



Penerapan Data Mining untuk Memprediksi Penjualan Bahan Bangunan (Ud Kiki Fatmala) Dengan Menggunakan Metode *Naive Bayes*

<u>INFO PENULIS</u>	<u>INFO ARTIKEL</u>
Muhyiddin A M Hayat Universitas Muhammadiyah Makassar muhyiddin@unismuh.ac.id	ISSN: 3026-3603 Vol.3, No. 1 April 2025 http://jurnal.ardenjaya.com/index.php/ajst
Titin Wahyuni Universitas Muhammadiyah Makassar titinwahyuni@unismuh.ac.id	
Haidul Universitas Muhammadiyah Makassar Haidul2001@gmail.com	

© 2025 Arden Jaya Publisher All rights reserved

Saran Penulisan Referensi:

Hayat, M. A M., Wahyuni, T., & Haidul. (2025). Penerapan Data Mining untuk Memprediksi Penjualan Bahan Bangunan (UD Kiki Fatmala) dengan Menggunakan Metode *Naive Bayes*. *Arus Jurnal Sains dan Teknologi*, 3(1),134 -140.

Abstrak

Penjualan bahan bangunan merupakan bidang usaha yang memiliki prospek menjanjikan dengan permintaan yang terus meningkat seiring pertumbuhan populasi dan pembangunan. Namun, permasalahan yang dihadapi oleh toko bahan bangunan, khususnya UD KIKI FATMALA, adalah kurangnya pemahaman terhadap minat pembeli, yang dapat menyebabkan penumpukan stok barang dan memperlambat perputaran modal. Untuk mengatasi permasalahan ini, penelitian ini menerapkan metode *Naive Bayes* untuk memprediksi penjualan bahan bangunan berdasarkan data historis. Penelitian ini bertujuan untuk menggunakan teknik data mining sebagai pendukung keputusan dalam menentukan barang yang memiliki potensi penjualan tinggi guna mengoptimalkan stok dan meningkatkan efisiensi penjualan. Metode penelitian mencakup pengumpulan data penjualan, preprocessing data, pembagian data menjadi training dan testing, serta penerapan algoritma *Naive Bayes*. Model yang dibangun dievaluasi menggunakan *confusion matrix* untuk menilai tingkat akurasi, *presisi*, *recall*, dan *F1-score*. Hasil penelitian menunjukkan bahwa metode *Naive Bayes* mampu memprediksi penjualan bahan bangunan dengan akurasi sebesar 84%. Implementasi model ini dilakukan menggunakan *Microsoft Excel* untuk pengolahan data awal dan *Jupyter Notebook* untuk analisis lebih lanjut. Kesimpulan dari penelitian ini adalah bahwa metode *Naive Bayes* dapat digunakan secara efektif untuk memprediksi produk yang memiliki peluang besar untuk terjual, sehingga membantu pengelolaan stok yang lebih optimal di UD KIKI FATMALA

Kata Kunci: *Naive Bayes*, prediksi penjualan, data mining, bahan bangunan, pengelolaan stok.

Abstract

Sales of building materials is a business sector that has promising prospects with increasing demand along with population growth and development. However, the problem faced by building material stores, especially UD KIKI FATMALA, is the lack of understanding of buyer interest, which can lead to stockpiling of goods and slowing down capital turnover. To overcome this problem, this study applies the Naïve Bayes method to predict sales of building materials based on historical data. This study aims to use data mining techniques as a decision support in determining goods that have high sales potential in order to optimize stock and increase sales efficiency. The research method includes collecting sales data, data preprocessing, dividing data into training and testing, and implementing the Naïve Bayes algorithm. The model built is evaluated using a confusion matrix to assess the level of accuracy, precision, recall, and F1-score. The results of the study showed that the Naïve Bayes method was able to predict sales of building materials with an accuracy of 84%. The implementation of this model was carried out using Microsoft Excel for initial data processing and Jupyter Notebook for further analysis. The conclusion of this study is that the Naïve Bayes method can be used effectively to predict products that have a high chance of being sold, thus helping to optimize stock management at UD KIKI FATMALA.

Keywords: Naïve Bayes, prediksi penjualan, data mining, bahan bangunan, pengelolaan stok.

