors)* Pada Prediksi Pemilihan Warna Untuk Produksi Mobil Berdasarkan Minat Konsumen

Muhammad Farid Salafudin Firdaus ^{a,1}, Agung Fatwanto ^{b,2}

^a Prodi Teknik Industri, Universitas Islam Negeri Sunan Kalijaga, Yogyakarta

^b Prodi Teknik Informatika, Universitas Islam Negeri Sunan Kalijaga, Yogyakarta

¹ salafudin.farid@gmail.com*; ² agung.fatwanto@uin-suka.ac.id

*corresponding author

ABSTRACT

Semakin banyaknya brand mobil yang ada di Indonesia membuat perusahaan mobil melakukan inovasi terhadap produknya untuk menarik minat pembeli. Salah satu inovasi yang ditawarkan adalah warna mobil, hal ini karena warna mobil sebagai inovasi paling murah dan berpengaruh terhadap minat beli. Sebelum menentukan warna yang akan diproduksi, perusahaan perlu melakukan analisa pasar terhadap pola pemilihan warna terhadap minat beli. Hal ini bertujuan untuk menentukan jumlah varian warna dan juga model mobil yang akan diproduksi, agar tidak terjadi kerugian akibat mobil kurang diminati dengan alasan model dan warna tidak sesuai dengan keinginan sebagian besar pembeli. Oleh karena itu, dalam penelitian ini dilakukan studi komperatif untuk pemilihan mobil berdasarkan warna menggunakan Algoritma *Decision Tree* dan *K-NN (K-Nearest Neighbors)*, dengan tujuan memberikan gambaran pola keinginan pembeli berdasarkan variabel yang sudah ditentukan yaitu Status (perkawinan), Model dan Umur. Pendekatan ini memberikan sudut pandang baru dalam memahami preferensi pasar otomotif melalui data mining. Pengujian ini menggunakan bantuan software Rapidminer dan menggunakan Algoritma *Decision Tree* dan *K-NN (K-Nearest Neighbors)*. Hasil yang didapatkan menggunakan *split data* dengan nilai proporsi 80% sebagai data pelatihan dan 20% data uji, didapatkan hasil uji *performance* dengan nilai *accuracy* sebesar 70,31% untuk Algoritma *Decision Tree* dan 78,12% untuk Algoritma *K-NN (K-Nearest Neighbors)*. Sehingga pola pemilihan warna menggunakan Algoritma *K-NN (K-Nearest Neighbors)* dapat digunakan sebagai pertimbangan perusahaan atau dealer mobil untuk mengukur kuantitas produksi berdasarkan warna dan mengukur stok persediaan mobil di dealer, sehingga mengurangi biaya penyimpanan.



ARTICLE INFO

Article history

Received: 11 Mei 2025

Revised: 23 Mei 2025

Accepted: 24 November 2025

Keywords:

Algoritma *Decision Tree*,
Pemilihan Mobil, Warna,
Rapidminer,
Algoritma *k-NN*,
K-Nearest Neighbors

1. Pendahuluan

Mobil adalah salah satu alat transportasi sehari-hari selain kendaraan bermotor yang digunakan masyarakat. Munculnya kendaraan listrik, menambah segmentasi mobil yang sebelumnya dikuasai oleh kendaraan berbahan bakar fosil. Hal ini menunjukkan bahwa persaingan dari penjualan mobil di Indonesia semakin ketat. Produsen mobil akan lebih dalam melakukan inovasi produk yang mereka jual. Hal ini dikarenakan untuk mempertahankan penjualan dengan banyaknya pesaing, para

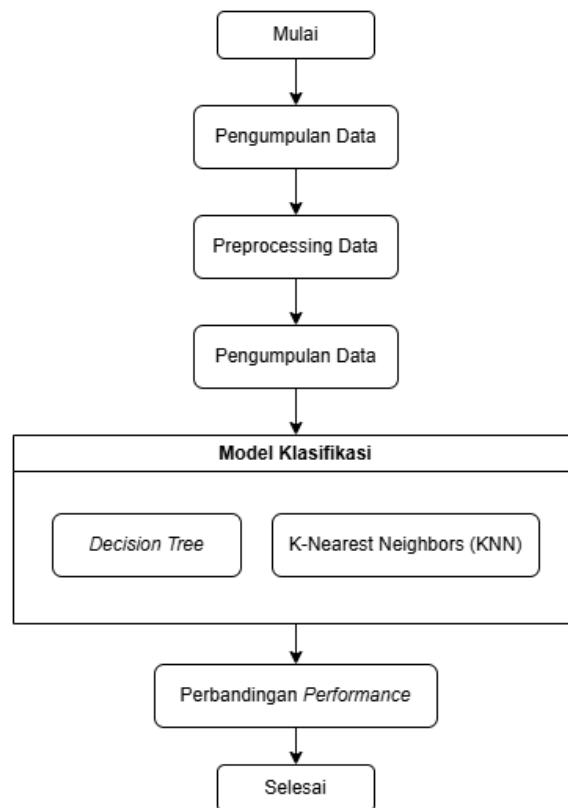


produsen mempelajari faktor-faktor yang mempengaruhi penjualan produk yang mereka buat dan memahami secara mendalam keinginan pembeli [1]. Saat ini banyak merk mobil berasal dari produsen yang membuat pabrik di Indonesia maupun langsung impor memiliki berbagai keunggulan masing-masing. Selain keunggulan fitur-fitur, salah satu faktor dari pembeli memilih mobil yaitu dari warna mobil yang semakin bervariasi. Warna menjadi hal pertama yang menarik perhatian seseorang dalam sebuah karya visual, terutama warna yang kontras akan mengundang indera penglihatan seseorang tertuju pada karya tersebut [2] [3]. Warna memegang peran penting dalam membentuk suasana yang mempengaruhi keputusan pembelian, memperkuat citra suatu produk serta meningkatkan reputasi bisnis [4] [5]. Hal ini terjadi pertama kali ketika GM (General Motor) membuktikan bahwa inovasi tidak semuanya tentang perubahan desain, akan tetapi dimulai dengan perubahan warna dan variasi yang digunakan membuat sebuah ketertarikan konsumen, sehingga menjadikan *Art and Colour Selection* menjadikan pengaruh yang tidak diragukan sejak 1937 [6]. Oleh karena itu banyak produsen mobil mulai membuat warna-warna mobil yang berbeda. Konsumen akan membayar lebih untuk tampilan yang berbeda dan menarik, karena hal tersebut yang paling mudah untuk membuat perbedaan harga dalam produksi massal [7]. Berdasarkan *Axalta Global Automotive Color Popularity Report 2024*, warna putih masih menjadi warna mobil paling populer secara global dengan proporsi 31%, diikuti oleh abu-abu (23%) dan hitam (21%) [8]. Dominasi warna netral tersebut menunjukkan bahwa preferensi konsumen terhadap warna yang bersifat universal dan mudah diterima masih sangat kuat. Pemilihan warna dalam industri otomotif memiliki implikasi luas terhadap strategi desain, pemasaran, serta perilaku konsumen. Oleh karena itu, analisis tren warna menjadi aspek penting yang perlu diperhatikan oleh produsen dalam merancang strategi produk yang berorientasi pasar.

