

Penerapan Algoritma Naïve Bayes Untuk Prediksi Penjualan Motor Terlaris Pada PT. Ramayana Mitra Sejahtera

**Nazla Ainayah¹, Pera Anjela Laoly², Fadhiya Cintana Soleka Putri³, Mulyadi⁴,
Kartika Mariskhana⁵**

^{1,2,3,4,5} Sistem Informasi, Universitas Bina Sarana Informatika
Jakarta, Indonesia

e-mail: nazlaainayahh@gmail.com¹, peraanjela@gmail.com², fadhiyacintana@gmail.com³,
mulyadi.myd@bsi.ac.id⁴, kartika.kma@bsi.ac.id⁵

ABSTRAK

Penjualan adalah aktivitas penting yang dapat meningkatkan pendapatan perusahaan. Namun, PT. Ramayana Mitra Sejahtera masih mengandalkan sistem manual dan belum memiliki metode yang efektif untuk mengidentifikasi produk terlaris. Hal ini menjadi kendala dalam menghadapi ketidakstabilan penjualan yang dipengaruhi oleh permintaan konsumen yang beragam. Penelitian ini bertujuan untuk membantu perusahaan dalam menentukan produk motor yang paling diminati menggunakan Algoritma Naïve Bayes. Berdasarkan hasil evaluasi, pada model yang menggunakan 60% data training (805 data) dan 40% data testing (537 data) menunjukkan kinerja yang sangat baik dengan akurasi mencapai 94,97%. Hasil ini menunjukkan bahwa model dengan perbandingan 60% data training dan 40% data testing merupakan paling optimal dalam mengukur akurasi pada data penjualan sepeda motor di PT. Ramayana Mitra Sejahtera.

Kata Kunci: akurasi, data training dan testing, naïve bayes, penjualan, prediksi produk terlaris

ABSTRACTS

Sales is an important activity that can increase company revenue. However, PT. Ramayana Mitra Sejahtera still relies on a manual system and does not have an effective method for identifying best-selling products. This becomes an obstacle in dealing with sales instability which is influenced by varying consumer demand. This research aims to help companies determine the most popular motorbike products using the Naïve Bayes algorithm. Based on the evaluation results, the model that uses 60% training data (805 data) and 40% testing data (537 data) shows very good performance with an accuracy of 94.97%. These results show that a model with a ratio of 60% training data and 40% testing data is the most optimal in measuring accuracy in motorbike sales data at PT. Ramayana Mitra Sejahtera.

Keywords: accuracy, data training and testing, naïve bayes, sales, top product prediction

1. PENDAHULUAN

Sepeda motor menjadi pilihan utama bagi banyak individu sebagai sarana transportasi karena harganya yang terjangkau serta dapat digunakan dalam jangka waktu yang lama. Oleh karena itu, tidak mengherankan penjualan

sepeda motor terus menunjukkan angka yang signifikan dari tahun ke tahun. Penjualan adalah aktivitas yang dapat memberikan keuntungan lebih bagi usaha retail. Usaha retail ini masih menggunakan sistem manual dan belum memiliki metode yang efektif untuk menentukan