A. Pendahuluan

Penjualan toko bahan bangunan merupakan usaha yang mempunyai prospek yang masih sangat baik di bidang perdagangan, dengan terus adanya pembangunan kebutuhan bahan dasar bangunan akan tetap dibutuhkan. Dengan pertumbuhan populasi dan pemukiman di seluruh dunia, industri bahan bangunan sedang berkembang, memungkinkan perusahaan bahan bangunan untuk masuk ke pasar yang ada dengan menjual barang seperti paku, semen, kayu, dan bahan bangunan lainnya. Karena menjanjikan keuntungan yang besar, bisnis bahan bangunan menjadi salah satu bisnis yang paling dicari.

Terdapat permasalahan yang dialami oleh Toko Bahan Bangunan Tersebut yaitu dalam menentukan apakah barang tersebut laku atau tidak dalam waktu ke depan, karena kurangnya pemahaman tentang minat pembeli terhadap barang tersebut dapat menyebabkan penumpukan barang di toko, yang memperlambat perputaran modal. Jika Anda memiliki pemahaman yang baik tentang barang yang sering terjual, Anda dapat meningkatkan jumlah barang yang paling diminati atau dicari.

Algoritma Naïve Bayes adalah salah satu algoritma yang digunakan untuk pengklasifikasian data. Terdapat beberapa penelitian terdahulu yang melakukan klasifikasi menggunakan metode Naive Bayes. Naive Bayes adalah teknik pembelajaran mesin yang menggunakan perhitungan probabilitas dan statistik yang diciptakan oleh ilmuwan Inggris Thomas Bayes; ini bertujuan untuk memprediksi kemungkinan di masa depan berdasarkan pengalaman di masa lalu.

1. Tinjauan Pustaka

A. Bahan bangunan

Bahan bangunan adalah semua bahan yang diperlukan untuk membangun bangunan tertentu, baik sebagai bahan pokok maupun penolong. Bahan bangunan juga didefinisikan sebagai bahan yang digunakan untuk membuat barang bangunan atau bahan yang memiliki sifat-sifat tertentu dalam teknik bangunan. Bahan bangunan tersebut di atas termasuk berbagai jenis kayu dan bambu, serta bahan galian dan produk yang dibuat darinya, seperti batu kapur, tras, tanah liat, dan lain-lain. Bahan logam, seperti paku, besi konstruksi, lembaran seng, dan lain-lain. Barang-barang yang memiliki sifat dan suara tertentu, seperti lembaran asbes untuk lantai karet, dan sebagainya berbagai cat yang melapisi dan mewarnai. Sifat material, tampilan fisik, dan durabilitasnya adalah cara terbaik untuk mengetahui berapa banyak bahan yang digunakan dalam konstruksi. Sifat fisik material terdiri dari sifat fisik, mekanis, kimiawi, dan sifat khusus.

Bahan bangunan berisi pengetahuan tentang bahan dan teknologi material bangunan yang meliputi sifat-sifat dan jenis bahan bangunan, penggunaannya, kekuatan dan pengawetan bahan bangunan, cara pemeliharaannya, pengetahuan rekayasa bahan bangunan dan bahan bangunan dalam perdagangan.

B. Prediksi

Prediksi adalah bagian penting dari proses pengambilan keputusan. Prediksi dilakukan selama proses produksi, yang merupakan bagian dari proses perencanaan dan pengendalian

produksi. Ini juga dilakukan untuk menentukan jumlah permintaan terhadap suatu produk. Untuk memenuhi kebutuhan pelanggan, kami menghadapi masalah stok barang yang umum, seperti stok barang yang tidak terjual atau stok barang dengan merek tertentu

C. Data Mining

Penambangan data, juga dikenal sebagai data mining, adalah proses yang menggunakan teknik statistik dan matematika untuk memanfaatkan kecerdasan buatan (AI) dan penggunaan data yang tidak terduga. Secara sederhana, data mining dapat didefinisikan sebagai rangkaian proses untuk mengekstraksi pola-pola yang menarik dari kumpulan data yang sangat besar yang terdiri dari pengetahuan buatan yang tidak diketahui. Perkembangan data mining disebabkan oleh akumulasi data seperti data penjualan, pembelian, pelanggan, transaksi, dan sebagainya selama bertahun-tahun. Data mining dilakukan untuk memprediksi kondisi masa depan dengan mempelajari perilaku data yang diamati atau apa yang terjadi.

