

Implementasi Data Mining Untuk Penjualan Mobil Menggunakan Metode Naive Bayes

M.Luthfi Rifky¹, Zahlan Nugraha², M.Bayu Saputra³, Dimas Pratama⁴, Errissya Raswir^{5*}, Yovi Pratama⁶

¹²³Fakultas Ilmu Komputer, Program Studi Teknik Informatika, Universitas Dinamika Bangsa, Kota Jambi, indonesiaEmail :

¹luthfikiki03@gmail.com ²zahlan@gmail.com ³bayaww18@gmail.com ⁴dimasdimasputra909@gmail.com,

⁵errissya.raswir@gmail.com, ⁶yovi.pratama@gmail.com

Email Penulis Korespondensi : ⁵errissya.raswir@gmail.com

Abstrak- Mobil adalah salah satu kendaraan yang paling sering dijumpai dengan berbagai *type* dan merek. Mobil memiliki spesifikasi yang beraneka ragam. Metode *Naive Bayes* adalah salah satu metode klasifikasi dan percabangan dari *artificial intelligence*. Berbagai merek tersebut akan dibentuk suatu *Classify* yaitu Laris dan Tidak Laris, sehingga para konsumen, produsen, dan peneliti dapat mengetahui merek mobil manakah yang paling laris berdasarkan kategori maupun *output*-nya. *Naive bayes* merupakan metode klasifikasi yang banyak digunakan karena sederhana dan akurasi yang tinggi dalam mengklasifikasi data. Penelitian ini menganalisis data sebanyak 639 data menjadi 511 data training dan 128 data *testing*, data ini didapatkan dari situs “Gabungan Industri Kendaraan Bermotor Indonesia” (GAIKINDO). Dengan atribut 19, untuk memudahkan penulis dalam penelitian, maka atribut yang digunakan adalah 8 (termasuk 1 atribut *Classify* yang ditambahkan penulis untuk memudahkan dalam pencarian mobil terlaris). Hasil penelitian yang dilakukan memberikan klasifikasi pembeda merek mobil yang paling banyak diminati para konsumen dan Kategori yang paling laris. Tingkat akurasi klasifikasi dengan Metode *Naive Bayes* menghasilkan nilai *accuracy* sebesar 92,19%, nilai *Precision*: 98,39% dan nilai *Recall*: 87,14% sehingga Metode *Naive Bayes* merupakan metode yang cukup baik dalam penelitian ini.

Kata kunci: Brand Mobil, Klasifikasi, *Naive Bayes*, *Accuracy*

Abstract- Cars are one of the vehicles most often found in various types and brands. Carshave various specifications. The Naive Bayes method is one of the classification and branching methods of artificial intelligence. The various brands will be in the form of a Class that are Selling and Not Selling, so that consumers, producers, and researchers can find out which car brands are best selling based on their category and output. Naive bayes is a widely used classification method because of its simple and high accuracy in classifying data. This study analyzed data as many as 639 data into 511 training data and 128 testing data, data was obtained from “the Indonesian Automotive Industries Association” (IAIS) site. With attribute 19, to facilitate the writer in the study the attributes used were 8 (including 1 Class attribute that author added to facilitate the search for the best-selling car). The results of the conducted research gave the classification of car brand different that were most in demand by consumers and the best-selling categories. The level of classification accuracy with the Naive Bayes Method produces accuracy values of 92, 19%, Precision values: 98, 39% and Recall values: 87, 14% so that the Naive Bayes Method is a pretty good method in this research.

Keywords: Car Brand, Classification, Naive bayes, Accuracy

1. PENDAHULUAN

Kendaraan adalah alat transportasi yang sangat dibutuhkan dalam kehidupan. dari berbagai macam kendaraan, mobil adalah salah satu kendaraan roda empat yang paling banyak dicari oleh para konsumen setelah motor. Sesuai kemajuan jaman, kini merek mobil-pun semakin banyak. Berdasarkan spesifikasinya, mobil memiliki jenis atau varian yang beraneka ragam. Dalam analisis ini spesifikasi yang digunakan hanya delapan atribut untuk memudahkan implementasi pada *Rapid Miner*. Spesifikasi yang digunakan dalam data tersebut yaitu *Category*, *Brand*, *Type Model*, *CC*, *Transmisi (Trans)*, *Origin Country*, *Output*, dan *Class*. Spesifikasi tersebut diambil penulis dari sebuah data yang didapatkan dari situs Gabungan Industri Kendaraan Bermotor Indonesia (GAIKINDO).