produk terlaris. Hal ini terjadi karena adanya peningkatan penjualan produk, yang disebabkan oleh permintaan konsumen yang sangat beragam. Dengan ratusan produk yang terjual, usaha ini mengalami kesulitan dalam mengecek struk belanja untuk menentukan produk terlaris berdasarkan data satu tahun terakhir. (Dewi et al., 2022) Demikian juga dengan perusahaan yang bergerak dibidang otomotif, penjualan suku cadang motor mengalami kesulitan dalam memprediksi fluktuasi penjualan, sehinggamodal yang terlalu besar dan penumpukan barang menyebabkan kerugian pada perusahaan. (Rezekika, 2020) Permasalahan serupa juga terjadi pada PT. Ramayana Mitra Sejahtera, dimana Perusahaan tersebut masih menggunakan cara manual untuk mengetahui produk motor apa yang paling laris. Oleh karena itu, peneliti ingin membantu Perusahaan dalam mengidentifikasi produk motor yang paling diminati oleh pelanggan dengan menggunakan metode Algoritma Naïve Bayes dengan menggunakan RapidMiner. RapidMiner adalah perangkat lunak yang kuat untuk ilmu data dan pembelajaran mesin. Perangkat lunak ini menyediakan beragam alat untuk persiapan data, pemodelan, evaluasi, dan implementasi. RapidMiner dirancang untuk mudah digunakan dan memungkinkan pengguna untuk dengan mudah membangun dan menguji berbagai model, bahkan tanpa pengalaman pemrograman. (Rafi Nahjan et al., 2023) Algoritma Naïve Bayes adalah metode pembelajaran berbasis formula yang digunakan untuk menyelesaikan masalah klasifikasi, terutama dalam pengelompokan teks yang terdiri dari kumpulan entitas dengan dimensi tinggi. (Rokhanah et al., 2023) Naïve Bayes merupakan salah satu algoritma pembelajaran induktif yang paling efektif dan efisien untuk pembelajaran mesin dan penambangan data. Efisiensi kompetitif Naïve Bayes dalam proses klasifikasi, meskipun menggunakan asumsi indenpedensi atribut (atau tidak ada korelasi antara atribut). Asumsi indenpedensi atribut ini jarang terjadi pada data nyata, tetap meskipun asumsi indenpedensi atribut ini dilanggar, efisiensi pengklasifikasi Naïve Bayes cukup tinggi sebagaimana dibuktikan oleh beberapa studi empiris. (Shinta et al., 2023)

Berdasarkan hasil penelitian sebelumnya Algoritma Naïve Bayes punya beberapa keunggulan dibandingkan algoritma lain, terutama karena mudah dan cepat digunakan dalam pengolahannya. Berbeda dari algoritma

lain yang lebih rumit seperti SVM atau neural networks, Naïve Bayes bisa memberikan hasil yang baik tanpa memerlukan banyak waktu atau dalam pengolahannya. Hal ini membuatnya cocok untuk data dalam jumlah besar. karena pengolahannya yang sederhana bahwa setiap fitur atau variabel tidak saling bergantung. Karena keunggulan-keunggulan ini, Naïve Bayes digunakan dalam mengidentifikasi produk motor yang diminati pelanggan pada PT. Ramayana mitra sejahtera. Adapun tujuan yang ingin dicapai diantaranya untuk mengidentifikasi prediksi penjualan motorterlaris menggunakan Algoritma *Naïve Bayes*, mengetahui jenis tipe motor yang paling diminati oleh masyarakat, serta meminimalisir ketersediaan stok motor agar data merek motor yang diterima lebih tepat sasaran.

