

Implementasi Algoritma *Naïve Bayes* Terhadap Data Penjualan untuk Mengetahui Pola Pembelian Konsumen pada Kantin

Raihan Putra Mohammad Rosidi ^{1*}, Kiki Setiawan ²

^{1*,2} Program Studi Teknik Informatika, Sekolah Tinggi Ilmu Komputer Cipta Karya Informatika, Kota Jakarta Timur, Daerah Khusus Ibukota Jakarta, Indonesia.

Email: raihanhugo71@gmail.com ^{1*}, ki2djoaz@gmail.com ²

Histori Artikel:

Dikirim 25 Juli 2023; *Diterima dalam bentuk revisi* 23 Agustus 2023; *Diterima* 1 November 2023; *Diterbitkan* 10 Januari 2024. Semua hak dilindungi oleh Lembaga Penelitian dan Pengabdian Masyarakat (LPPM) STMIK Indonesia Banda Aceh.

Abstrak

Implementasi algoritma *Naïve Bayes* pada data penjualan untuk menentukan pola pembelian konsumen di kantin merupakan pendekatan yang berupaya mengatasi keterbatasan pemahaman tentang perilaku konsumen di kantin. Tujuan dari pendekatan ini adalah untuk menganalisis data penjualan dan mengidentifikasi pola pembelian untuk mendapatkan wawasan tentang perilaku konsumen, mengoptimalkan strategi penjualan, dan meningkatkan kepuasan pelanggan. Dengan menerapkan algoritme *Naïve Bayes* pada data penjualan, penelitian ini berupaya mengidentifikasi faktor-faktor yang memengaruhi perilaku konsumen, seperti harga, kenyamanan, dan penawaran menu, serta memberikan wawasan yang dapat digunakan untuk mengoptimalkan penawaran menu, strategi penetapan harga, dan alokasi sumber daya. Selain itu, pendekatan tersebut berusaha untuk meningkatkan kepuasan pelanggan dan membedakan kantin dari pesaing dengan mempersonalisasi penawaran menu dan meningkatkan pengalaman pelanggan secara keseluruhan. Penelitian ini juga bertujuan untuk berkontribusi pada bidang penelitian perilaku konsumen dengan menerapkan algoritma *Naïve Bayes* pada data penjualan di kantin, berpotensi memberikan wawasan yang dapat diterapkan dalam konteks lain. Secara keseluruhan, penerapan algoritma *Naïve Bayes* pada data penjualan di kantin dapat memberikan pendekatan berbasis data untuk memahami perilaku konsumen dan meningkatkan strategi penjualan dan kepuasan pelanggan.

Kata Kunci: Algoritma *Naïve Bayes*; Strategi Penjualan; Pola Pembelian; Data Mining.

Abstract

Implementation of the *Naïve Bayes* algorithm on sales data to determine consumer buying patterns in canteens is an approach that seeks to overcome the limited understanding of consumer behavior in canteens. The goal of this approach is to analyze sales data and identify purchasing patterns to gain insight into consumer behavior, optimize sales strategies, and improve customer satisfaction. By applying the *Naïve Bayes* algorithm to sales data, this research seeks to identify factors that influence consumer behavior, such as price, convenience, and menu offerings, and provide insights that can be used to optimize menu offerings, pricing strategies and resource allocation. Additionally, the approach seeks to increase customer satisfaction and differentiate canteens from competitors by personalizing menu offerings and enhancing the overall customer experience. This research also aims to contribute to the field of consumer behavior research by applying the *Naïve Bayes* algorithm to canteen sales data, potentially providing insights that can be applied in other contexts. Overall, the application of the *Naïve Bayes* algorithm to canteen sales data can provide a data-driven approach to understanding consumer behavior and improving sales strategy and customer satisfaction.

Keyword: *Naïve Bayes* Algorithm; Sales Strategies; Purchasing Patterns; Data Mining.