D. Naive Bayes

Naive Bayes adalah salah satu algoritma pembelajaran induktif terbaik untuk pembelajaran mesin dan penambangan data. Efisiensi pengklasifikasi *Naive Bayes* kompetitif dalam proses klasifikasi dengan menggunakan asumsi independensi atribut. Namun, meskipun asumsi independensi atribut jarang terjadi pada data nyata, efisiensi pengklasifikasi Naive Bayes masih cukup baik.

metode Naive Bayes menggabungkan frekuensi dan kombinasi nilai dari kumpulan data yang diberikan. Algoritma ini, yang didasarkan pada teorema Bayes, menganggap bahwa setiap atribut adalah independen atau tidak bergantung satu sama lain, dengan asumsi nilai-nilai variabel kelas. Persamaan Naive Bayes sangat sederhana, jadi metode ini dapat digunakan untuk mengklasifikasikan data. Naive Bayes adalah metode klasifikasi statistik yang didasarkan pada teorema Bayes.

E. Software Pendukung

a. Jupyter Notebook

Jupyter berasal singkatan dari tiga bahasa pemrograman yakni Julia(Ju), Python(Py), dan R. Jupyter notebook merupakan aplikasi web gratis yang paling banyak digunakan oleh data scientist. Aplikasi ini dipakai untuk membuat dan membagikan dokumen yang memiliki kode, hasil hitungan, visualisasi dan teks

F. Confusion matrix

Confusion matrix adalah hasil evaluasi klasifikasi data mining yang digambarkan dalam tabel dan berisi perhitungan jumlah objek pengujian data yang diharapkan akan dimasukkan ke dalam sebuah kelas yang memiliki klasifikasi yang nyata. Nilai Ketepatan dan Recall berkisar antara 0 dan 1, dengan nilai yang lebih tinggi menunjukkan kinerja yang lebih baik.

- Accuracy* adalah persentase ketepatan *record* data yang diidentifikasi benar.
- Recall* adalah proporsi kasus positif yang diidentifikasi dengan benar.
- Precision* adalah proporsi kasus yang diidentifikasi positif dengan benar.

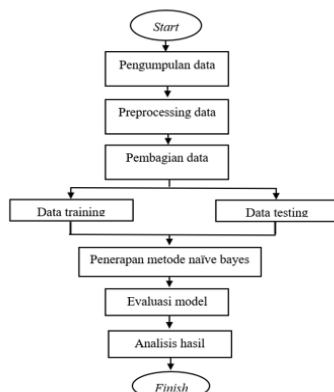
Menggunakan tabel matriks, matriks confusion menunjukkan bahwa dataset hanya terdiri dari dua kelas akan dianggap sebagai positif atau negatif.

B. Metodologi

a. Waktu dan Tempat Penelitian

Penelitian ini dilakukan di toko (UD KIKI FATMALA) Karamasa, Desa Berutallasa, Kec Biringbulu, Kab Gowa.