Banyaknya merek mobil yang terjun di industri pasar otomotif, namun pihak GAIKINDO sendiri dalam situsnya tidak memberikan klasifikasi pembeda manakah merek (Brand) mobil yang paling banyak diminati para konsumen dan kategori manakah yang paling laris. Oleh sebab itu diperlukan suatu klasifikasi pembeda, agar para konsumen mengetahui merek mobil yang terlaris dari segi kategori yang paling diminati. Selain itu memacu para

produsen mobil dalam memproduksi mobil sesuai yang paling banyak diminati para konsumen[1].XXXXXXXXXXXXXXXXXX

Metode *Naive Bayes* adalah salah satu metode klasifikasi dan percabangan dari *artificial intelligence*. *Artificial intelligence* yaitu kemampuan sistem untuk menafsirkan data eksternal dengan benar, untuk belajar dari data dan menggunakan pembelajaran tersebut guna mencapai tujuan [2].

Oleh karena itu, bermacam merek dalam penelitian ini akan dibentuk suatu *Class* yaitu Laris dan Tidak Laris, sehingga para konsumen, produsen, dan peneliti dapat mengetahui merek mobil manakah yang paling laris berdasarkan kategori maupun *output*-nya [3]. Berdasarkan pada permasalahan diatas peneliti melakukan penelitian tentang “Implementasi Data Mining Untuk Penjualan Mobil Menggunakan Metode Naive Bayes” Implementasi metode *Naive Bayes* dapat menghasilkan akurasi yang maksimal dengan data latih yang sedikit.

2. METODOLOGI PENELITIAN

Pada Metode penelitian ini menjelaskan tentang naive bayes, analisis penerapan *naive bayes*, pengumpulan data mobil pada GAIKINDO, data penjualan mobil GAIKINDO, skenario pengambilan data pada GAIKINDO, seleksi data dan penerapan teknik mining Menggunakan metode *naive bayes*.

2.1 Naive Bayes

Naive Bayes adalah metode yang cocok untuk klasifikasi biner dan *multiclass*[4]. Metode yang juga dikenal sebagai *Naive Bayes Classifier* ini menerapkan teknik *supervised* klasifikasi objek di masa depan dengan menetapkan label kelas ke *instance*/catatan menggunakan probabilitas bersyarat [5]. Probabilitas bersyarat adalah ukuran peluang suatu peristiwa yang terjadi berdasarkan peristiwa lain yang telah (dengan asumsi, praduga, pernyataan, atau terbukti) terjadi [6].

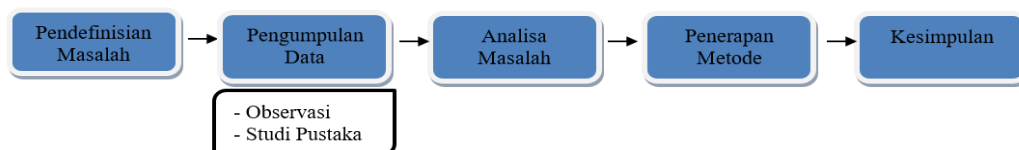
2.2 Analisis Penerapan Naive Bayes

Untuk memperoleh suatu kesimpulan dari hasil pengumpulan data maka dilakukan analisis terhadap semua data yang terkumpul dan metode yang digunakan sesuai kebutuhan [7].

2.2 Pengumpulan Data Mobil GAIKINDO

Penelitian dilakukan terhadap data set yang diperoleh dari situs GAIKINDO yang berfokus pada bulan Januari, Februari, Maret, dan April 2018, dengan jumlah data 639. Metode pengumpulan (link) data (Gambar 1) yang dilakukan dalam penelitian ini adalah [1].