Saat ini dengan adanya *database* atau riwayat penjualan yang dimiliki oleh produsen mobil, kita dengan mudah bisa menentukan kecenderungan pembeli mobil akan memilih mobil berdasarkan warna dengan mengesampingkan harga. Selain itu juga hasil ini dapat digunakan untuk produsen menentukan jumlah produksi mobil sesuai dengan warna mobil yang sering dibeli. Oleh karena itu, dengan kemajuan teknologi informasi kita dapat melakukan prediksi pemilihan mobil berdasarkan warna dengan bantuan *software* analisis dan pengolahan data. Pengolahan dan analisa data digunakan untuk melakukan prediksi yaitu menggunakan Algoritma *Decision Tree* dan K-NN (*K-Nearest Neighbors*). Algoritma *Decision Tree* digunakan untuk menentukan beberapa model dengan kelas yang dipisahkan dan digunakan untuk pembeda antara tiap kelas [9]. Pemilihan Metode dengan Algoritma *Decision Tree* karena dapat menghasilkan sebuah pohon keputusan yang memetakan kondisi dan tindakan yang harus diambil berdasarkan kondisi [10]. Selain itu ada juga Algoritma K-NN (*K-Nearest Neighbors*) yaitu metode regresi nonparametrik yang melakukan prediksi pengamatan di masa mendatang berdasarkan situasi serupa saat ini [11].

2. Metodologi Penelitian

Penelitian ini meliputi beberapa tahapan, seperti terlihat pada Gambar 1 dibawah ini. Tahapan atau metodologi penelitian mengacu pada proses perolehan data dan informasi hingga pengolahan data untuk mencari serta memperoleh data yang akurat untuk digunakan dalam penelitian ini.



Gambar 1. Diagram Alir Penelitian

2.1. Identifikasi Masalah

Dalam penelitian ini, tahapan pertama yang dilakukan adalah identifikasi masalah. Berdasarkan hasil identifikasi permasalahan yang telah dilakukan, dapat disimpulkan sebagai berikut:

1. Mengetahui pola pengaruh warna pada pemilihan mobil dengan menggunakan Model Algoritma *Decision Tree* dan K-NN (*K-Nearest Neighbors*).
2. Melakukan pengujian Model Algoritma *Decision Tree* dan K-NN (*K-Nearest Neighbors*) dengan *performance* klasifikasi.

2.2. Studi Literatur

Setelah dilakukan identifikasi masalah, tahap selanjutnya adalah penulis melakukan studi literatur untuk menyelesaikan permasalahan dengan mengumpulkan informasi dari berbagai sumber. Sumber tersebut seperti jurnal dan informasi lain yang berkaitan dengan topik dalam penelitian ini yaitu seputar implementasi Algoritma *Desicion Tree* dan K-NN (*K-Nearest Neighbors*) untuk melakukan identifikasi

2.3. Pengumpulan Data

Metode pengumpulan data yang digunakan dalam penelitian ini adalah data sekunder yang diperoleh dari *website*. Dataset tersebut didapatkan dari kaggle.com, yaitu sebuah website yang berisi dataset dari komunitas online, kompetisi dan data praktisi pembelajaran dibawah Google LCC. Dataset yang didapatkan adalah Dataset Calon Pembeli Mobil. Dataset tersebut berisi warna mobil, model mobil, umur pembeli, status pembeli dan kelamin pembeli.

2.4. Model Penelitian

Penelitian ini akan dilakukan pengujian dengan menggunakan metode klasifikasi *data mining* dengan Algoritma *Decision Tree* dan K-NN (*K-Nearest Neighbors*) terhadap pola pengaruh warna terhadap pemilihan mobil. Data yang tersedia akan diolah untuk kemudian dilihat tingkat akurasi dari setiap algoritma yang digunakan. Adapun hasil dari pengolahan data menggunakan Algoritma *Decision Tree* dan K-NN (*K-Nearest Neighbors*) akan menghasilkan model dan model tersebut akan dilakukan pengujian *performance* klasifikasi untuk mendapatkan nilai *Accuracy*.

2.5. Pengolahan Data Awal

Tahapan ini merupakan tahap awal dalam pengolahan *data mining*, dimana data yang telah diperoleh akan diolah ke dalam format yang dibutuhkan sesuai dengan kebutuhan, seperti pengelompokan data dan penentuan atribut data. Adapun tahapan pengolahan data awal adalah sebagai berikut.

2.5.1 Select Data

Pada tahapan ini dilakukan seleksi atau pemilihan variabel data yang akan dianalisis, karena tidak semua data akan digunakan untuk pengujian dan hanya data tertentu saja yang akan digunakan sesuai dengan kebutuhan dan tujuan penulisan pada penelitian ini.

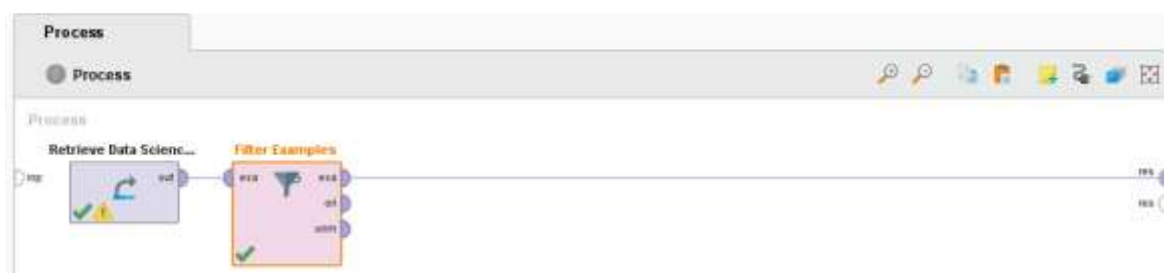
2.5.2 Preprocessing Data

Pada tahapan ini dataset yang akan digunakan akan dilakukan pembersihan data terlebih dahulu menggunakan operator *Replace Missing Value* yang tersedia di Rapidminer, sehingga menghasilkan data yang bebas dari *missing value*. *Missing value* akan dideteksi dengan dengan membandingkan data dengan nilai rata-rata. Namun, operator *Replace Missing Value* tidak dapat memperoleh data dari label kosong, sehingga disini memerlukan penggunaan operator tambahan yaitu *Filter Example* dengan *label is not missing*.