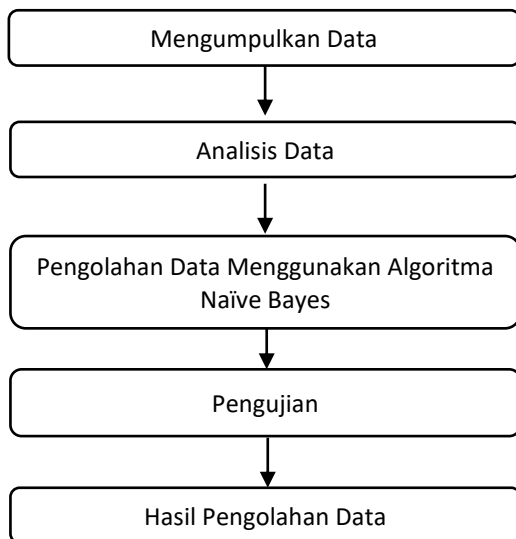
2. METODE PENELITIAN

Penelitian ini dilakukan dalam beberapa tahapan yaitu pengumpulan data, analisis data, pengelolaan menggunakan algoritma naive bayes, dan pengujian data seperti yang terlihat pada gambar 1. Pada tahapan pengumpulan data, peneliti melakukan wawancara, observasi dan studi pustaka. Wawancara dilakukan secara langsung dengan Bapak Lukmanul Hakim sebagai Branch Manager, Informasi didapat melalui tanya jawab untuk mengetahui motor yang terlaris di PT. Ramayana Mitra Sejahtera selama 3 tahun terakhir. Sehingga peneliti mendapatkan data yang signifikan dan akurat. Peneliti juga melakukan observasi secara langsung mengenai penjualan motor pada PT. Ramaya Mitra Sejahtera. Dalam kesempatan ini peneliti melakukan riset secara langsung mengenai penjualan sepeda motor terlaris di 3 tahun terakhir, yaitu data penjualan tahun 2021, tahun 2022 dan tahun 2023. Studi pustaka dilakukan peneliti untuk mengumpulkan data yang diarahkan kepada pencarian data dan informasi melalui dokumen, baik dokumen tertulis, foto, gambar, serta dokumen elektronik yang dapat mendukung dalam proses penelitian ini. Selanjutnya data yang telah dikumpulkan dianalisa lalu diolah menggunakan algoritma naive bayes.

Pada tahapan pengujian menggunakan RapidMiner, peneliti melakukan beberapa langkah penting. Pertama, data yang telah dianalisis diimpor ke dalam RapidMiner untuk proses pengolahan lebih lanjut. Peneliti membersihkan data untuk memastikan kualitas data yang akan diuji, termasuk menghapus duplikasi atau data yang tidak relevan.

Selanjutnya, peneliti menggunakan modul data preparation untuk membagi dataset menjadi dua bagian, yaitu training set dan testing set, dengan perbandingan tertentu. Algoritma Naive Bayes diterapkan pada training set untuk membangun model prediksi. Kemudian, model tersebut diuji menggunakan testing set untuk mengukur akurasi dan performanya. Hasil pengujian berupa metrik seperti akurasi, presisi, recall dievaluasi untuk menentukan seberapa efektif model tersebut dalam mengidentifikasi motor terlaris. Tahapan ini bertujuan memastikan bahwa model yang dihasilkan tidak hanya akurat tetapi juga dapat diandalkan untuk pengambilan keputusan.

Pengujian dalam penelitian ini dilakukan menggunakan RapidMiner dengan menggunakan data telah yang telah dipersiapkan sebelumnya. Proses ini bertujuan untuk menentukan klasifikasi produk penjualan motor ke dalam kategori Laris dan Tidak Laris. Pengujian pada penelitian ini dilakukan dengan menggunakan metode Algoritma Naive Bayes dengan data penjualan dari 3 tahun (2021 – 2023) dengan data sebanyak 1324 record di PT. Ramayana Mitra Sejahtera. Dengan atribut yang ada yaitu Tahun, Bulan, Tanggal, Nomor surat jalan, Customer, Tipe Motor, Warna, Nomor mesin dan Metode pembayaran.



Gambar 1. Tahapan Penelitian

3. HASIL DAN PEMBAHASAN

Data penjualan sepeda motor di PT. Ramayana Mitra Sejahtera selama tiga tahun yaitu pada tahun 2021 sampai tahun 2023 yang berhasil dikumpulkan terdiri dari 1.324 record, yang mencakup 10 atribut penting terkait

penjualan yang tercantum pada tabel 1. Data ini digunakan sebagai dasar sebelum dilakukan proses penelitian untuk menentukan klasifikasi.

Tabel 1. Atribut data sebelum diolah

No.	Nama Atribut	Tipe data
1	Tahun	Numerik
2	Bulan	Karakter
3	Tanggal	Date
4	No. Surat Jalan	Numerik
5	Nama Pelanggan	Karakter
6	Tipe Motor	Karakter
7	Warna	Karakter
8	Nomor Mesin	Karakter
9.	Metode Pembayaran	Karakter
10.	Harga Jual	Numerik

Tabel 1 berisi atribut beserta tipe datanya sebelum diolah yang berisi atribut-atribut awal dalam dataset yang terdiri dari data numerik dan karakter. Atribut numerik mencakup Tahun, No Surat Jalan, harga jual. Pada atribut Date pada Tanggal dan sedangkan atribut karakter meliputi Bulan, Nama Pelanggan, Tipe Motor Warna, No Mesin, dan Metode Pembayaran. Data ini merupakan informasi mentah yang nantinya akan diproses lebih lanjut untuk keperluan analisis penelitian.

