

PREDIKSI REKOMENDASI KATEGORI PRODUK MENGGUNAKAN METODE PREDICTIVE ANALYTICS NAIVE BAYES PADA PT KILAUMAS LESTARI MULIA

Nico Pratama¹, Rino^{2*}

^{1,2} Program Studi Teknik Informatika, Fakultas Sains dan Teknologi, Universitas Buddhi Dharma

*Corresponding Author, email: rino@ubd.ac.id

ABSTRAK

Di tengah lanskap bisnis digital yang kompetitif, pemanfaatan data strategis untuk pengambilan keputusan menjadi esensial. *Predictive analytics*, sebuah teknik canggih yang menganalisis data historis untuk mengidentifikasi pola dan memprediksi tren masa depan, terbukti efektif membantu organisasi merespons dinamika pasar lebih cepat dan tepat. Bagi PT Kilaumas Lestari Mulia, penerapan *predictive analytics* sangat relevan untuk memprediksi kategori produk diminati dari data pencarian pelanggan yang belum optimal, mengubahnya menjadi wawasan berharga. Penelitian ini mengusulkan algoritma Naive Bayes, metode klasifikasi probabilistik yang dikenal sederhana dan cepat dalam menangani data berbasis teks. Tujuannya adalah merancang sistem rekomendasi yang mampu secara akurat mengenali preferensi dan minat pelanggan. Dengan pemahaman ini, PT Kilaumas Lestari Mulia diharapkan dapat mengoptimalkan operasionalnya, seperti penyesuaian stok, perumusan strategi promosi, dan pengembangan pemasaran yang lebih personal. Sistem rekomendasi ini tidak hanya akan meningkatkan pengalaman pelanggan dengan menyajikan produk relevan, tetapi juga memberikan nilai strategis signifikan melalui wawasan tren permintaan pasar. Untuk memastikan validitas dan fungsionalitas, penelitian ini akan melibatkan *Black-Box Testing* untuk memverifikasi fungsionalitas internal dan *User Acceptance Testing* (UAT) untuk mengevaluasi penerimaan dan kepuasan pengguna akhir. Melalui pendekatan komprehensif ini, sistem yang dikembangkan diharapkan mampu meningkatkan daya saing dan kinerja bisnis PT Kilaumas Lestari Mulia di pasar yang terus berkembang.

Kata kunci: Data Pencarian, *E-commerce*, Naive Bayes, *Predictive Analytics*, Sistem Rekomendasi.

I. PENDAHULUAN

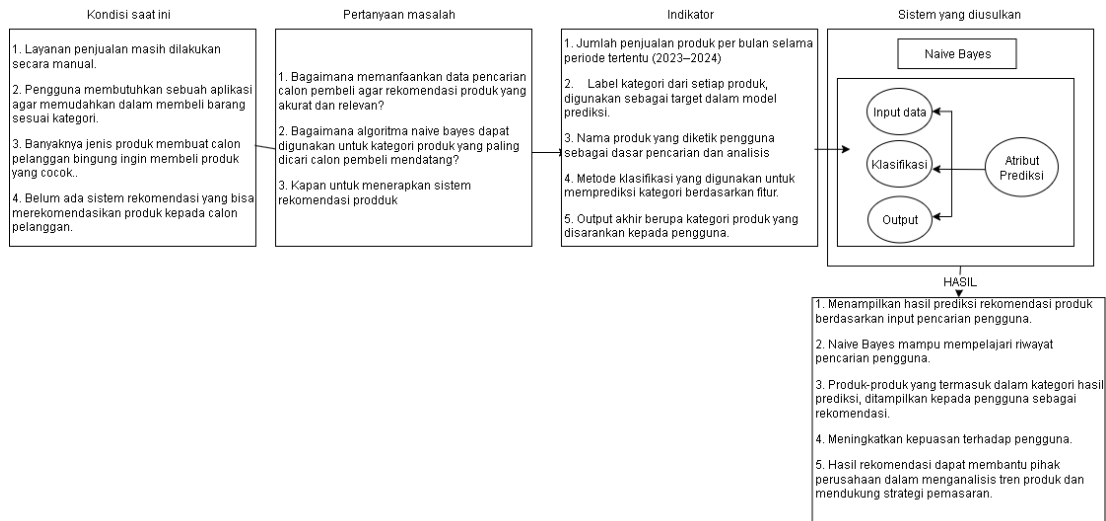
Era digital menuntut pemanfaatan data untuk keunggulan kompetitif. *Predictive analytics* krusial untuk meramalkan tren pasar dari data historis dan meminimalkan risiko rantai pasok (Aljohani, 2023). Bagi perusahaan, ini relevan mengoptimalkan data pencarian pelanggan guna prediksi kategori produk yang diminati. Data ini dapat menjadi dasar sistem rekomendasi kuat, memungkinkan perusahaan menyesuaikan stok, promosi, dan pemasaran secara efisien. Ini juga dapat meningkatkan loyalitas dan