1. Pendahuluan

Memahami perilaku konsumen sangat penting untuk bisnis yang ingin mengoptimalkan strategi penjualan mereka dan meningkatkan kepuasan pelanggan. Dalam kantin, mengidentifikasi pola pembelian konsumen dapat memberikan wawasan mengenai faktor-faktor yang mempengaruhi keputusan pembelian, seperti harga, kenyamanan, dan penawaran menu. Namun, metode analisis data tradisional mungkin tidak cukup untuk sepenuhnya menangkap sifat kompleks dari perilaku konsumen. Pada kesempatan ini, peneliti akan menggunakan algoritme *machine learning*, seperti algoritme Naïve Bayes, untuk menganalisis data penjualan dan mengidentifikasi pola perilaku konsumen. Algoritma Naïve Bayes adalah teknik yang umum digunakan dalam data *mining* dan *machine learning* yang dapat mengklasifikasikan data berdasarkan probabilitas dan pengalaman sebelumnya. Dalam kasus ini, algoritma dapat dilatih pada data penjualan untuk memprediksi pola pembelian di masa mendatang dan memberikan rekomendasi untuk penawaran menu dan strategi penetapan harga. Secara keseluruhan, penerapan algoritma Naïve Bayes pada data penjualan untuk menentukan pola pembelian konsumen di Kantin berpotensi memberikan wawasan berharga bagi bisnis yang ingin mengoptimalkan strategi penjualan mereka dan meningkatkan kepuasan pelanggan.

Beberapa penelitian telah memanfaatkan algoritma *machine learning*, khususnya algoritma Naïve Bayes, untuk mengungkap pola dan perilaku konsumen. Sejumlah studi telah dilakukan dalam berbagai bidang industri dan toko untuk memahami perilaku pembelian konsumen. Aksoy, Atas, dan Karabatak (2019) mengeksplorasi kebiasaan berbelanja menggunakan algoritma klasifikasi data mining [1], sedangkan Nurdianawati dan Salim (2018) menerapkan metode Naïve Bayes Classifier pada penjualan barang guna optimalisasi strategi pemasaran [2]. Di sisi lain, Jiaxian dan Gengming (2019) mengusulkan algoritma Naïve Bayesian yang ditingkatkan untuk memprediksi perilaku mikroblog [3]. Selain itu, penelitian juga dilakukan pada tingkat toko atau outlet, seperti studi Rozaq (2021) tentang analisis transaksi penjualan pada 212 Mart di Palembang [4], dan penelitian Syamsudin, Halundaka, dan Nugroho (2020) yang memprediksi status konsumen produk celana menggunakan Naïve Bayes [5].

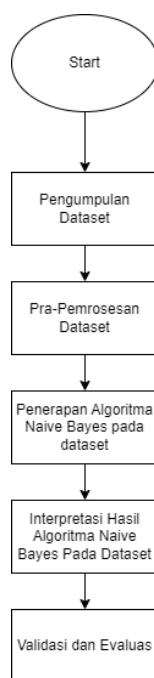
Penelitian lain juga berfokus pada aplikasi algoritma Naïve Bayes dalam memprediksi penjualan produk di toko tertentu. Yulianto (2019) melakukan studi kasus pada Couple Store di Yogyakarta [6], sementara Sari (2022) mengimplementasikan data mining untuk memprediksi penjualan menggunakan metode Naïve Bayes [7]. Selain itu, Nawangsih dan Setyaningsih (2020) menerapkan algoritma Naïve Bayes untuk menentukan klasifikasi produk terlaris dalam penjualan kacang-kacangan [8], sedangkan Ariska, Hasibuan, dan Purba (2020) fokus pada perhitungan nilai *Point Of Sales* (POS) dari penjualan produk fashion di CV. Sumber Makmur [9]. Penelitian Juwita dan Damanik (2022) yang memprediksi penjualan di toko VJCakes di Pematang Siantar menggunakan algoritma Naïve Bayes [10]. Begitu pula dengan penelitian Pransiska, Mirza, dan Andri (2019) menunjukkan bahwa algoritma Naïve Bayes tidak hanya berperan dalam pemahaman perilaku konsumen secara umum, tetapi juga dapat disesuaikan dengan kebutuhan dan karakteristik unik dari setiap bisnis [11]. Semua penelitian ini menunjukkan signifikansi algoritma Naïve Bayes dalam menganalisis perilaku konsumen dan memprediksi tren penjualan. Penelitian ini menciptakan kerangka konseptual yang kokoh dalam penerapan algoritma Naïve Bayes untuk meningkatkan pemahaman terhadap perilaku konsumen dan mengoptimalkan strategi penjualan.