1. Pendefinisian Masalah : Mengklarifikasi karakteristik situasi dari penjualan brand mobil GAIKINDO
2. Pengumpulan Data : mencari data yang diperlukan untuk penjualan brand mobil GAIKINDO, ada 2 cara untuk mengumpulkan data ini yaitu :
 - Observasi : mengumpulkan data yang dilakukan dengan pengamatan dan pencatatan secara langsung maupun tidak langsung terhadap objek yang akan dibahas.
 - Studi Pustaka : melakukan studi terhadap *literature* (buku-buku, jurnal, dan skripsi) dan mempelajari referensi lain yang berkaitan dengan masalah yang akan dibahas untuk dapat dijadikan acuan dalam analisis ini. Dan juga mencari dan mengumpulkan data-data yang diperlukan dari GAIKINDO
3. Analisa Masalah : menganalisa masalah dan memahami permasalahan mana saja yang kritis pada penjualan brand mobil GAIKINDO
4. Penerapa metode : memahami banyaknya data pelanggan dan menawarkan penjualan brand mobil GAIKINDO
5. Kesimpulan : ringkasan analisis dari seluruh bab

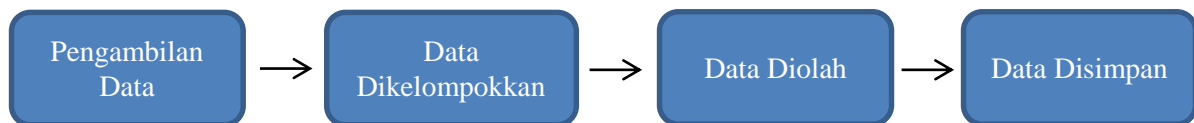


Gambar 1. Alur Analisis

2.3 Data Penjualan Mobil GAIKINDO

Data yang digunakan nantinya adalah semua data dari *Category*, *Brand*, dan *Type Model* sebanyak 639 data, yang diperoleh dari situs GAIKINDO. Data tersebut nantinya akan dikelompokkan menjadi 8 atribut termasuk atribut kelas. Variabel yang digunakan yaitu *Category*, *Brand*, *Type Model*, *CC*, *Transmisi (Trans)*, *Output*, dan *Class*. Klasifikasi dilakukan dengan menerapkan Algoritma *Naive Bayes*[1].

2.4 Skenario Pengambilan Data Pada GAIKINDO

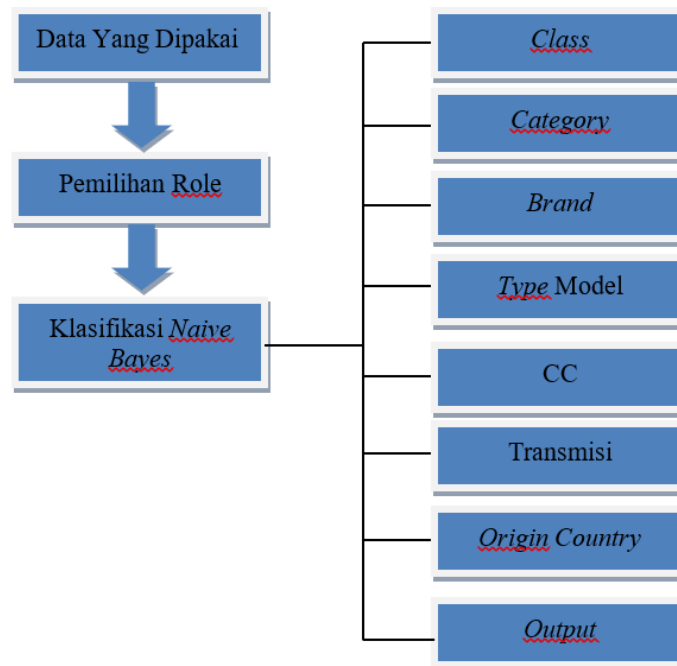


Skenario dalam pengambilan data:

1. Pengambilan data diambil yang diperoleh dari situs GAIKINDO.
2. Data dikelompokkan menjadi 8 atribut berdasarkan data aslinya.
3. Data diolah sedemikian rupa dengan meminimalkan/menghilangkan beberapa atribut.
4. Data disimpan dalam file berekstensi .xlsx.

A. Seleksi Data

Ada tahap ini akan dilakukan penyeleksian terhadap data yang kurang relevan terhadap penelitian (menghilangkan atau menghapus data)[8]. Alur kerja *Naive Bayes* dapat dilihat pada Gambar 3. Setelah diseleksi, dilakukan penggabungan seluruh data yang telah diperoleh. Setelah itu data akan disimpan dalam satu file dengan ekstensi .xlsx berdasarkan variabel sesuai dengan kriteria atributnya[9].