penjualan, juga memberi wawasan strategis tentang tren permintaan (Fayyaz et al., 2020). Studi (Karabila et al., 2023) menunjukkan analisis sentimen dapat meningkatkan rekomendasi berbasis *collaborative filtering*, sementara (Shahbazi et al., 2020) mengeksplorasi *Extreme Gradient Boosting* untuk akurasi prediksi. Algoritma Naive Bayes dipilih dari kesederhanaan, kecepatan komputasi, dan kemampuannya dalam klasifikasi teks, menggunakan pendekatan probabilistik untuk menentukan kata kunci termasuk dalam kategori produk (Yang et al., 2023). Inovasi *Three-Way Incremental Naive Bayes* (TWNB) (Yang et al., 2023) meningkatkan performa dalam pembelajaran berkelanjutan. Penerapan *trust* dalam sistem rekomendasi menggunakan Naive Bayes turut meningkatkan akurasi dan kualitas rekomendasi (Rrmoku et al., 2022).

Pentingnya *big data* dan *predictive analytics* dalam *business intelligence* telah dikaji (Chen et al., 2022) dan (Bokonda et al., 2020). Berbagai aplikasi *machine learning* di industri telah dianalisa, dari prediksi penyakit jantung (Haryanto, 2024), peningkatan kualitas kredit perbankan (Kurniawan et al., 2022), dan tinjauan penerapan *ensemble learning* untuk prediksi penyakit (Mahajan et al., 2023). Ini selaras dengan tinjauan literatur sistematis di bidang sistem informasi dan ilmu komputer oleh (Priharsari, 2022), menegaskan posisinya dalam ilmu pengetahuan saat ini. Penelitian ini berfokus pada penerapan Naive Bayes untuk membangun sistem rekomendasi kategori produk berdasarkan data pencarian pelanggan, bertujuan menghasilkan sistem yang akurat dan efisien, serta meningkatkan daya saing PT Kilaumas Lestari Mulia.

II. METODOLOGI

Dalam melakukan perancangan aplikasi prediksi rekomendasi, peneliti memerlukan kerangka berpikir yang menjadikan tolak ukur penelitian dan pengembangan aplikasi.



Gambar 1. Tahapan Penelitian

Dapat disimpulkan sistem rekomendasi produk berbasis Naive Bayes sangat dibutuhkan. Sistem ini memanfaatkan data pencarian dan riwayat pembelian untuk merekomendasikan produk relevan. Tujuannya adalah meningkatkan kepuasan pengguna dan mendukung strategi pemasaran perusahaan melalui analisis tren. Penelitian ini mengadopsi pendekatan *predictive analytics* dengan algoritma Naive Bayes untuk membangun sistem rekomendasi kategori produk. Metode ini dipilih dalam mengklasifikasikan data historis penjualan untuk memprediksi kategori produk yang relevan. *Predictive analytics* merupakan gabungan statistik, *data mining*, *machine learning*, dan pemodelan prediktif untuk meramalkan kejadian masa depan dari data historis (Bokonda et al., 2020). Proses ini melibatkan pengumpulan data, pra-pemrosesan, pemodelan, dan evaluasi. Dimulai dari pengumpulan data bulanan produk dari perusahaan dalam format CSV. Data ini kemudian melewati pra-pemrosesan yang meliputi normalisasi, penyesuaian kolom, dan penanganan nilai kosong, serta penambahan total penjualan kumulatif sebagai fitur prediksi. Selanjutnya, model Naive Bayes dilatih menggunakan data penjualan sebagai fitur dan kategori produk sebagai label (Yang et al., 2023). Proses pelatihan ini mencakup penentuan label (Tinggi/Rendah) berdasar median total penjualan, perhitungan *prior* dan statistik per