Penelitian yang memanfaatkan algoritma Naïve Bayes untuk menganalisis perilaku konsumen dan meramalkan tren penjualan memiliki relevansi strategis yang signifikan bagi perusahaan dalam mencapai optimalisasi kinerja. Dengan mendalaminya pemahaman terhadap pola perilaku konsumen, perusahaan dapat menyelaraskan strategi penjualan, penawaran produk, dan penetapan harga secara lebih presisi. Keterkaitan penelitian ini terfokus pada penerapan teknik analisis data, khususnya algoritma Naïve Bayes, untuk mengungkap pola yang mungkin sulit terdeteksi dengan metode konvensional. Hasil dari penelitian ini menghasilkan informasi bagi pemangku kepentingan bisnis, mendukung mereka dalam pengambilan keputusan yang lebih terinformasi terkait strategi pemasaran, manajemen persediaan, dan pengembangan produk. Dengan memanfaatkan algoritma Naïve Bayes, perusahaan dapat mengoptimalkan strategi penjualan, meningkatkan efisiensi operasional, dan

mengantisipasi kebutuhan konsumen dengan tingkat akurasi yang lebih tinggi. Hal ini tidak hanya berdampak positif pada peningkatan penjualan, tetapi juga memperkuat kepuasan pelanggan dengan menyajikan pengalaman yang lebih personal dan sesuai dengan preferensi mereka.

2. Metode Penelitian

Metode yang digunakan dalam penelitian ini adalah Algoritma Naïve Bayes merupakan metode pengklasifikasian paling populer digunakan dengan tingkat keakuratan yang baik. Metode ini menggunakan data penjualan dan data mining untuk digunakan pada Algoritma Naïve Bayes. Berikut adalah metodologi penelitian yang dilakukan penulis.



Gambar 1. Tahapan Penelitian

- 1) Pengumpulan Dataset
Mengumpulkan dataset dari kantin selama periode waktu tertentu, dataset mencakup informasi tentang makanan dan minuman yang dibeli, jumlah penjualan, waktu pembelian, dan harga produk
- 2) Pra-pemrosesan Dataset
Mengubah format data ke dalam struktur yang sesuai, seperti file CSV atau spreadsheet kemudian menyusun atribut-atribut yang akan digunakan dalam analisis, seperti jenis produk, tanggal pembelian, jumlah pembelian, dan harga.
- 3) Penerapan Algoritma Naïve Bayes Pada Dataset
Menganalisis output dari Algoritma Naïve Bayes untuk mengidentifikasi pola pembelian konsumen dan juga menganalisis factor-faktor yang mempengaruhi keputusan pembelian, seperti harga atau pilihan menu khusus.
- 4) Interpretasi Hasil Algoritma Naïve Bayes Pada Dataset
Mengidentifikasi area potensial untuk peningkatan strategi penjualan dan penawaran makanan dan minuman setelah itu membuat rekomendasi berdasarkan wawasan yang diperoleh untuk meningkatkan kepuasan konsumen dan efektivitas penjualan
- 5) Validasi dan Evaluasi
Mengevaluasi efektivitas implementasi Algoritma Naïve Bayes dalam mengungkap pola pembelian konsumen pada kantin.

3. Hasil dan Pembahasan

3.1 Perhitungan Algoritma Naïve Bayes

Data penjualan selama ini tidak tersusun rapi, sehingga data penjualan selama ini tidak dimanfaatkan oleh baik dengan koperasi yang akhirnya hanya menjadi arsip data dan pembukuan koperasi dan tidak diketahui apa manfaat dari data data yang ada. Melalui perhitungan algoritma apriori kita bisa membuat suatu analisis asosiasi atau association rule mining, analisis asosiasi adalah teknik data mining untuk menemukan aturan suatu kombinasi item. Pentingnya suatu asosiasi dapat dinilai menggunakan dua metrik, yaitu support dan confidence. Support (nilai penunjang) mengacu pada persentase kemunculan kombinasi item tersebut dalam database, sedangkan confidence (nilai kepastian) mengukur kekuatan hubungan antara item dalam aturan asosiasi.

1) Data Transaksi Penjualan Barang

Tabel 1. Transaksi Penjualan

No	Nama Produk	Jenis Jajanan	Stok Produk	Output
1	A1	Makanan	Normal	Terjual
2	A2	Makanan	Normal	Terjual
3	A3	Makanan	Normal	Terjual
4	A4	Makanan	Normal	Terjual
5	A5	Makanan	Normal	Terjual
6	A6	Minuman	Normal	Terjual
7	A7	Minuman	Normal	Terjual
8	A8	Minuman	Normal	Terjual
9	A9	Minuman	Normal	Terjual
10	A10	Minuman	Normal	Terjual
11	A11	Minuman	Tinggi	Tidak Terjual
12	A12	Minuman	Rendah	Tidak Terjual
13	A13	Minuman	Normal	Terjual
14	A14	Minuman	Tinggi	Tidak Terjual
15	A15	Makanan	Rendah	Tidak Terjual
16	A16	Makanan	Rendah	Tidak Terjual
17	A17	Minuman	Rendah	Terjual
18	A18	Minuman	Rendah	Terjual
19	A19	Minuman	Rendah	Terjual
20	A20	Makanan	Rendah	Terjual