b. Perancangan sistem



Gambar 1 Flowchart Perancangan Sistem

1. Pengumpulan data
Langkah pertama adalah mengumpulkan data penjualan bahan bangunan dari UD KIKI FATMALA. Data yang dikumpulkan mencakup berbagai informasi yang ada seperti Tahun penjualan, Nama barang, Jenis bahan bangunan, Transaksi, dan Kategori penjualan, dan faktor-faktor lain yang mungkin mempengaruhi penjualan.
2. Processing data
Data yang telah dikumpulkan akan diproses dalam tahap preprocessing. Tujuan dari tahap praproses data adalah untuk mempersiapkan data mentah yang ada untuk dianalisis atau di proses lebih lanjut. data dibersihkan dari nilai yang hilang atau tidak valid, dan dilakukan normalisasi atau standarisasi jika diperlukan. Data kemudian diubah ke dalam format yang sesuai untuk analisis lebih lanjut, termasuk mengkodekan variabel kategori menjadi numerik atau melakukan transformasi lainnya.
3. Pembagian data
Pembagian data melibatkan pembagian dataset yang telah diproses menjadi data Training dan data Testing. Pelatihan data digunakan untuk melatih model algoritma Naive Bayes, yang memungkinkan model untuk belajar dan memahami pola data yang tersedia.
4. Penerapan Algoritma Naive Bayes
Algoritma Naive Bayes digunakan untuk membuat model prediksi. Algoritma ini menghitung kemungkinan setiap jenis penjualan, atau kelas, berdasarkan fitur-fitur dalam data latih. Selain itu, model dilatih untuk mengidentifikasi pola dalam data yang dapat digunakan untuk memprediksi penjualan di masa mendatang.
5. Evaluasi Model
Setelah model dibangun, langkah selanjutnya adalah menilainya dengan data Testing. Ini dilakukan dengan menilai akurasi, presisi, recall, dan skor F1 untuk mengetahui seberapa baik model tersebut dalam memprediksi penjualan. Jika model tidak memenuhi persyaratan akurasi yang diharapkan, mungkin diperlukan penyesuaian atau tuning.
6. Analisis Hasil
Untuk memprediksi penjualan bahan bangunan di masa mendatang, model yang telah dievaluasi dan dianggap cukup akurat digunakan. Analisis hasil melibatkan membandingkan hasil prediksi dengan data aktual dan menemukan pola atau tren dalam hasil prediksi. Prediksi dibuat untuk mengetahui seberapa akurat model dalam kondisi nyata.

C. Hasil dan Pembahasan

Penelitian ini menghasilkan nilai akurasi dengan menggunakan metode *naive bayes* dalam memprediksi penjualan bahan bangunan dengan melakukan perhitungan dan pengolahan data secara manual menggunakan *microsoft excel* dan menggunakan *software jupyter notebook*.

1. Dataset penjualan terdiri dari tujuh atribut: dari tanggal transaksi tahun, nama barang, jenis barang, permintaan barang, penjualan, dan kategori penjualan. Persiapan data mencakup semua langkah-langkah yang diperlukan untuk membuat dataset penjualan yang akan digunakan dalam alat pemodelan. Dataset ini dibangun dari data mentah awal sebelum proses data mining.
2. Seleksi Data Seleksi data adalah proses memilih dan menganalisis data yang relevan dari database. Karena tidak semua data dalam database akan digunakan untuk pengujian, seleksi data dilakukan untuk memilih dan menganalisis data tersebut. enam data yang dipilih sebagai hasil seleksi ini digunakan sebagai indikator penelitian: dari tahun, nama barang, jenis barang, permintaan barang, penjualan, dan kategori penjualan.
3. Analisa Proses Algoritma Pada tahap ini, yaitu melakukan pemodelan data, penelitian ini menggunakan probabilitas (perkiraan) dengan algoritma Naive Bayes untuk mengolah data untuk memanfaatkan peluang. Metode ini dapat digunakan untuk menilai prospek masa depan. Dengan pengalaman yang sebelumnya sebagai referensi. Data yang akan diuji dibagi menjadi dua bagian, pelatihan dan pengujian, dan kemudian dianalisis dengan menggunakan software Jupyter notebook. untuk data pelatihan 400 record (tahun 2021-2023) dan data testing 100 record.
4. Probabilitas kriteria Tahun Berdasarkan data penjualan barang bahan bangunan UD Kiki Fatmala pada data set di atas, diketahui bahwa ada 500 data. Dari 500 data tersebut, 79% adalah penjualan barang yang laku dan 21% adalah penjualan barang yang tidak laku pada

tahun 2021. 57% penjualan barang laku dan 43% penjualan barang tidak laku Pada 2022, 65% penjualan barang laku dan 35% penjualan barang tidak laku Pada tahun 2023.

Tabel 1 probabilitas Kriteria Tahun

P(T)	Laku	Tidak	Jumlah
2021	79%	21%	100%
2022	57%	43%	100%
2023	65%	35%	100%

5. Probabilitas Kriteria jenis Barang

Berdasarkan data penjualan barang dari dataset di atas, diketahui bahwa ada beberapa jenis barang tersebut, kriteria jenis barang tercantum dalam tabel Probabilitas berikut:

Tabel 2 Probabilitas Kriteria Jenis Barang

P(JB)	Laku	Tidak
Kran	1,30%	0,51%
Balon	0,97%	1,03%
Ban	1,95%	1,03%
Baut	3,25%	4,14%
Barbel	0,32%	0,51%
Tang	0,32%	1,03%
Terminal	0,65%	2,59%
Tungku	0,32%	0,95%
Tjpr	0,32%	0%
waring	0,32%	0,51%