Gambar 2. Alur Kerja *Naive Bayes*

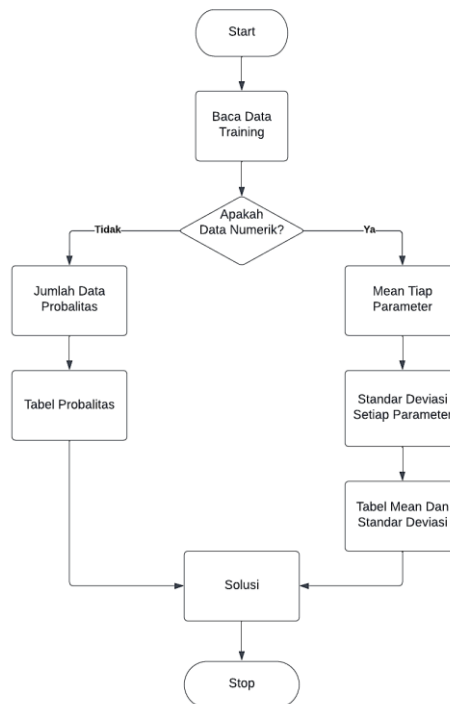
B. Penerapan Teknik Mining

Pada tahap ini data penjualan brand mobil GAIKINDO menjadi data yang sangat relevan, data dari berbagai merek akan diklasifikasikan ke beberapa tipe dengan Algoritma *Naive Bayes*. [1].

3. HASIL PEMBAHASAN

3.1 Implementasi Algoritma *Naive Bayes*

Pada bagian ini Ketika data siap di proses, algoritma *Naive Bayes* memproses data [10] pada gambar dibawah ini



Gambar 3. Tahapan Algoritma *Naive Bayes*

Data penjualan mobil ditampilkan dalam tabel dibawah ini:

Tabel 1. Training Set Yang Telah Diolah

Category	Brand	Type Model	CC	Transmisi	Origin Country	Output	Class
Sedan Type	Honda	All-New City IVTECE AT	1497	AT	Thailand	364	LARIS
Sedan Type	Honda	All-New City IVTECE AT	1497	AT	Thailand	0	TIDAK LARIS
Sedan Type	Honda	All-New Civic	1500	AT	Thailand	403	LARIS
Sedan Type	Honda	All-New Civic Prestige	1500	AT	Thailand	422	LARIS
Sedan Type	Toyota	Vios-E	1500	MT	Indonesia	4	TIDAK LARIS
Sedan Type	Toyota	Vios-E AT	1500	AT	Indonesia	5	TIDAK LARIS
Sedan Type	Toyota	Vios-E G	1500	MT	Indonesia	43	TIDAK LARIS
Sedan Type	Toyota	Vios-E G AT	1500	AT	Indonesia	61	TIDAK LARIS
Sedan Type	Audi	A4 2.0 Quattro TFSI	1984	AT	Germany	1	TIDAK LARIS
Sedan Type	Audi	A5 2.0 TFSI	1984	AT	Germany	0	TIDAK LARIS