Name	Type	Missing	Missing
COLOR	String	20	Missing
MODEL	String	0	Missing
UMR	Integer	0	Missing
STATUS	Integer	0	Missing
KELAMIN	Integer	0	Missing

Gambar 2. Dataset awal

Gambar 2 menunjukkan bahwa hasil dari tahapan pengolahan dataset awal, dimana pada kolom *Missing* nilai yang tercantum adalah 20 untuk label data dan 0 untuk atribut data lainnya, hal ini menandakan bahwa terdapat data yang mengalami *missing value*. Oleh karena itu perlu dilakukan proses *Filter Examples* dengan memasukkan atribut dengan dengan memilih *label is not missing*.



Gambar 3. Model *Filter Examples* di Rapidminer

Gambar 3 menunjukan model penggunaan *filter examples* pada Rapidminer, kemudian setelah dilakukan *running* akan menghasilkan hasil seperti pada Gambar 4.

Name	Type	Missing	Missing
COLOR	String	0	Missing
MODEL	String	0	Missing
UMR	Integer	0	Missing
STATUS	Integer	0	Missing
KELAMIN	Integer	0	Missing

Gambar 4. Hasil Pengolahan Dataset setelah dilakukan *Filter Examples*

2.5.3 Data Transformation

Tahapan ini adalah tahap mengubah data menjadi bentuk yang sesuai untuk selanjutnya diproses dalam *data mining*, dimana data sebelumnya telah melewati tahap *preprocessing* atau pembersihan data. Sehingga diperoleh data yang siap untuk dilakukan pengujian dengan menggunakan algoritma *Decision Tree* dan K-NN (*K-Nearest Neighbors*).

2.5.4 Pengujian dengan RapidMiner

Rapidminer adalah sebuah perangkat lunak yang memiliki fungsi sebagai media belajar *data mining* [12]. Aplikasi tersebut dibuat oleh Dr. Markus Hofmann dari Institute of Technology Blanchardstown dan Ralf Klinkenberg dari rapid-i.com dengan tampilan GUI (*Graphical User Interface*) sehingga memudahkan pengguna dalam menggunakan perangkat lunak ini [13]. Perangkat lunak ini bersifat *open source* dan dibuat dengan menggunakan program Java di bawah lisensi *GNU Public Licence* dan *Rapid Miner* dapat dijalankan di sistem operasi manapun [14]. Dengan menggunakan Rapid Miner tidak dibutuhkan kemampuan koding khusus, karena semua fasilitas sudah disediakan. Rapid Miner dikhususkan untuk penggunaan *data mining*. Model yang disediakan juga cukup banyak dan lengkap, seperti Model *Bayesian*, *Modelling*, *Tree Induction*, *Neural Network* dan lain-lain

2.5.5 Evaluasi dan Validasi Hasil

Proses validasi ini menggunakan *Confusion matrix*, yang merupakan sebuah alat yang digunakan untuk melakukan analisis tingkat *accuracy*, *precision*, dan *recall* terhadap suatu model dari algoritma [15]. Dengan menggunakan *confusion matrix* akan membantu peneliti dalam memberikan rincian klasifikasi terhadap kelas yang diprediksi. Sedangkan *accuracy* merupakan nilai dari ketepatan presentasi antara nilai yang diprediksi dengan nilai yang sebenarnya. *Precision* merupakan nilai akurasi dari kelas yang sudah diprediksi, dan *recall* merupakan nilai tingkat keberhasilan dari algoritma yang digunakan dan mengukur efektivitas model dalam mengklasifikasi kelas minoritas dengan menilai akurasi model terhadap kasus positif [16].

3. Hasil dan Pembahasan

Dalam implementasi hasil akhir dari penerapan Algoritma *Decision Tree* dan K-NN (*K-Nearest Neighbors*) dibagi dalam dua tahapan, yaitu pertama proses pengolahan dataset dengan Algoritma *Decision Tree* dan K-NN (*K-Nearest Neighbors*) serta yang kedua proses *performance* dari dataset untuk mendapatkan tingkat akurasi dari proses *training data* dan pengujian dataset. Berdasarkan penelitian menggunakan dataset skunder yang diperoleh dari kaggle.com, yang selanjutnya dilakukan klasifikasi menggunakan Algoritma *Decision Tree* dan K-NN (*K-Nearest Neighbors*) untuk mendapatkan pola pengaruh warna terhadap pemilihan mobil dengan bantuan *software* RapidMiner didapatkan hasil sebagai berikut.

3.1. Dataset

Dataset calon pembeli mobil yang nantinya akan dilakukan proses modeling dengan Algoritma *Decision Tree* dan K-NN (*K-Nearest Neighbors*) untuk mendapatkan pola pengaruh warna pada pemilihan mobil. Dalam penelitian ini, terdapat beberapa kategori data dengan keterangan sebagai berikut. Warna netral terdiri atas hitam dan putih, sedangkan warna primer mencakup biru, hijau, merah, dan lainnya. Variabel status pernikahan 0 untuk belum menikah, 1 untuk sudah menikah, 2 untuk sudah menikah dan memiliki anak, serta 3 untuk duda atau janda. Sementara itu, variabel jenis kelamin diberi kode 0 untuk laki-laki dan 1 untuk perempuan.