6. Implementasi Naive Bayes

Naive Bayes adalah sebuah algoritma klasifikasi yang didasarkan pada Teorema Bayes dengan asumsi "naive" atau sederhana bahwa setiap fitur dalam data bersifat independen satu sama lain, atau dengan kata lain, tidak ada hubungan antara satu fitur dengan fitur lainnya. Ini adalah salah satu algoritma yang paling sederhana namun sering kali sangat efektif untuk masalah klasifikasi, terutama ketika dataset yang digunakan memiliki banyak fitur atau atribut.

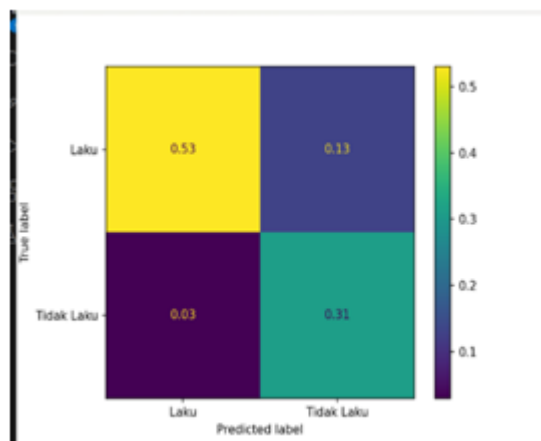
Tabel 3 Laporan Klasifikasi

akurasi 84.00%

laporan klasifikasi

	precision	recall	f1-score	support
laku	0.95	0.8	0.87	66
tidak	0.7	0.91	0.79	34
accuracy			0.84	100
macro avg	0.83	0.86	0.83	100
weighted avg	0.86	0.84	0.84	100

Hasil akurasi yang tercantum adalah 84.00%, yang menunjukkan bahwa model klasifikasi berhasil memprediksi dengan benar 84 dari 100 data uji. Ini adalah ukuran performa keseluruhan model, tetapi untuk memahami lebih dalam, kita perlu melihat metrik lainnya dalam table.



Gambar 2 Confusion Matrix

Gambar diatas menjelaskan sebuah confusion matrix yang menggambarkan performa model klasifikasi saya dalam memprediksi dua kelas: "Laku" dan "Tidak Laku".

D. Kesimpulan

Penelitian ini berhasil menerapkan metode Naive Bayes dalam memprediksi penjualan bahan bangunan dengan menggunakan dataset yang telah diproses secara manual menggunakan Microsoft Excel dan diimplementasikan menggunakan Jupyter Notebook. Dengan akurasi model sebesar 84%, hasil penelitian menunjukkan bahwa metode Naive Bayes cukup efektif untuk memprediksi apakah barang akan laku atau tidak berdasarkan fitur-fitur yang ada seperti jenis barang dan permintaan barang.

E. Saran

Saran untuk penelitian selanjutnya yaitu: Pengumpulan Data Lebih Lanjut Untuk meningkatkan performa model, disarankan untuk memperluas dataset dengan memasukkan data dari tahun-tahun sebelumnya atau atribut tambahan yang mungkin mempengaruhi penjualan, seperti promosi atau musim penjualan.