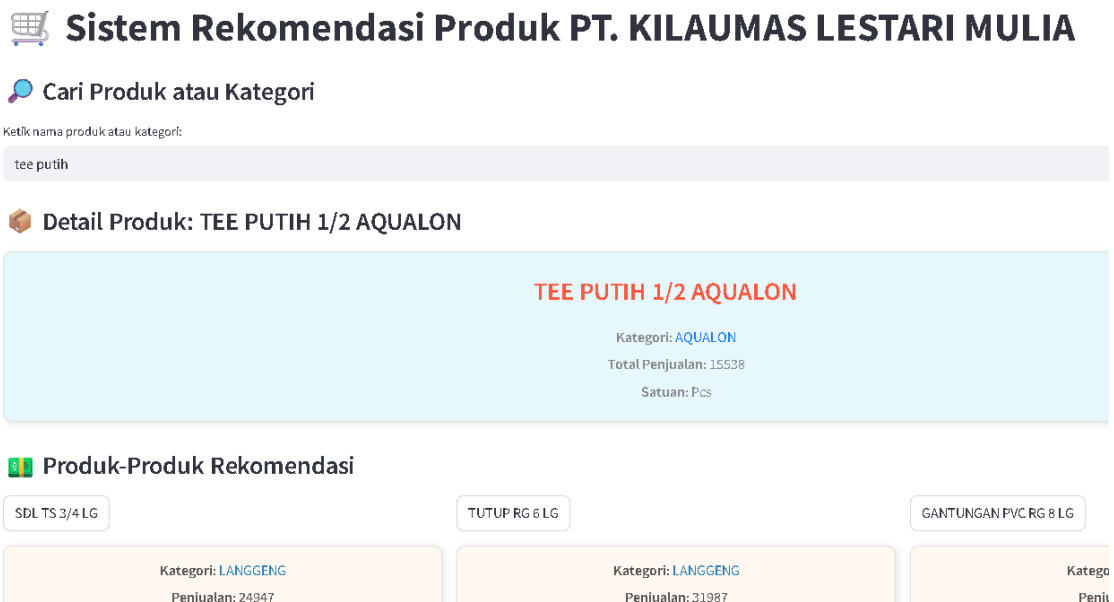
kelas, serta probabilitas Gaussian dan *posterior*. Stabilitas prediksi model sangat bergantung pada jumlah dan pola data yang memadai.

Setelah model dilatih, sistem menerima input pencarian pengguna untuk prediksi kategori produk. Implementasinya berupa aplikasi web interaktif menggunakan Streamlit (Buga et al., 2025). Streamlit didukung fitur dan integrasi pustaka populer. Lingkungan pengembangan terintegrasi (IDE) yang efisien, termasuk fitur kolaborasi *real-time* (Fan et al., 2019), juga mendukung proses implementasi. Pengujian sistem dilakukan melalui *Black-Box* Testing untuk memvalidasi output fungsional tanpa memeriksa struktur internal, dan *User Acceptance Testing* (UAT) yang melibatkan pengguna akhir. UAT memastikan aplikasi memenuhi kebutuhan bisnis dan mudah digunakan, sejalan dengan pentingnya pengalaman pengguna dan interaktivitas sistem personalisasi untuk keberhasilan penerimaan system (Troussas et al., 2023). Berikut adalah perumusan naïve bayes,

$$P(C|X) = \frac{P(X|C) \cdot P(C)}{P(X)}$$

III. HASIL DAN PEMBAHASAN

Berikut adalah tampilan dari aplikasi,



Gambar 2. Tampilan Setelah Mencari Produk

Tampilan ini menunjukkan setelah melakukan pencarian produk dengan nama TEE PUTIH 1/2 AQUALON maka akan menunjukkan detail dari produk tersebut dan dibawahnya akan menampilkan produk – produk yang direkomendasikan. Jadi jika calon pembeli atau pengguna melakukan pencarian suatu produk maka akan merekomendasikan produk yang kemungkinan besar dibeli bersamaan dengan produk yang dipilih oleh pengguna.