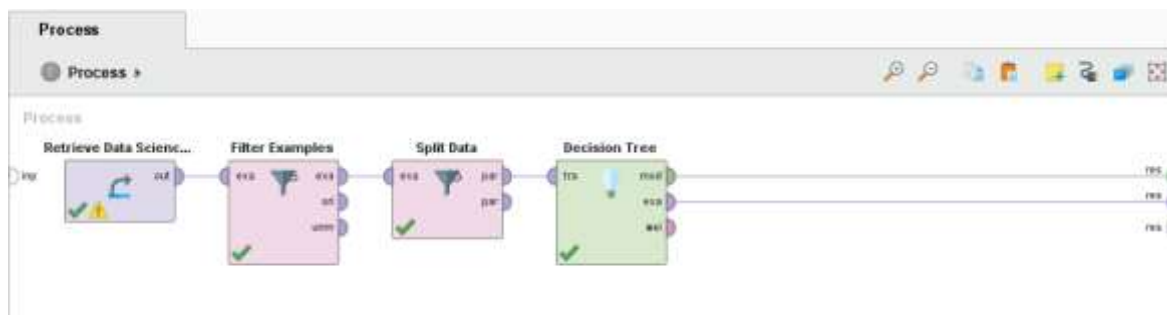
Tabel 1. Dataset

NO	COLOR	MODEL	UMUR	STATUS	KELAMIN
1	WARNA NETRAL	SUV	32	1	0
2	WARNA NETRAL	MINIBUS	49	2	1
3	WARNA NETRAL	SUV	52	2	0
...
337	WARNA NETRAL	SEDAN	33	2	0
338	WARNA PRIMER	SUV	33	2	1

NO	COLOR	MODEL	UMUR	STATUS	KELAMIN
339	WARNA NETRAL	SUV	33	1	0

3.2. Model Algoritma Decision Tree

Proses pengolahan dataset dengan membuat model *training* dan uji menggunakan bantuan *software* Rapidminer dengan algoritma *Decision Tree* ditunjukkan pada gambar dibawah ini.



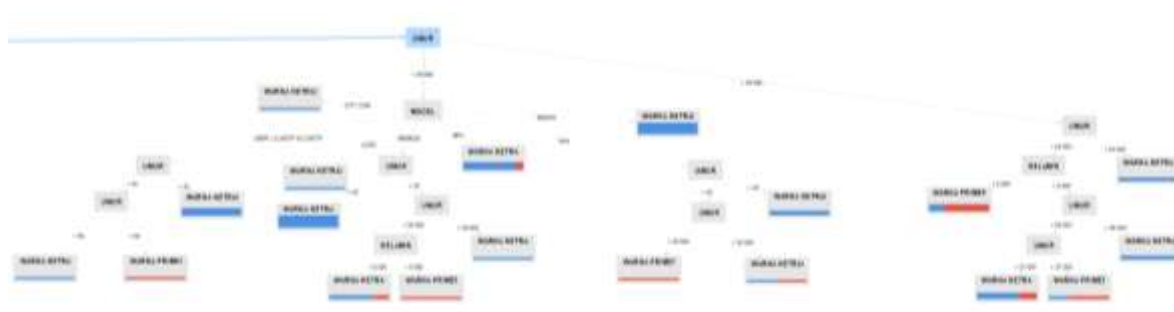
Gambar 5. Model Algoritma Decision Tree

Dari hasil modeling yang telah dibuat sesuai dengan Gambar 5, maka dilakukan proses *running* untuk mendapatkan hasil *Tree* dan dapat melihat hubungan yang mempengaruhi dengan atribut.



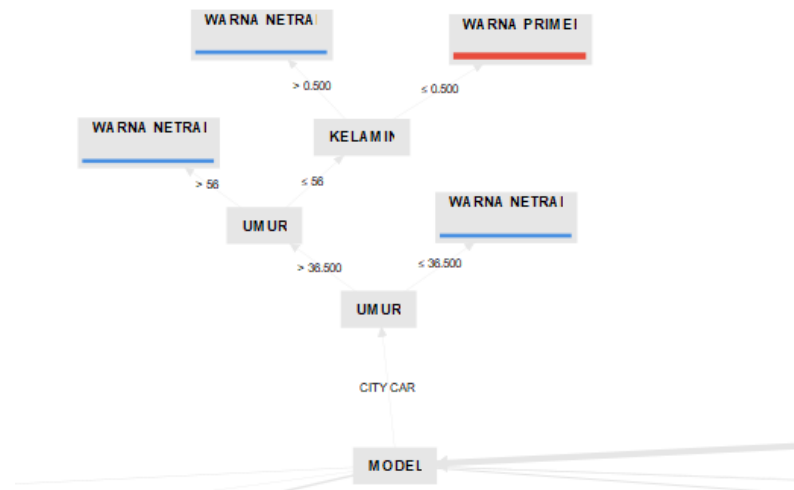
Gambar 6. Pohon Keputusan Hasil Modeling

Gambar 6 menunjukkan hasil dari *running* pada aplikasi Rapidminer terdapat 3 atribut yang signifikan, yaitu Status (Perkawinan), Umur dan Model. Node yang didapatkan sebagai hasil percabangan adalah dengan Status (Perkawinan) sebagai atribut utama dengan hasil ≤ 1.5 untuk variabel Umur dan ≥ 1.5 untuk variabel Model. Sehingga hasil yang didapatkan adalah pohon keputusan dengan akar banyak cabang, hal ini dipengaruhi karena variabel yang digunakan banyak sehingga mempengaruhi hasil pohon keputusan.



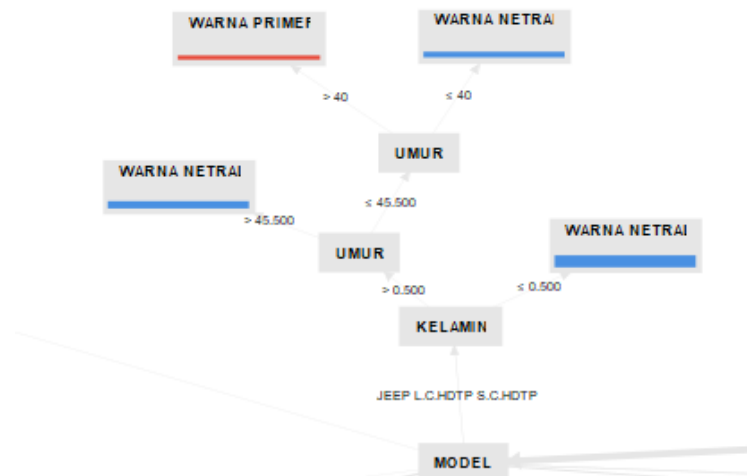
Gambar 7. Pohon keputusan variabel umur dengan status ≤ 1.5

Hasil variabel umur dengan status ≤ 1.5 ditunjukkan pada Gambar 7. Pada gambar tersebut menunjukkan bahwa kelompok umur yang lebih muda cenderung memilih warna primer, sedangkan umur yang lebih tua lebih banyak memilih warna netral. Hal ini menunjukkan bahwa variabel umur memiliki pengaruh terhadap preferensi warna kendaraan pada kelompok status tersebut.