F. Referensi

- Ali Muthohar, A. D., Pribadi, T., & Sulistiawan, A. (2024). PENERAPAN METODE NAIVE BAYES UNTUK PREDIKSI STOK BARANG BANGUNAN DI TOKO BANGUNAN REJO MULYO. *Multidisciplinary Applications of Quantum Information Science (Al-Mantiq)*, 4(1), 1-7.
- Apriansyah, M. R. (2019). Pengembangan Media Pembelajaran Video Berbasis Animasi Mata Kuliah Ilmu Bahan Bangunan Di Program Studi Pendidikan Teknik Bangunan Fakultas Teknik Universitas Negeri Jakarta (Doctoral dissertation, UNIVERSITAS NEGERI JAKARTA).
- Ardiansyah, D. (2019). Algoritma c4. 5 untuk klasifikasi calon peserta lomba cerdas cermat siswa smp dengan menggunakan aplikasi rapidminer. *Jurnal Inkopar*, 1(2).
- Asyrofi, R. R., & Asyrofi, R. (2023). Implementasi Aplikasi Jupyter Notebook Sebagai Analisis Kreteria Plagiasi Dengan Teknik Simantik. *JUPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)*, 8(2), 627-637.
- Djamaludin, M. A., Triayudi, A., & Mardiani, E. (2022). Analisis sentimen tweet kri nanggala 402 di twitter menggunakan metode naïve bayes classifier. *Jurnal JTIC (Jurnal Teknologi Informasi dan Komunikasi)*, 6(2), 161-166.
- Fairul, R. S. P., & Saputri, G. (2023). PENERAPAN DATA MINING UNTUK PREDIKSI PENJUALAN CAT MENGGUNAKAN METODE NAIVE BAYES: STUDI KASUS: MITRA 10 GADING SERPONG. *LOGIC: Jurnal Ilmu Komputer dan Pendidikan*, 1(3), 571-578.
- Fatmawati, K., & Windarto, A. P. (2018). Data Mining: Penerapan rapidminer dengan K-means cluster pada daerah terjangkau demam berdarah dengue (DBD) berdasarkan provinsi. *Comput. Eng. Sci. Syst. J*, 3(2), 173.
- Hayami, R., & Oktaviandi, I. (2021). Penerapan metode single exponential smoothing pada prediksi penjualan bed sheet. *Jurnal CoSciTech (Computer Science and Information Technology)*, 2(1), 32-39.
- Hutahaean, M. (2022). Penerapan Data Mining untuk Memprediksi Penjualan Obat di Klinik Harapan Kita Batam (Doctoral dissertation, Prodi Teknik Informatika).
- Jollyta, D., Siddik, M., Mawengkang, H., & Efendi, S. (2021). Teknik Evaluasi Cluster Solusi Menggunakan Python Dan Rapidminer. Deepublish.
- Kalua, A. L., Mantiri, R., Rumondor, C., & Mogogibung, E. (2024). Sistem Informasi Pendaftaran Beasiswa dan Jadwal Legalisir Berbasis Website Responsif (Studi Kasus: Dinas Pendidikan Sulawesi Utara). *Journal of Information Technology, Software Engineering and Computer Science*, 2(2), 58-74.
- Komariyah, K., Dasuki, R., Saputra, D. B., Anwar, S., & Dwilestari, G. (2020). Klasifikasi Stok Barang Menggunakan Algoritma Naïve Bayes Pada Pt. Dharma Electrindo Manufacturing. *KOPERTIP: Scientific Journal of Informatics Management and Computer*, 4(2), 35-41.
- Laluan, N. T., & Raintung, M. C. (2022). PENGARUH STRATEGI PEMASARAN UNTUK MENINGKATKAN VOLUME PENJUALAN BAHAN BANGUNAN (STUDI KASUS UD

- BANGUN NUSANTARA MINAHASA). Jurnal EMBA: Jurnal Riset Ekonomi, Manajemen, Bisnis dan Akuntansi, 10(3), 783-794.
- Panennungi, P., & Pertiwi, N. (2018). Ilmu Bahan Bangunan.
- Salmawati, Yuyun, H. (2021). Klasifikasi Mahasiswa Berpotensi Drop Out Menggunakan Algoritma Decision Tree C4 . 5 Dan Naive Bayes Di Universitas Jambi. Jurnal Ilmiah Ilmu Komputer, 8(2), 115032. <https://repository.unja.ac.id/25341/>
- Samasil, S., Yuyun, Y., & Hazriani, H. (2022). Klasifikasi Mahasiswa Berpotensi Drop Out Menggunakan Algoritma Naive Bayes Dan Decision Tree. Jurnal Ilmiah Ilmu Komputer Fakultas Ilmu Komputer Universitas Al Asyariah Mandar, 8(2), 108-114.
- Sandy, S. (2022). Penerapan Algoritma Naive Bayes dalam Memprediksi Penjualan Material Bangunan (Doctoral dissertation, Prodi Teknik Informatika).
- Susanti, N. A., Walid, M., & Hoiriyah, H. (2022). Klasifikasi Data Tweet Ujaran Kebencian Di Media Sosial Menggunakan Naive Bayes Classifier. JATI (Jurnal Mahasiswa Teknik Informatika), 6(2), 538-543.