2) Pengolahan data

Untuk menghitung probabilitas yang Terjual dan Tidak Terjual, rumus yang digunakan adalah sebagai berikut:

$$P(A) = \frac{n(A)}{n(S)}$$

Keterangan:

P(A) = Probabilitas terjadinya kejadian A

n(A) = Banyaknya kejadian yang dimaksud

n(S) = Jumlah Keseluruhan Data

$$P(\text{Terjual}) = 15/20 = 0,75$$

$$P(\text{Tidak Terjual}) = 5/20 = 0,25$$

Untuk hasil probabilitas dapat dilihat pada data berikut

Tabel 2. Hasil probabilitas Terjual/Tidak Terjual

	Terjual	Tidak Terjual	Terjual	Tidak Terjual
P(Terjual/Tidak Terjual)	15	5	0,75	0,25

Setelah probabilitas dari data telah diketahui, selanjutnya penulis menghitung masing-masing probabilitas dari setiap variabel yang digunakan dalam penelitian ini. Variabel yang digunakan yakni ukuran, rasa, jenis produk, dan stok. Penentuan probabilitas setiap variabel, dengan melakukan perhitungan untuk setiap nilai pada variabel.

3) Probabilitas variabel Jenis Jajanan

Berikut adalah penyelesaian variabel Jenis Jajanan berdasarkan data yang sudah disediakan pada tabel 1.

Tabel 3. Hasil probabilitas variabel Jenis Jajanan

P(J=↓ ...	Terjual	Tidak Terjual	Terjual	Tidak Terjual
Makanan	6	2	0,3	0,1
Minuman	9	3	0,45	0,15

Dari tabel diatas dapat diketahui bahwa untuk variabel Jenis Jajanan, nilai (Makanan | Terjual) adalah 0,3, nilai (Minuman | Terjual) adalah 0,45, nilai(Makanan | Tidak Terjual) adalah 0,1, nilai(Minuman | Tidak Terjual) adalah 0,15, dan total dari keseluruhan probabilitas adalah 1 untuk setiap kategori terjual dan tidak terjual.

4) Probabilitas variabel Stok Produk

Berikut adalah penyelesaian variabel Stok Produk berdasarkan data yang sudah disediakan pada tabel 1.

Tabel 4. Hasil probabilitas variabel Stok Produk

P(S=↓ ...	Terjual	Tidak Terjual	Terjual	Tidak Terjual
Rendah	4	3	0,2	0,15
Normal	11	0	0,55	0
Tinggi	0	2	0	0,1

Dari tabel diatas diketahui bahwa untuk variabel Stok Produk, nilai(Rendah | Terjual) adalah 0,2, nilai (Normal | Terjual) adalah 0,55, nilai (Tinggi | Terjual) adalah 0, nilai(Rendah | Tidak Terjual) adalah 0,15, nilai(Normal | Tidak Terjual) adalah 0, nilai(Tinggi | Tidak Terjual) adalah 0,1, dan total dari keseluruhan probabilitas adalah 1 untuk setiap kategori terjual dan tidak terjual.. Dari pengujian per masing-masing variabel yang menggunakan algoritma Naïve Bayes menghasilkan nilai yang hampir sama dengan nilai angka-angka yang telah ditentukan. Berdasarkan hasil tersebut maka system ini dapat membantu pola pembelian konsumen berdasarkan probabilitas variabel yang dibeli ataupun dihindari pembeli, kemudian dari hasil pengujian ini juga dapat membantu pihak kantin dalam mengelola produk-produk yang tersedia.

4. Kesimpulan

Peneliti menyimpulkan bahwa penerapan algoritma Naïve Bayes pada data penjualan kantin merupakan pendekatan yang bermanfaat dalam meningkatkan strategi penjualan dan efisiensi pengelolaan kantin. Hasil penelitian ini memberikan wawasan yang berharga bagi pengelola kantin dalam mengambil keputusan bisnis yang lebih cerdas dan berbasis data. Namun, penelitian ini juga memiliki beberapa batasan. Salah satunya adalah keterbatasan data yang mungkin mempengaruhi akurasi analisis. Selain itu, faktor-faktor lain di luar data penjualan, seperti tren pasar atau perubahan kebijakan, juga dapat mempengaruhi perilaku konsumen di kantin. Rekomendasi untuk penelitian mendatang adalah untuk memperluas cakupan data dan mempertimbangkan faktor-faktor eksternal yang dapat memengaruhi pola pembelian konsumen. Selain itu, dapat dipertimbangkan pula penggunaan algoritma.