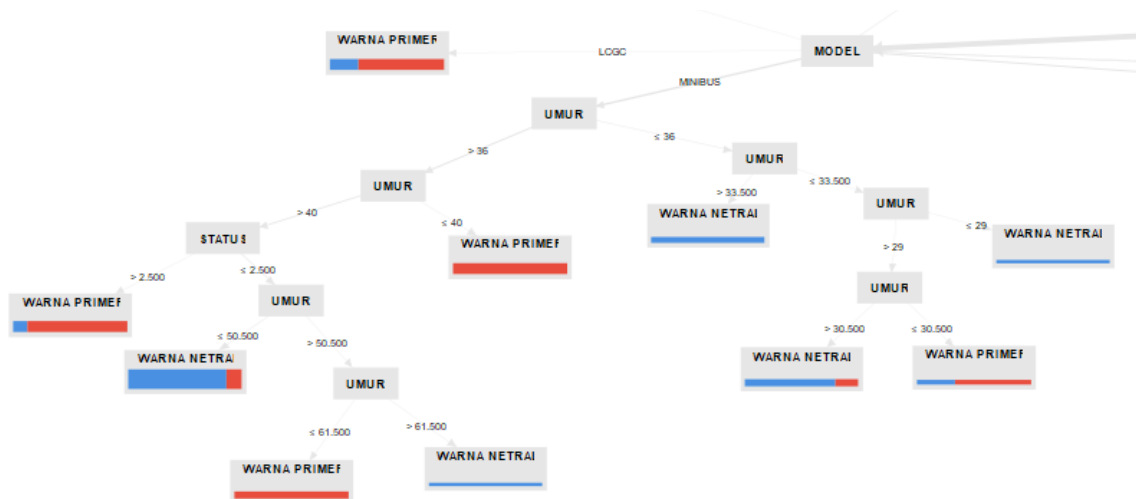
Gambar 8. Pohon Keputusan Variabel Model CITY CAR dengan Status ≥ 1.5

Gambar 8 menunjukkan umur di bawah 36,5 tahun cenderung memilih warna netral, sedangkan pada kelompok umur 36,5 hingga 56 tahun terdapat perbedaan preferensi berdasarkan jenis kelamin. Laki-laki lebih banyak memilih warna netral, sementara perempuan cenderung memilih warna primer. Responden dengan umur di atas 56 tahun umumnya memilih warna netral.



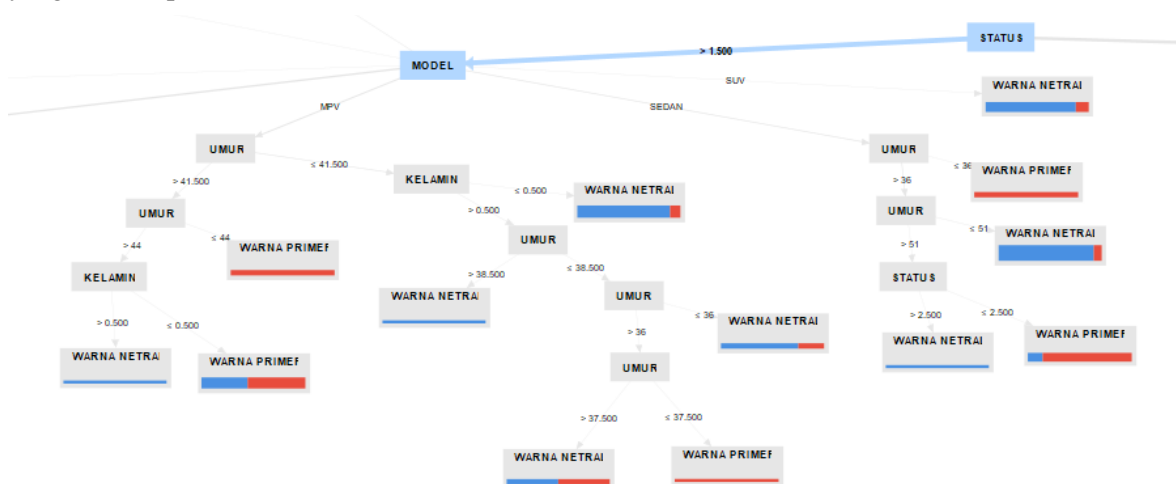
Gambar 9. Pohon Keputusan Variabel Model JEEP L.C.HDTP S.C HDTP dengan Status ≥ 1.5

Pola yang terbentuk berdasarkan Gambar 9 menunjukkan bahwa umur menjadi faktor dominan dalam menentukan preferensi warna kendaraan pada model ini. Umur di bawah 40 tahun lebih cenderung memilih warna netral, sedangkan yang berumur di atas 40 tahun menunjukkan kecenderungan pada warna primer. Namun pada rentang umur 40 – 45,5 tahun dan di atas 45 tahun preferensi kembali ke warna netral baik untuk laki-laki maupun perempuan.



Gambar 10. Pohon Keputusan Variabel Model LCGC dan MINIBUS dengan Status ≥ 1.5

Pada Gambar 10 menunjukkan untuk kelompok umur di atas 40 tahun dipengaruhi oleh variabel status, di mana status lebih tinggi menunjukkan kecenderungan terhadap warna primer, sedangkan yang lain tetap memilih warna netral.



Gambar 11. Pohon Keputusan Variabel Model MPV, SEDAN dan SUV dengan Status ≥ 1.5

Hasil pada Gambar 11 menunjukkan faktor umur tetap menjadi variabel yang paling berpengaruh, diikuti oleh jenis kelamin dan status. Pada model MPV, umur di bawah 41,5 tahun cenderung memilih warna primer, terutama kelompok laki-laki, sedangkan yang berumur lebih muda dominan memilih warna netral. Pada model Sedan dan SUV, pola preferensi warnanya terlihat lebih bervariasi. Umur ≤ 36 tahun lebih memilih warna netral, sementara yang berumur antara 36 – 51 tahun cenderung tertarik pada warna primer, khususnya untuk kelompok dengan status sosial lebih tinggi. Secara keseluruhan, ini memperkuat hasil pada Gambar 10, bahwa meskipun umur tetap menjadi faktor utama, kombinasi antara jenis kelamin dan status sosial mulai memperlihatkan pengaruh yang signifikan dalam menentukan pilihan warna, terutama pada jenis kendaraan dengan segmen menengah ke atas seperti MPV, Sedan, dan SUV.