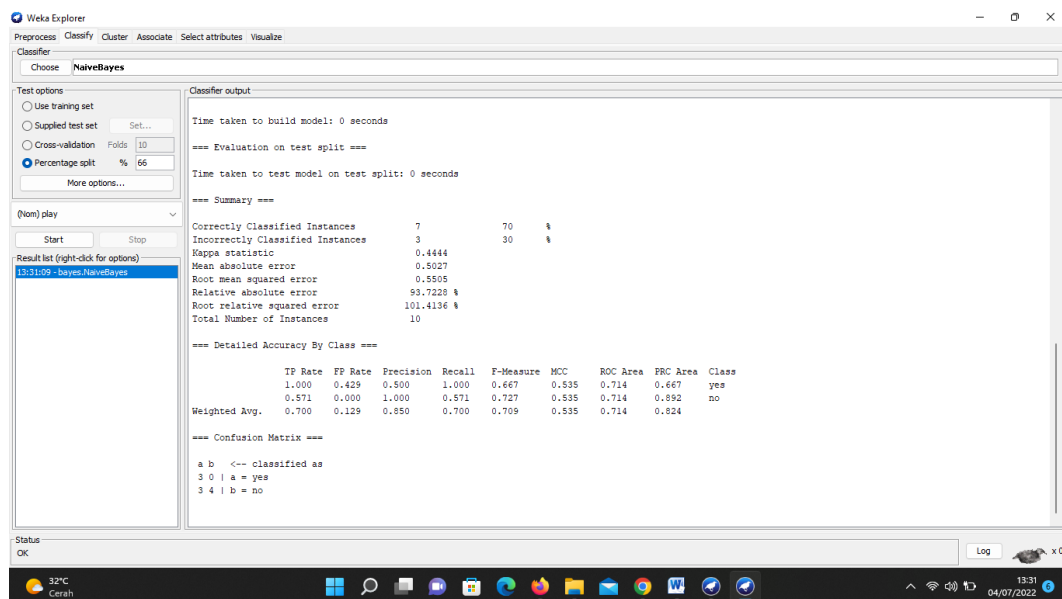
Tabel 2. Testing Set yang telah diolah

Category	Brand	Type Model	CC	Transmisi	Origin	Output	Class
----------	-------	------------	----	-----------	--------	--------	-------

	Country	DOI	XXXXXXXXXXXXXXXXXXXX			
4x2 TYPE	Chevrolet	Trailblazer	2500	AT	Thailand	172
4x2 TYPE	Honda	All-New CR-V	2000	AT	Indonesia	61
4x2 TYPE	Honda	CIVIC 5D S	2400	AT	Thailand	60
4x2 TYPE	Honda	CIVIC 5D S	2400	AT	Thailand	1035
4x2 TYPE	Honda	HR-V Prestige	1800	CVT	Indonesia	1731
4x2 TYPE	Honda	Odyssey Prestige	2354	AT	Japan	101
4x2 TYPE	Hyundai	All-New Tucson	2000	MT	Korea	37
4x2 TYPE	Hyundai	All-New Tucson CRDI	2000	AT	Korea	15
4x2 TYPE	Hyundai	New Santa FE CRDI	2200	AT	Korea	33

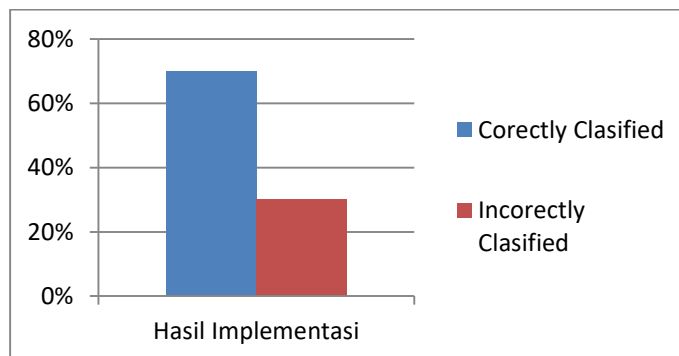
3.2 Implementasi Klasifikasi Naive Bayes Pada Weka

Dengan Menggunakan *tools* Weka, data akan diuji seperti pada gambar 4 berikut :



Gambar 4. Klasifikasi Metode Naive Bayes

Berdasarkan Gambar 4 diatas dapat dilihat persentase untuk *correctly classified instance* adalah sebesar 70% sementara persentase untuk *Incorrectly Classified Instance* adalah sebesar 30%. Dimana 10 data dari penjualan mobil ada 7 data penjualan mobil yang berhasil diklasifikasikan dengan benar dan sebanyak 3 data dari penjualan mobil tidak berhasil diklasifikasikan benar.



Gambar 5. Hasil Data Implementasi

Dapat disimpulkan bahwa *corectly clasified* yang didapatkan adalah 70% sedangkan untuk *Incorectly Clasified* yang didapatkan adalah 30%.