Prediksi Kategori Berdasarkan Pencarian

Melakukan prediksi untuk kategori yang terakhir dicari: AQUALON

Hasil Perhitungan Naive Bayes Manual

Prior kelas tinggi: 0.4848

Prior kelas rendah: 0.5152

Posterior kelas tinggi: 1.0000

Posterior kelas rendah: 0.0000

Kategori penjualan produk AQUALON diprediksi: TINGGI

Gambar 3. Tampilan Prediksi Kategori

Gambar 3 menunjukkan bahwa hasil prediksi kategori dari kategori produk yang dicari adalah TINGGI. Dari hasil prediksi maka penjualan dari kategori produk AQUALON diprediksikan TINGGI. Hasil perhitungan:

Produk 'TEE PUTIH 1/2 AQUALON' dari kategori 'AQUALON'. Produk ini memiliki total penjualan aktual 15538, berdasarkan median penjualan di kategorinya, diklasifikasikan sebagai kelas TINGGI.

1. Persiapan Data dan Parameter:

Probabilitas Prior ($P(C)$) adalah probabilitas umum suatu kelas penjualan tanpa mempertimbangkan data spesifik produk. Probabilitas Kelas Tinggi ($P(CTinggi)$): 0.4848. Probabilitas Kelas Rendah ($P(CRendah)$): 0.5152. Data Penjualan Produk (X): 24 bulan data penjualan spesifik dari produk 'TEE PUTIH 1/2 AQUALON'. Parameter Gaussian (μ dan σ^2): Rata-rata (mean) dan variansi (variance) penjualan tiap bulan,

dihitung secara terpisah untuk semua produk di kelas 'Tinggi' dan kelas 'Rendah' kategori 'AQUALON'.

2. Menghitung Likelihood ($P(X|C)$):

$$P(X|C) = P(X_1|C)P(X_2|C) \dots P(X_{24}|C)$$

Dimana $P(x_i|C)$ dihitung menggunakan fungsi kepadatan probabilitas (PDF)

Gaussian:

$$P(x_i|C) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x_i - \mu)^2}{2\sigma^2}\right)$$

Untuk Kelas Tinggi ($P(X | C_{Tinggi})$), Setelah menghitung PDF Gaussian untuk setiap dari 24 bulan penjualan produk 'TEE PUTIH 1/2 AQUALON' berdasarkan mean dan variansi kelas Tinggi, lalu mengalikan semua hasilnya, kita mendapatkan nilai yang sangat kecil. Dari perhitungan logaritmik sebelumnya (yang lebih stabil), nilai ini setara dengan $\exp(-193.6410)$.

$$P(X|C_{Tinggi}) \approx 1.34 \times 10^{-85}$$

Untuk Kelas Rendah ($P(X | C_{Rendah})$), Sama seperti di atas, kita mengalikan 24 nilai PDF Gaussian untuk setiap bulan berdasarkan mean dan variansi kelas Rendah. Berdasarkan perhitungan logaritmik sebelumnya, nilai ini setara dengan $\exp(-3659.1749)$. $P(X|C_{Rendah}) \approx 0$ (Angka ini sangat mendekati nol sehingga komputer menampilkannya sebagai 0 karena keterbatasan presisi floating-point, ini disebut *underflow*). Menghitung Pembilang

$$(P(C) \times P(X | C))$$

Ini merupakan bagian dari rumus Bayes yang membandingkan "skor" setiap kelas. Pembilang untuk Kelas Tinggi ($P(C_{Tinggi}) \times P(X | C_{Tinggi})$) $0.4848 \times (1.34 \times 10^{-85})$, Hasilnya adalah $\approx 6.50 \times 10^{-86}$. Pembilang untuk Kelas Rendah ($P(C_{Rendah}) \times P(X | C_{Rendah})$) $= 0.5152 \times (0)$. Hasilnya adalah ≈ 0

3. Menghitung Penyebut ($P(X)$)

Ini dihitung dengan menjumlahkan pembilang dari semua kelas.

$$P(X) = (P(C_{Tinggi}) \times P(X | C_{Tinggi})) + (P(C_{Rendah}) \times P(X | C_{Rendah}))$$

4. Menghitung Probabilitas Posterior ($P(C | X)$)

Ini adalah hasil akhir, probabilitas bahwa produk termasuk dalam kelas C,

diberikan data penjualannya X. Probabilitas Kelas Tinggi($P(C_{Tinggi} | X)$) = Pembilang Tinggi/ $P(X) = (6.50 \times 10^{-86}) / (6.50 \times 10^{-86})$ Hasilnya adalah 1.0000. Probabilitas Kelas Rendah ($P(C_{Rendah} | X)$) = Pembilang Rendah/ $P(X) = (0) / (6.50 \times 10^{-86})$ Hasilnya adalah 0.0000. Keputusan Prediksi Untuk kelas Tinggi (1.0000) lebih tinggi daripada kelas Rendah (0.0000). Jadi, model memprediksi bahwa produk TEE PUTIH 1/2 AQUALON termasuk dalam kategori penjualan TINGGI.