3.3. Hasil Description Tree

```
STATUS > 1.500
|
| MODEL = CITY CAR
| |
| | UMUR > 36.500
| | |
| | | UMUR > 56: WARNA NETRAL {WARNA NETRAL=2, WARNA PRIMER=0}
| | | UMUR ≤ 56
| | | |
| | | | KELAMIN > 0.500: WARNA NETRAL {WARNA NETRAL=2, WARNA PRIMER=0}
| | | | KELAMIN ≤ 0.500: WARNA PRIMER {WARNA NETRAL=0, WARNA PRIMER=4}
| | | UMUR ≤ 36.500: WARNA NETRAL {WARNA NETRAL=2, WARNA PRIMER=0}
|
| MODEL = JEEP L.C.HDTP S.C.HDTP
| |
| | KELAMIN > 0.500
| | |
| | | UMUR > 45.500: WARNA NETRAL {WARNA NETRAL=4, WARNA PRIMER=0}
| | | UMUR ≤ 45.500
| | | |
| | | | UMUR > 40: WARNA PRIMER {WARNA NETRAL=0, WARNA PRIMER=2}
```



```

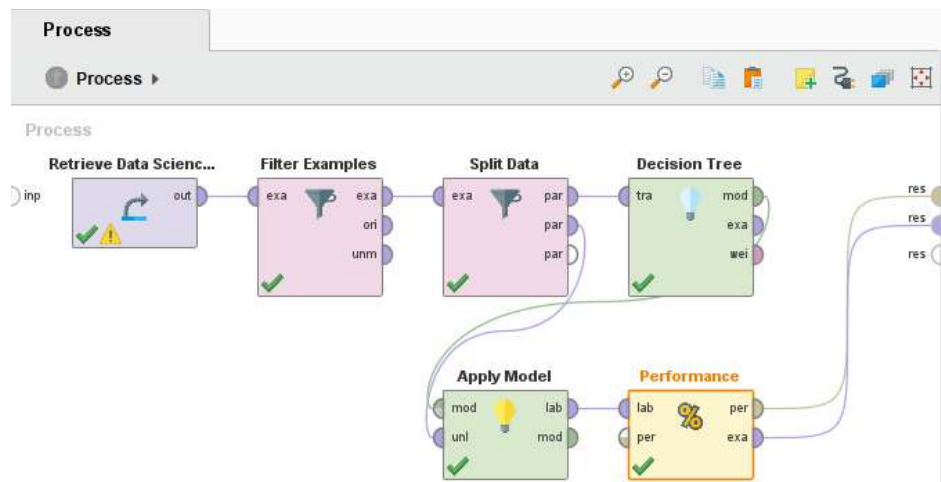
UMUR ≤ 40: WARNA NETRAL {WARNA NETRAL=3, WARNA PRIMER=0}
KELAMIN ≤ 0.500: WARNA NETRAL {WARNA NETRAL=7, WARNA PRIMER=0}
MODEL = LCGC: WARNA PRIMER {WARNA NETRAL=2, WARNA PRIMER=6}
MODEL = MINIBUS
  UMUR > 36
    UMUR > 40
      STATUS > 2.500: WARNA PRIMER {WARNA NETRAL=1, WARNA PRIMER=7}
      STATUS ≤ 2.500
        UMUR > 50.500
          UMUR > 61.500: WARNA NETRAL {WARNA NETRAL=2, WARNA PRIMER=0}
          UMUR ≤ 61.500: WARNA PRIMER {WARNA NETRAL=0, WARNA PRIMER=5}
        UMUR ≤ 50.500: WARNA NETRAL {WARNA NETRAL=13, WARNA PRIMER=2}
      UMUR ≤ 40: WARNA PRIMER {WARNA NETRAL=0, WARNA PRIMER=8}
    UMUR ≤ 36
      UMUR > 33.500: WARNA NETRAL {WARNA NETRAL=4, WARNA PRIMER=0}
      UMUR ≤ 33.500
        UMUR > 29
          UMUR > 30.500: WARNA NETRAL {WARNA NETRAL=4, WARNA PRIMER=1}
          UMUR ≤ 30.500: WARNA PRIMER {WARNA NETRAL=1, WARNA PRIMER=2}
        UMUR ≤ 29: WARNA NETRAL {WARNA NETRAL=2, WARNA PRIMER=0}
MODEL = MPV
  UMUR > 41.500
    UMUR > 44
      KELAMIN > 0.500: WARNA NETRAL {WARNA NETRAL=2, WARNA PRIMER=0}
      KELAMIN ≤ 0.500: WARNA PRIMER {WARNA NETRAL=4, WARNA PRIMER=5}
    UMUR ≤ 44: WARNA PRIMER {WARNA NETRAL=0, WARNA PRIMER=4}
  UMUR ≤ 41.500
    KELAMIN > 0.500
      UMUR > 38.500: WARNA NETRAL {WARNA NETRAL=2, WARNA PRIMER=0}
      UMUR ≤ 38.500
        UMUR > 36
          UMUR > 37.500: WARNA NETRAL {WARNA NETRAL=2, WARNA PRIMER=2}
          UMUR ≤ 37.500: WARNA PRIMER {WARNA NETRAL=0, WARNA PRIMER=2}
        UMUR ≤ 36: WARNA NETRAL {WARNA NETRAL=3, WARNA PRIMER=1}
      KELAMIN ≤ 0.500: WARNA NETRAL {WARNA NETRAL=9, WARNA PRIMER=1}
MODEL = SEDAN
  UMUR > 36
    UMUR > 51
      STATUS > 2.500: WARNA NETRAL {WARNA NETRAL=2, WARNA PRIMER=0}
      STATUS ≤ 2.500: WARNA PRIMER {WARNA NETRAL=1, WARNA PRIMER=6}
    UMUR ≤ 51: WARNA NETRAL {WARNA NETRAL=12, WARNA PRIMER=1}
    UMUR ≤ 36: WARNA PRIMER {WARNA NETRAL=0, WARNA PRIMER=4}
MODEL = SUV: WARNA NETRAL {WARNA NETRAL=7, WARNA PRIMER=1}
STATUS ≤ 1.500
  UMUR > 29.500
    MODEL = CITY CAR: WARNA NETRAL {WARNA NETRAL=2, WARNA PRIMER=0}
    MODEL = JEEP L.C.HDTP S.C.HDTP
      UMUR > 45
        UMUR > 56: WARNA NETRAL {WARNA NETRAL=2, WARNA PRIMER=0}
        UMUR ≤ 56: WARNA PRIMER {WARNA NETRAL=0, WARNA PRIMER=2}
      UMUR ≤ 45: WARNA NETRAL {WARNA NETRAL=9, WARNA PRIMER=0}
    MODEL = LCGC: WARNA NETRAL {WARNA NETRAL=2, WARNA PRIMER=0}
    MODEL = MINIBUS
      UMUR > 35: WARNA NETRAL {WARNA NETRAL=15, WARNA PRIMER=0}
      UMUR ≤ 35
        UMUR > 30.500
          KELAMIN > 0.500: WARNA NETRAL {WARNA NETRAL=3, WARNA PRIMER=1}
          KELAMIN ≤ 0.500: WARNA PRIMER {WARNA NETRAL=0, WARNA PRIMER=2}
        UMUR ≤ 30.500: WARNA NETRAL {WARNA NETRAL=2, WARNA PRIMER=0}
      MODEL = MPV: WARNA NETRAL {WARNA NETRAL=8, WARNA PRIMER=1}
      MODEL = SEDAN: WARNA NETRAL {WARNA NETRAL=14, WARNA PRIMER=0}
      MODEL = SUV
        UMUR > 42
          UMUR > 52.500: WARNA PRIMER {WARNA NETRAL=0, WARNA PRIMER=2}
          UMUR ≤ 52.500: WARNA NETRAL {WARNA NETRAL=1, WARNA PRIMER=1}
        UMUR ≤ 42: WARNA NETRAL {WARNA NETRAL=5, WARNA PRIMER=0}
  UMUR ≤ 29.500
    UMUR > 24.500
      KELAMIN > 0.500: WARNA PRIMER {WARNA NETRAL=2, WARNA PRIMER=6}