Tabel 2 menyajikan informasi awal terkait penjualan sepeda motor di PT. Ramayana Mitra Sejahtera. Data ini mencakup beberapa atribut utama, yaitu Tahun, Bulan, dan Tanggal untuk mencatat waktu transaksi; Nomor Surat Jalan sebagai identitas dokumen penjualan; Customer sebagai pembeli, Tipe Motor dan Warna untuk spesifikasi produk, Nomor Mesin sebagai identifikasi unik kendaraan, Metode Transaksi untuk cara pembayaran; serta Harga Jual yang mencatat nilai transaksi. Data ini merupakan informasi mentah yang menjadi dasar untuk proses pengolahan dan analisis lebih lanjut.

3.1. Data Reduction

Pada tahapan ini atribut yang dianggap kurang relevan untuk analisis akan dihapus atau dikurangi untuk meningkatkan efisiensi model, serta menyederhanakan struktur dataset. Langkah ini dilakukan untuk memastikan bahwa hanya atribut yang signifikan dan memiliki pengaruh langsung terhadap hasil penelitian yang digunakan. Hasil dari tahapan ini disajikan pada tabel 3. Dengan demikian, proses analisis dapat berjalan lebih optimal.

Tabel 2. Data Mentah Penjualan Sepeda Motor PT. Ramayana Mitra Sejahterah

Tahun	Bulan	Tanggal	No Surat Jalan	Customer	Tipe Motor	Warna	No. Mesin	Metode Transaksi	Harga Jual
2021	Jan	02/01/2021	SJ-2021/00000001	UDI	BEAT SPORTY	BIRU HITAM	JM91E1316991	Cash	17.647.000
2021	Jan	02/01/2021	SJ-2021/00000003	AHMAD MURSAN QO SANTI APRIYANTI	BEAT SPORTY	SILVER	JM91E1360310	Kredit	17.747.000
2021	Jan	02/01/2021	SJ-2021/00000005	RUSMIATI	ADV	RED	KF51E1057438	Kredit	34.893.000
2021	Jan	04/01/2021	SJ-2021/00000007	MATREJO SUKIMAN	Vario 125	WHITE	JM51E1703985	Cash	21.656.000
2021	Jan	04/01/2021	SJ-2021/00000009	MUHAMMAD SANDI	Scoopy	WHITE	JM02E1046354	Cash	21.317.000
2021	Jan	04/01/2021	SJ-2021/00000011	KURDIN QQ AAS NURASIAH	PCX	HITAM	KF21E1465459	Kredit	29.994.000
2021	Jan	05/01/2021	SJ-2021/00000013	SUPRAPTO	PCX	WHITE	KF21E1427287	Kredit	29.994.000
2021	Jan	05/01/2021	SJ-2021/00000015	WARNO	BEAT SPORTY	SILVER	JM91E1374737	Kredit	17.747.000
2021	Jan	05/01/2021	SJ-2021/00000017	AHMAD SURURI AFIF	Scoopy	WHITE	JM02E1038966	Kredit	21.317.000
2021	Jan	06/01/2021	SJ-2021/00000019	SUPARBEDO QO MUIFAH ROSMAYANTI	Scoopy	KREM COKLAT	JM01E1044971	Kredit	20.517.000
2023	Dec	30/12/2023	SJ-2023/00004101	ARI PRATAMA	Vario 125	WHITE	JMD1E1407109	Kredit	25.173.000
2023	Dec	30/12/2023	SJ-2023/00004104	NURUL AZMI	BEAT SPORTY	GREEN	JM91E3456051	Kredit	19.493.000

Tabel 3. Data Penjualan Dengan Atribut Yang Sudah Dikurangi

No	Type	Warna	Penjualan	Harga
1	ADV	MERAH HITAM	2	Rp 77.632.000,00
2	ADV	PUTIH HITAM	2	Rp 62.797.863,00
3	ADV	RED	2	Rp 69.786.000,00
4	BEAT SPORTY	BIRU HITAM	1	Rp 17.647.000,00
5	BEAT SPORTY	BIRU PUTIH	6	Rp 99.612.000,00
6	BEAT SPORTY	HITAM	19	Rp 332.613.000,00
7	BEAT SPORTY	MAGENTA HITAM	1	Rp 17.647.000,00
...
...
...
1342	Vario 160	WHITE	1	Rp 30.543.000,00