5. Ucapan Terima Kasih

Terima kasih kepada pihak kantin yang telah memberikan layanan dan dukungan selama saya berada di lingkungan kantor. Dukungan penuh dari kantin telah membuat segala kebutuhan studi menjadi lebih mudah terpenuhi. Terima kasih atas ketersediaan produk dan layanan yang diberikan dengan profesionalisme tinggi. Kantin telah menjadi bagian yang tak terpisahkan dari perjalanan studi dan hidup saya di lingkungan sekolah. Selain itu, dukungan dan kerjasama yang diberikan oleh tim Kantin telah membantu dalam menavigasi proses penelitian ini dengan lebih lancar. Respon cepat terhadap pertanyaan dan permintaan yang diajukan sangatlah berarti dalam mempercepat proses penelitian dan analisis data. Semoga hubungan kerjasama ini dapat terus berlanjut dan memberikan manfaat bagi kedua belah pihak.

6. Daftar Pustaka

- [1] Aksoy, G., Ataş, P., & Karabatak, M. (2019, November). Investigation of shopping habits using data mining classification algorithms. In *2019 1st International Informatics and Software Engineering Conference (UBMYK)* (pp. 1-5). IEEE. DOI: <https://doi.org/10.1109/UBMYK48245.2019.8965647>.
- [2] Nurdiawan, O., & Salim, N. (2018). Penerapan Data Mining Pada Penjualan Barang Menggunakan Metode Metode Naive Bayes Classifier Untuk Optimasi Strategi Pemasaran. *Jurnal Teknologi Informasi dan Komunikasi*, 13(1), 84-95.
- [3] Jiaxian, Y., & Gengming, Z. (2019, March). Algorithm based on improved naive Bayesian for predicting microblog behavior. In *Proceedings of the 2019 3rd International Conference on Innovation in Artificial Intelligence* (pp. 182-187). DOI: <https://doi.org/10.1145/3319921.3319956>.
- [4] ROZAQ, A. J. (2021). IMPLEMENTASI METODE KLASIFIKASI ALGORITMA NAIVE BAYES DALAM MENGANALISIS TRANSAKSI PENJUALAN PADA 212 MART KUTO PALEMBANG.
- [5] Syamsudin, D., Halundaka, Y. C. D., & Nugroho, A. (2020). Prediksi Status Konsumen Produk Celana Menggunakan Naïve Bayes. *JOINTECS (Journal of Information Technology and Computer Science)*, 5(3), 177-184. DOI: <https://doi.org/10.31328/jointecs.v5i3.1435>.

- [6] Yulianto T. (2019). Prediksi Penjualan Produk Menggunakan Algoritma Naïve Bayes (Studi kasus Couple Store Yogyakarta). Skripsi. Fakultas Teknologi Informasi dan Elektro Universitas Teknologi Yogyakarta.
- [7] Sari, R M. (2022). Implementasi Data Mining Untuk Memprediksi Penjualan menggunakan Metode Naïve Bayes. Skripsi. Fakultas Sains dan Teknologi Universitas Pembangunan Panca Budi.
- [8] Nawangsih, I., & Setyaningsih, A. (2020). THE APPLICATION OF THE NAÏVE BAYES ALGORITHM TO DETERMINE THE CLASSICATION OF BEST SELLING PRODUCTS ON PULSES SALES. *Incomtech*, 9(1), 39-45.
- [9] Ariska, P., Hasibuan, N. A., & Purba, B. (2020). Penerapan Algoritma Naïve Bayes Untuk Perhitungan Nilai Point Of Sales (Pos) Dari Penjualan Produk Fashion (Studi Kasus: CV. Sumber Makmur). *KOMIK (Konferensi Nasional Teknologi Informasi dan Komputer)*, 4(1). DOI: <http://dx.doi.org/10.30865/komik.v4i1.2717>
- [10] Juwita, J., Safii, M., & Damanik, B. E. (2022). Naïve Bayes Algorithm For Predicting Sales at the Pematang Siantar VJCakes Store. *JOMLAI: Journal of Machine Learning and Artificial Intelligence*, 1(4), 337-346.
- [11] Pransiska, N., Mirza, A. H., & Andri, A. (2019). PENERAPAN DATA MINING PREDIKSI PENJUALAN BARANG ELEKTRONIK TERLARIS MENGGUNAKAN ALGORITMA NAÏVE BAYES (Study Kasus: Planet Cash And Credit Cabang Muara Enim). In *Bina Darma Conference on Computer Science (BDCCS)*. 1(6), 2157-2169.