```

				KELAMIN ≤ 0.500
				UMUR > 26.500
				UMUR > 27.500 : WARNA NETRAL {WARNA NETRAL=5, WARNA PRIMER=2}
				UMUR ≤ 27.500 : WARNA PRIMER {WARNA NETRAL=1, WARNA PRIMER=2}
				UMUR ≤ 26.500 : WARNA NETRAL {WARNA NETRAL=4, WARNA PRIMER=0}
				UMUR ≤ 24.500 : WARNA NETRAL {WARNA NETRAL=4, WARNA PRIMER=0}

3.4. Hasil Pengujian Performance

Hasil pengujian untuk melihat model validasi *performance* ditunjukkan pada Gambar 12 dibawah ini.



Gambar 12. Model Uji Performance Algoritma Decision Tree

accuracy: 70.31%

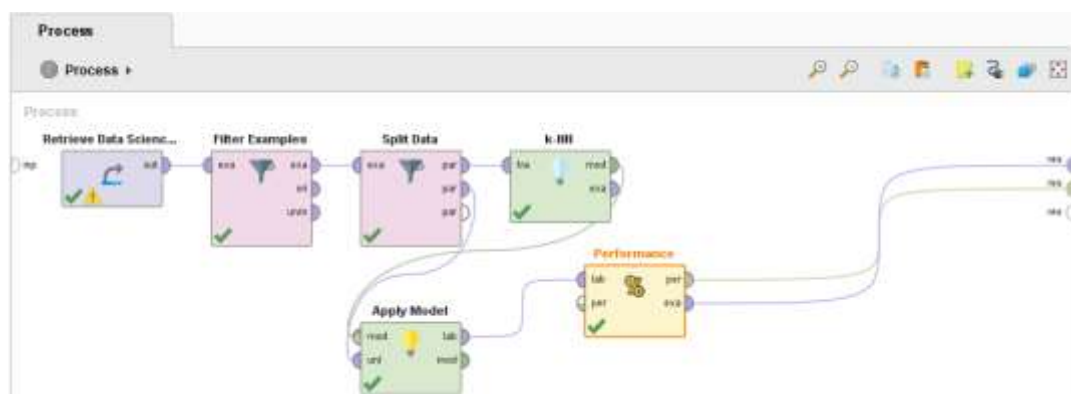
	true WARNA NETRAL	true WARNA PRIMER	class precision
pred. WARNA NETRAL	35	11	76.09%
pred. WARNA PRIMER	8	10	55.56%
class recall	81.40%	47.52%	

Gambar 13. Hasil Uji Performance Algoritma Decision Tree

Berdasarkan dari Gambar 12 dan Gambar 13 diatas dengan *Criterion Information Gain* yang digunakan, dapat ditentukan nilai *accuracy* paling tinggi didapatkan dengan nilai 70.31%.

3.5. Model Algoritma K-NN

Berikut adalah proses pengolahan dataset dengan membuat model *training* dan uji menggunakan bantuan *software* Rapidminer dengan Algoritma *Decision Tree*.



Gambar 14. Model Algoritma k-NN

3.6. Hasil Pengujian Performance

Pada pengujian menggunakan Algoritma K-NN dengan 3 buah K dapat dilihat pada Tabel 2 sebagai berikut:

Tabel 2. Hasil Performa K-NN

Performa K-NN	
K	Accuracy
4	73.44%
6	78.12%
8	73.44%



Gambar 15. Hasil Uji Performance Algoritma K-NN dengan K=6

Berdasarkan Tabel 2 dapat dilihat bahwa nilai *accuracy* tertinggi dalam pengujian menggunakan Algoritma K-NN ketika nilai $K=6$, yaitu 78.12% dengan hasil uji performancenya pada Gambar 15. Nilai ini lebih tinggi dibandingkan dengan pengujian pada $K = 4$ dan $K = 8$, yang masing-masing hanya mencapai 73,44%. Kinerja K-NN unggul disebabkan oleh karakteristik data yang relatif terstruktur dan memiliki hubungan jarak antar atribut yang jelas, seperti umur, status (perkawinan), jenis kelamin, dan model kendaraan. Karena K-NN bekerja berdasarkan kedekatan jarak antar data, maka ketika distribusi data tidak terlalu bervariasi dan tidak banyak outlier, algoritma ini mampu melakukan klasifikasi dengan cukup akurat.

Namun K-NN memiliki keterbatasan yaitu sensitif terhadap skala data dan outlier, sehingga hasil prediksi bisa berubah jika proses normalisasi tidak tepat. Untuk pengembangan di masa mendatang, performa model dapat ditingkatkan dengan optimasi nilai K menggunakan teknik *cross-validation* untuk mendapatkan nilai K yang paling stabil. Serta melakukan *feature selection* untuk mengidentifikasi atribut paling berpengaruh terhadap hasil klasifikasi warna kendaraan.