3.2. Data Integration

Pada tahapan ini, jumlah atribut pada dataset berjumlah sepuluh atribut dengan 1342 record data. Setelah melalui proses reduksi, jumlah atribut berkurang menjadi lima atribut yaitu bulan, tipe, harga, warna, dan jumlah penjualan. Kemudian, dua atribut ditambahkan, yaitu varian dan kelas/label status penjualan. Status penjualan berisi data laris atau tidak laris.

3.3. Data Transformation

Range harga jual dikategorikan dari data mentah menjadi tiga kategori yaitu murah, sedang dan mahal dengan range yang dapat dilihat pada tabel 4. Tujuannya adalah untuk

mempermudah pengolahan data dan meningkatkan akurasi dalam pengklasifikasian. Proses transformasi ini dilakukan pada atribut harga, yang dikelompokkan menjadi kategori murah, sedang, dan mahal. Penentuan kategori harga ini berdasarkan hasil wawancara dengan Branch Manager PT. Ramayana Mitra Sejahtera.

Tabel 4. Range Harga Jual

Range Harga	Kategori
<Rp.20.000.000	Murah
>Rp.20.000.000 – Rp.25.000.000	Sedang
>Rp.25.000.000	Mahal

3.4. Pengolahan Data Algoritma Naïve Bayes

Proses ini mencakup perhitungan probabilitas untuk setiap kelas berdasarkan atribut dalam dataset. Implementasi dilakukan menggunakan Microsoft Excel untuk mempermudah perhitungan dan analisis data secara efektif. Rumus yang digunakan untuk mendapatkan hasil probabilitas kelas pada perhitungan algoritma naïve bayes di Excel menggunakan beberapa formula dasar, yang menghasilkan probabilitas untuk label "Laris" sebesar 0,083457526 dan untuk label "Tidak Laris" sebesar 0,916542474 seperti yang terlihat pada tabel 5. Probabilitas ini menunjukkan kecenderungan data penjualan untuk diklasifikasikan ke dalam kategori "Tidak Laris".

Tabel 5. Probabilitas Kelas Klasifikasi

Klasifikasi	Jumlah
Laris	0,083457526
Tidak Laris	0,916542474

3.5. Pengujian Aplikasi RapidMiner

Pengujian selanjutnya akan menggunakan aplikasi Rapidminer untuk mempermudah implementasi algoritma naïve bayes. Pada pengujian ini, akan digunakan data penjualan sebagai dataset uji. Dalam penelitian ini, percobaan dilakukan menggunakan metode Algoritma naïve bayes dengan perangkat lunak Rapidminer. Pengujian ini dilakukan dengan menggunakan 1342 record data. Gambar 2 menyajikan hasil evaluasi kinerja model dengan Perbandingan 60% dan 40% yang menghasilkan 474 jumlah sampel yang benar-benar positif dan diprediksi sebagai positif, 9 jumlah sampel yang sebenarnya positif tetapi diprediksi sebagai negative. 18 jumlah sampel yang sebenarnya negatif tetapi diprediksi sebagai positif dan 36 jumlah sampel yang benar-benar negatif dan diprediksi sebagai negative. Confusion Matrix ini menunjukkan bahwa model memiliki kinerja yang baik, dengan accuracy, precision, recall, dan F1-score yang tinggi. Angka 474 dan 36 menunjukkan sebagian besar prediksi model sudah benar, sementara jumlah kesalahan prediksi (9 FN dan 18 FP) cukup kecil.

$$Accuracy: \frac{(TP+TN)}{(TP+TN+FP+FN)} \times 100\% = \frac{(474 + 36)}{(474 + 36 + 9 + 18)} \times 100\% = 94,97\%$$