IV. SIMPULAN

Penelitian ini menyimpulkan sistem rekomendasi kategori produk berbasis algoritma Naive Bayes dan *predictive analytics* dapat membantu PT Kilaumas Lestari Mulia dalam mengenali preferensi pelanggan dan memprediksi produk yang relevan dari data pencarian. Sistem ini meningkatkan efisiensi promosi dan pengelolaan stok. Dengan prediksi yang akurat dari *predictive analytics* dan klasifikasi kata kunci yang baik oleh Naive Bayes, sistem mendukung pengambilan keputusan berbasis data yang cepat. Penggunaan Streamlit juga menghasilkan *dashboard* interaktif yang *user-friendly*, mempermudah admin dan *user*. Sistem ini merupakan alat strategis yang mendorong analisis cepat, akurat, dan transformasi bisnis yang lebih efisien.

DAFTAR PUSTAKA

- Aljohani, A. (2023). *Predictive Analytics and Machine Learning for Real-Time Supply Chain Risk Mitigation and Agility*.
- Bokonda, P. L., Touhami, K. O., & Souissi, N. (2021). *Predictive Analytics Using Machine Learning: Review of Trends and Methods*.
- Buga, R., Buzea, C. G., Agop, M., Ochiuz, L., Vasincu, D., Popa, O., . . . Eva, L. (2025). *Streamlit Application and Deep Learning Model for Brain Metastasis Monitoring After Gamma Knife Treatment*.
- Chen, Y., Li, C., & Wang, H. (2021). *Big Data and Predictive Analytics for Business Intelligence: A Bibliographic Study (2000–2021)*.
- F, Fan, H., Li, K., Li, X., Song, T., Zhang, W., . . . Du, B. (2019). *CoVSCode: A Novel Real-Time Collaborative Programming Environment for Lightweight IDE*.

- Fayyaz, Z., Ebrahimian, M., Nawara, D., Ibrahim, A., & Kashef, R. (2020). *Recommendation Systems: Algorithms, Challenges, Metrics, and Business Opportunities*.
- Haryanto. (2024). *Analisis Big Data dan Artificial Intelligence (AI): dalam Industri Khususnya Prediksi Penyakit Jantung dengan Phyton*.
- Karabila, I., Nossayba, N., El-Ansari, A., & Alami, N. (2023). *Enhancing Collaborative Filtering-Based Recommender System Using Sentiment Analysis*.
- Kurniawan, A., Rifa'i, A., Nafis, M. A., Sefrida, N., & Patria, H. (2022). *Pemilihan Metode Predictive Analytics dengan Machine Learning untuk Analisis dan Strategi Peningkatan Kualitas Kredit Perbankan*.
- Mahajan, P., Udin, S., Hajati, F., & Moni, M. A. (2023). *Ensemble Learning for Disease Prediction: A Review*.
- Priharsari, D. (2022). *Systematic Literature Review di Bidang Sistem Informasi dan Ilmu Komputer*.
- Rrmoku, K., Selimi, B., & Ahmedi, L. (2022). *Application of Trust in Recommender Systems—Utilizing*.
- Shahbazi, Z., Hazra, D., Park, S., & Byun, Y. C. (2020). *Toward Improving the Prediction Accuracy of Product Recommendation System Using Extreme Gradient. Department of Computer Engineering*.
- Troussas, C., Krouska, A., Koliarakis, A., & Sgouropoulou, C. (2023). *Harnessing the Power of User-Centric Artificial Intelligence: Customized Recommendations and Personalization in Hybrid Recommender Systems*.
- Yang, Z., Ren, J., Zhang, Z., Sun, Y., Zang, C., Wang, M., & Wang, L. (2023). *A New Three-Way Incremental Naive Bayes Classifier*.