4. KESIMPULAN

Berdasarkan penelitian tentang memprediksi penjualan mobil dapat ditarik beberapa kesimpulan yaitu metode *Naive Bayes* memanfaatkan data training untuk menghasilkan probabilitas setiap kriteria untuk *class* yang berbeda, sehingga nilai-nilai probabilitas dari kriteria tersebut dapat dioptimalkan untuk memprediksi penjualan mobil berdasarkan proses klasifikasi yang dilakukan oleh metode *Naive Bayes* itu sendiri. Berdasarkan data penjualan mobil yang dijadikan data training, metode *Naive Bayes* berhasil mengklasifikasikan 7 data dari 10 data yang diuji. Sehingga metode *Naive Bayes* berhasil memprediksi besarnya penjualan mobil dengan persentase keakuratan sebesar 70%

REFERENCES

- [1] I. Romli, E. Pusnawati, and U. P. Bangsa, "Penentuan Tingkat Penjualan Mobil Di Indonesia," *Pros. Semin. Nas. Teknol. dan Sains*, vol. x, no. x, 2019.
- [2] R. Sari and R. Y. Hayuningtyas, "Penerapan Algoritma Naive Bayes Untuk Analisis Sentimen Pada Wisata TMII Berbasis Website," *Indones. J. Softw. Eng.*, vol. 5, no. 2, pp. 51–60, 2019, doi: 10.31294/ijse.v5i2.6957.
- [3] F. Rizki, A. Faisol, and F. Santi Wahyuni, "Penerapan Metode Naive Bayes Untuk Memprediksi Penjualan Pada Ud. Hikmah Pasuruan Berbasis Web," *JATI (Jurnal Mhs. Tek. Inform.)*, vol. 4, no. 1, pp. 26–34, 2020, doi: 10.36040/jati.v4i1.2379.
- [4] L. Indriyani and W. Susanto, "Analisis Penerapan Naive Bayes Untuk Memprediksi Resiko Kredit Anggota Koperasi Keluarga Guru," *J. Inform.*, vol. 6, no. 2, pp. 262–270, 2019, doi: 10.31311/ji.v6i2.5724.
- [5] Y. Prasiska, "Perancangan Aplikasi Rekam Medis Pada Uptd Puskesmas Durian Luncuk," *J. Inform. Dan Rekayasa Komput. ...*, vol. 1, no. April, pp. 90–98, 2022, [Online]. Available: <https://ejournal.unama.ac.id/index.php/jakakom/article/view/57>
- [6] J. Informatika, D. Rekayasa, K. Jakakom, H. Prastiwi, J. Pricilia, and E. Raswir, "Implementasi Data Mining Untuk Menentuksn Persediaan Stok Barang Di Mini Market Menggunakan Metode K-Means Clustering Jurnal Informatika Dan Rekayasa Komputer (JAKAKOM)," vol. 1, no. April, pp. 141–148, 2022.
- [7] J. H. Jaman, J. H. Jaman, . C., and A. R. Sanjaya, "Klasifikasi jenis mobil paling diminati Di indonesia menggunakan algoritma Naive bayes," *Fakt. Exacta*, vol. 13, no. 1, p. 18, 2020, doi: 10.30998/faktorexacta.v13i1.5573.
- [8] H. Annur, "Penerapan Data Mining Menentukan Strategi Penjualan Variasi Mobil Menggunakan Metode K-Means Clustering," *J. Inform. Upgris*, vol. 5, no. 1, 2019, doi: 10.26877/jiu.v5i1.3091.
- [9] Farhan Abie Ardandy, Immanuel Morries Pohan, Ariq Mitsal, Finandra Nusantara, and Muhammad Deka Ruliansyah, "Perbandingan Algoritma Naive Bayes dan Linear Discriminant Analysis dengan Dataset Car Evaluation," *J. Rekayasa Elektro Sriwij.*, vol. 3, no. 1, pp. 213–217, 2021, doi: 10.36706/jres.v3i1.45.
- [10] D. Darwis, N. Siskawati, and Z. Abidin, "Penerapan Algoritma Naive Bayes Untuk Analisis Sentimen Review Data Twitter Bmkg Nasional," *J. Tekno Kompak*, vol. 15, no. 1, p. 131, 2021, doi: 10.33365/jtk.v15i1.744.
- [11] W. Santoso, V. C. Mawardi, and T. Sutrisno, "Implementasi Aplikasi Jual Beli Mobil Bekas Dengan Metode Analytical Hierarchy Process Dan Naive Bayes," *J. Ilmu Komput. dan Sist. Inf.*, vol. 9, no. 1, p. 215, 2021, doi: 10.24912/jiksi.v9i1.11597.