4. Kesimpulan

Berdasarkan pengujian menggunakan dataset Calon Pembeli Mobil didapatkan bahwa pengujian pola pemilihan warna pada mobil dengan menggunakan Algoritma *Decision Tree*, dengan data 80% untuk data pelatihan dan 20% untuk data uji adalah hasil dari uji coba ini memberikan hasil *accuracy* dengan menggunakan *Criteria Information Gain* dengan nilai 70.31%, sedangkan ketika menggunakan Algoritma K-NN dengan $K=6$ dan proporsi 80% data pelatihan, 20% untuk data uji mendapatkan hasil *accuracy* 78.12% . Oleh karena itu dari pengujian menggunakan 2 algoritma dapat disimpulkan bahwa pengujian menggunakan algoritma K-NN merupakan algoritma yang terbaik dibandingkan dengan Algoritma *Decision Tree*. Hasil yang didapatkan oleh Algoritma K-NN tingkat kebenarannya sudah mendekati 100%. Sehingga Pola pemilihan warna menggunakan Algoritma K-NN dapat digunakan sebagai bahan pertimbangan perusahaan atau dealer mobil untuk mengukur kuantitas produksi berdasarkan warna dan mengukur stok persediaan mobil di dealer, sehingga mengurangi biaya penyimpanan karena *dead stock* atau barang menumpuk karena proses *outbound* yang lama.

References

- [1] M. Afrizki, I. P. Jadmiko, A. H. Nugroho, R. F. Irsyadi, and N. D. Anggraeni, "Pengaruh Keahlian Berbicara Sales dalam Mendorong Tingkat Penjualan Mobil Dealer Toyota pada Bidang Industri Otomotif Masa Kini," *Jurnal Bisnis dan Komunikasi Digital*, vol. 1, no. 4, p. 12, 2024.

- [2] Y. Erlyana, S. Everlin, and I. F. Yuwono, "Analisis Color Palette Berdasarkan Rasa Warna Sebagai Penguat Daya Tarik Emosional Dalam Video Anak," 2023. [Online]. Available: <http://publikasi.dinus.ac.id/index.php/andharupa/index>
- [3] W. Handayani, R. Ulum, and N. Khofia, "Psikologi Warna Dalam Kehidupan Sehari-Hari: Pengaruh Warna Terhadap Emosi, Persepsi, Dan Perilaku Konsumen," *PSIKIS: Jurnal Ilmu Psikiatri Dan Psikologi*, vol. 1, no. 1, pp. 39–48, 2025.
- [4] S. N. Fadijah and Satriadi, "Peran Warna Dalam Meningkatkan Daya Tarik Visual Logo," *PARATIWI: Jurnal Seni Rupa dan Desain*, vol. 3, 2024.
- [5] X. Zhou, C. Xiao, S. Yoon, and H. Zhu, "The color of status: color saturation, brand heritage, and perceived status of luxury brands," *Journal of Consumer Research*, p. ucaf029, 2025.
- [6] I. Sharon Gauci, "The Importance and Evolution of Design at General Motors," *Innovation Summer*, 2020.
- [7] A. Muhiban and E. K. Putri, "Pengaruh Tampilan Produk dan Electronic Word of Mouth Terhadap Keputusan Pembelian Pada E-commerce Shopee (Studi Kasus Konsumen Shopee di PT. Gucci Ratu Textile Kota Cimahi)," *Jurnal EMT KITA*, vol. 6, no. 2, 2022.
- [8] Axalta Coating System, "Axalta Global Automotive Color Popularity Report 2024," 2024. Accessed: Oct. 29, 2025. [Online]. Available: https://www.axalta.com/content/dam/New%20Axalta%20Corporate%20Website/Public/Documents/US/axalta-global-automotive-color-popularity_2024.pdf
- [9] H. S. S. Rubianto, U. Ristian, J. Rekayasa Sistem Komputer, and F. H. MIPA Universitas Tanjungpura Jalan Hadari Nawawi Pontianak, "Penerapan Metode Decision Tree Untuk Mangklasifikasikan Mutu Jeruk Berdasarkan Fitur Warna dan Ukuran," 2021.
- [10] B. H. Mawaridi and M. Faisal, "Rekomendasi Merk Mobil Untuk Calon Pembeli Menggunakan Algoritma Decision Tree," *Jurnal Informatika*, vol. 10, no. 2, pp. 157–162, Oct. 2023, doi: 10.31294/inf.v10i2.16000.
- [11] N. W. Mardiyah, N. Rahaningsih, and I. Ali, "Penerapan data mining menggunakan algoritma k-nearest neighbor pada prediksi pemberian kredit di sektor finansial," *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 8, no. 2, pp. 1491–1499, 2024.
- [12] V. R. Prasetyo, H. Lazuardi, A. A. Mulyono, and C. Lauw, "Penerapan Aplikasi RapidMiner Untuk Prediksi Nilai Tukar Rupiah Terhadap US Dollar Dengan Metode Regresi Linier," *Jurnal Nasional Teknologi dan Sistem Informasi (TEKNOSI)*, vol. 7, no. 1, pp. 8–17, 2021.
- [13] B. G. Sudarsono, M. I. Leo, A. Santoso, and F. Hendrawan, "Analisis Data Mining Data Netflix Menggunakan Aplikasi Rapid Miner," *JBASE-Journal of Business and Audit Information Systems*, vol. 4, no. 1, 2021.
- [14] Y. F. S. Y. Damanik, S. Sumarno, I. Gunawan, D. Hartama, and I. O. Kirana, "Penerapan Data Mining Untuk Pengelompokan Penyebaran Covid-19 Di Sumatera Utara Menggunakan Algoritma K-Means," *Jurnal Ilmu Komputer Dan Informatika*, vol. 1, no. 2, 2021.
- [15] H. C. Husada and A. S. Paramita, "Analisis Sentimen Pada Maskapai Penerbangan di Platform Twitter Menggunakan Algoritma Support Vector Machine (SVM)," *Teknika*, vol. 10, no. 1, pp. 18–26, 2021.
- [16] A. Fatwanto, M. Nur Aslam, R. Ndugi, and M. Syafrudin, "An Investigation Towards Resampling Techniques and Classification Algorithms on CM1 NASA PROMISE Dataset for Software Defect Prediction," *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, vol. 8, no. 5, pp. 631–643, Oct. 2024, doi: 10.29207/resti.v8i5.5910.