Recall:

$$Pred. Tidak Laris : \frac{(474)}{(474+18)} \times 100\% = 96,34\%$$

$$Pred. Laris : \frac{(36)}{(36+9)} \times 100\% = 80,00\%$$

Precision:

$$Pred. Tidak Laris = \frac{(474)}{(474+9)} \times 100\% = 98,14\%$$

$$Pred. Laris = \frac{(36)}{(36+18)} \times 100\% = 66,67\%$$

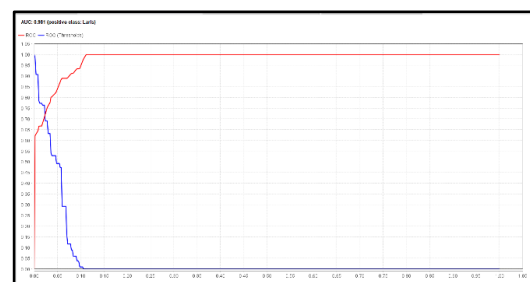
Prediction Sample			
Actual sample		True	False
	True	474	9
	False	18	36

Gambar 2. Confusion Matrix

Accuracy 94.97%			
	Test Tidak Laris	Test Laris	Class Precision
pred. Tidak Laris	474	9	98.14%
pred. Laris	18	36	66.67%
class recall	96.34%	80.00%	

Gambar 3. Nilai Accuracy

Gambar 3 menunjukkan hasil perhitungan algoritma naïve bayes menggunakan aplikasi RapidMiner dengan konfigurasi 60% data training dan 40% data testing yang menunjukkan akurasi tinggi sebesar 94,97%. Model ini efektif dalam memprediksi kategori "Laris" dengan Recall 80,00% dan Precision 66,67%, serta memiliki performa lebih baik pada kategori "Tidak Laris" dengan Recall 96,34% dan Precision 98,14%.



Gambar 4. Kurva ROC

Gambar 4 menunjukkan hasil nilai AUC sebesar 0,981, yang memperkuat bahwa model ini memiliki performa yang sangat baik dalam melakukan prediksi. Nilai AUC yang mendekati 1 menunjukkan bahwa model ini mampu

membedakan dengan jelas antara kelas *Laris* dan *Tidak Laris* dalam penjualan sepeda motor di PT. Ramayana Mitra Sejahtera. Hal ini menunjukkan bahwa model yang digunakan dalam penelitian ini sangat optimal dan dapat diandalkan untuk memberikan prediksi yang akurat terkait penjualan motor di perusahaan tersebut.

Berdasarkan evaluasi model, Naïve Bayes dengan konfigurasi 60% data training dan 40% data testing dalam prediksi penjualan sepeda motor di PT. Ramayana Mitra Sejahtera sangat baik. Namun, hasil evaluasi ini perlu dibahas lebih lanjut dalam konteks penelitian-penelitian serupa. Misalnya, dalam penelitian sebelumnya yang menggunakan metode K-Nearest Neighbor (K-NN) untuk prediksi penjualan, akurasi yang dicapai adalah sekitar 96,15%. Perbandingan ini menunjukkan bahwa metode K-NN juga memberikan hasil yang baik.

4. KESIMPULAN

Berdasarkan hasil implementasi algoritma Naïve Bayes menggunakan aplikasi RapidMiner, dapat disimpulkan bahwa model dengan perbandingan 60% data training (805 data) dan 40% data testing (537 data) menunjukkan kinerja yang sangat baik, dengan akurasi mencapai 94,97%. Hal ini menunjukkan bahwa pengolah ini merupakan yang paling efektif untuk mengukur akurasi pada dataset penjualan sepeda motor di PT. Ramayana Mitra Sejahtera. Model ini efektif dalam mendeteksi kategori "Laris" dengan Recall sebesar 80,00% dan Precision 66,67%, yang berarti model cukup baik dalam mengenali data "Laris". Untuk kategori "Tidak Laris", model menunjukkan performa lebih tinggi dengan Recall 96,34% dan Precision 98,14%, yang mengindikasikan kemampuan model yang sangat baik. Selain itu, analisis terhadap dataset penjualan sepeda motor juga mengungkapkan bahwa produk terlaris di PT. Ramayana Mitra Sejahtera adalah tipe motor Beat Sporty dengan varian warna hitam.

5. REFERENSI

Apriyani, H., & Kurniati, K. (2020). Perbandingan Metode Naïve Bayes Dan Support Vector Machine Dalam Klasifikasi Penyakit Diabetes Melitus. *Journal of Information Technology Ampere*, 1(3), 133–143. <https://doi.org/10.51519/journalita.volume1.issue3.year2020.page133-143>

- Butsianto, S., & Mayangwulan, N. T. (2020). Penerapan Data Mining Untuk Prediksi Penjualan Mobil Menggunakan Metode K-Means Clustering. *Jurnal Nasional Komputasi Dan Teknologi Informasi (JNKTI)*, 3(3), 187–201. <https://doi.org/10.32672/jnkti.v3i3.2428>
- Dewi, S. P., Nurwati, N., & Rahayu, E. (2022). Penerapan Data Mining Untuk Prediksi Penjualan Produk Terlaris Menggunakan Metode K-Nearest Neighbor. *Building of Informatics, Technology and Science (BITS)*, 3(4), 639–648. <https://doi.org/10.47065/bits.v3i4.1408>
- igual L. (2020). Penerapan data science. *Вестник Росздравнадзора*, 4(1), 9–15.
- Pramana, I., Sudiarsa, I. W., & ... (2023). Penerapan Algoritma Naive Bayes Untuk Prediksi Penjualan Produk Terlaris Pada CV Akusara Jaya Abadi. *JATISI (Jurnal Teknik ...)*, 10(4), 518–534. <https://jurnal.mdp.ac.id/index.php/jatisi/article/view/6498%0Ahttps://jurnal.mdp.ac.id/index.php/jatisi/article/download/6498/1694>
- Rafi Nahjan, M., Nono Heryana, & Apriade Voutama. (2023). Implementasi Rapidminer Dengan Metode Clustering K-Means Untuk Analisa Penjualan Pada Toko Oj Cell. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 7(1), 101–104. <https://doi.org/10.36040/jati.v7i1.6094>
- Rezekika, D. (2020). Penerapan Algoritma Naïve Bayes Untuk Memprediksi Penjualan Spare Part Sepeda Motor. *Jurnal Pelita Informatika*, 8(3), 326–329.
- Rianti, W., & Harahap, E. (2021). Pengolahan Data Hasil Penjualan Online Menggunakan Aplikasi Microsoft Excel. *Matematika: Jurnal Teori Dan Terapan Matematika*, 20(2), 69–76. <https://journals.unisba.ac.id/index.php/matematika/article/view/1553>
- Ridho Handoko, M. (2021). Sistem Pakar Diagnosa Penyakit Selama Kehamilan Menggunakan Metode Naive Bayes Berbasis Web. *Jurnal Teknologi Dan Sistem Informasi (JTSI)*, 2(1), 50–58. <http://jim.teknokrat.ac.id/index.php/JTSI>
- Ridwan, A. (2020). Penerapan Algoritma Naïve Bayes Untuk Klasifikasi Penyakit Diabetes Mellitus. *Jurnal SISKOM-KB (Sistem Komputer Dan Kecerdasan Buatan)*, 4(1), 15–21. <https://doi.org/10.47970/siskom-kb.v4i1.169>

Rokhanah, S., Hermawan, A., & Avianto, D. (2023). Pengaruh Principal Component Analysis Pada Naïve Bayes dan K-Nearest Neighbor Untuk Prediksi Dini Diabetes Melitus Menggunakan Rapidminer. *EVOLUSI : Jurnal Sains Dan Manajemen*, 11(1).
<https://doi.org/10.31294/evolusi.v11i1.14728>

Shinta, R., Fairul, P., & Saputri, G. (2023). Penerapan Data Mining Untuk Prediksi Penjualan Cat Menggunakan Metode Naïve (Studi Kasus : Mitra 10 Gading Serpong). *LOGIC: Jurnal Ilmu Komputer Dan ...*, 1(3), 571–578.