

Integrasi Machine Learning dalam Ritel: Tinjauan Komprehensif tentang Prediksi Harga, Analisis Data Pelanggan, dan Pemanfaatan Media Sosial

Andi Wafda^{1*}

¹ Universitas Islam Indonesia, Yogyakarta, Indonesia

* 22917022@students.uii.ac.id

Abstract

Penerapan teknologi machine learning dalam industri ritel telah memberikan dampak yang signifikan dalam beberapa aspek sehingga memiliki potensi besar untuk meningkatkan efisiensi operasional, meningkatkan kepuasan pelanggan, dan mengoptimalkan strategi pemasaran di sektor ritel. Tujuan dari tinjauan literatur ini adalah untuk mengevaluasi kontribusi penerapan machine learning dalam industri ritel, dengan fokus pada aspek Prediksi dan Optimasi Harga, Analisis Data Pelanggan dan Segmentasi Pasar RFM, serta Pemanfaatan Sosial Media. Metode yang digunakan melibatkan pencarian terhadap studi-studi terbaru yang relevan dan berkualitas tinggi dalam domain tersebut, dengan mempertimbangkan kriteria kebaruan, relevansi, dan kualitas metodologi penelitian. Selain itu, metode analisis lainnya mencakup evaluasi kritis terhadap setiap pendekatan dan teknik yang digunakan dalam setiap studi, untuk memastikan bahwa hasil yang dihasilkan dapat diandalkan dan dapat diterapkan secara praktis. Hasil tinjauan menunjukkan dalam aspek Prediksi dan Optimasi Harga, teknik seperti Artificial Neural Network (ANN) dan Random Forest telah terbukti meningkatkan akurasi prediksi harga produk, memungkinkan perusahaan ritel untuk menetapkan harga yang kompetitif dan merespons pasar secara efisien. Analisis Data Pelanggan dan Segmentasi Pasar RFM menggunakan algoritma clustering seperti K-Means dan Fuzzy C-Means untuk membagi pelanggan berdasarkan pola pembelian mereka, yang mendukung pengembangan strategi retensi pelanggan yang lebih efektif dan personalisasi pengalaman pelanggan. Sementara itu, Pemanfaatan Sosial Media melalui analisis sentimen dengan metode seperti Support Vector Machine (SVM) dan Naïve Bayes memberikan wawasan mendalam mengenai persepsi konsumen terhadap produk dan merek, yang membantu perusahaan untuk mengarahkan strategi pemasaran yang lebih tepat. Kesimpulan dari tinjauan literatur ini menegaskan pentingnya terus mengembangkan dan menerapkan teknologi machine learning dalam industri ritel. Meskipun telah ada kemajuan signifikan, masih ada beberapa gap penelitian yang perlu diisi, seperti integrasi data multi-channel yang lebih holistik, pengembangan personalisasi yang lebih mendalam, serta perlunya perhatian yang lebih besar terhadap aspek keamanan dan etika dalam penggunaan data konsumen. Selain itu, penelitian mendatang juga dapat mengeksplorasi kemungkinan penggunaan teknologi baru seperti deep learning dan ensemble learning untuk meningkatkan prediksi dan analisis dalam konteks ritel yang semakin dinamis dan kompetitif.

Keywords: Machine Learning; Ritel; Prediksi Harga, Analisis Data Pelanggan, Pemanfaatan Media Sosial

Pendahuluan

Industri ritel merupakan salah satu sektor yang terus bertransformasi secara signifikan dalam menghadapi era digitalisasi, terutama dengan masuknya teknologi digital yang mengubah lanskap bisnis secara fundamental. Perubahan ini didorong oleh pertumbuhan pesat teknologi informasi dan komunikasi, yang mempengaruhi cara perusahaan berinteraksi dengan pelanggan serta memberikan akses baru terhadap data yang sangat berharga. Dengan meningkatnya volume data yang besar dari berbagai sumber seperti transaksi penjualan, interaksi pelanggan, dan media sosial, kemampuan untuk menganalisis dan menggunakan data ini secara efektif menjadi kunci untuk mempertahankan keunggulan kompetitif (Sudirjo et al., 2023). Data-data ini menjadi aset berharga yang dapat memberikan wawasan mendalam tentang perilaku konsumen, pola pembelian, serta tren pasar yang dapat dimanfaatkan untuk meningkatkan efisiensi operasional dan mengoptimalkan strategi bisnis.

Salah satu alat yang memiliki peran krusial dalam ekstraksi nilai dari data ini adalah teknologi machine learning. Penerapan teknologi machine learning dalam industri ritel dapat memberikan solusi terhadap tantangan yang dihadapi retailer modern untuk lebih inovatif dan efisien dalam operasional serta strategi bisnis mereka. Teknologi machine learning menyediakan alat yang sangat kuat untuk memahami pola dan tren dalam data tersebut. Hal ini memungkinkan retailer membuat keputusan yang lebih tepat dan responsif terhadap perubahan pasar.

Tantangan yang dihadapi oleh retailer modern tidak hanya terbatas pada pengelolaan volume data yang besar, tetapi juga pada kebutuhan untuk menghadapi persaingan yang semakin ketat di pasar global. Dengan memanfaatkan teknologi machine learning secara efektif, perusahaan ritel dapat memperoleh keunggulan kompetitif dengan cara yang lebih efektif dari sebelumnya. Oleh karena itu, tinjauan literatur ini bertujuan untuk memberikan gambaran yang komprehensif mengenai penerapan teknologi machine learning dalam industri ritel, dengan fokus khusus pada tiga area utama: prediksi dan optimasi harga, analisis data pelanggan dan segmentasi pasar menggunakan model RFM (recency, frequency, monetary), serta pemanfaatan media sosial untuk pengambilan keputusan bisnis. Ketiga aspek ini masing-masing memiliki peran penting dalam peningkatan efisiensi dan efektivitas operasional serta pengalaman pelanggan.

Prediksi harga dalam konteks ritel adalah proses estimasi harga produk di masa depan berdasarkan analisis data historis dan faktor-faktor yang mempengaruhi harga, seperti tren pasar, pola musiman, perilaku konsumen, serta kondisi ekonomi dan kompetisi. Prediksi ini bertujuan untuk membantu bisnis merencanakan strategi penetapan harga yang lebih akurat dan mengantisipasi perubahan yang mungkin terjadi di pasar (Mohamed et al., 2022). Sedangkan Optimasi harga dalam konteks ritel adalah proses penentuan harga yang paling efektif untuk mencapai tujuan bisnis tertentu seperti memaksimalkan keuntungan, meningkatkan pangsa pasar, atau mengoptimalkan volume penjualan. Proses ini melibatkan penggunaan teknik analisis data dan model matematis untuk menemukan harga optimal yang mempertimbangkan faktor-faktor seperti permintaan, biaya, dan kompetisi (Dong et al., 2019). Aspek ini akan mengeksplorasi berbagai teknik dan model yang digunakan untuk memprediksi harga produk dan mengoptimalkan harga jual. Ini termasuk studi tentang bagaimana machine learning dapat membantu dalam menentukan harga yang tepat untuk produk yang berbeda dan

bagaimana model ini dapat diterapkan dalam skenario dunia nyata untuk meningkatkan margin keuntungan dan efisiensi operasional.

Analisis Data Pelanggan adalah proses pengumpulan, pengolahan, dan interpretasi data terkait dengan perilaku dan preferensi pelanggan untuk memahami kebutuhan, keinginan, dan pola pembelian mereka. Sedangkan Segmentasi Pasar adalah proses membagi pasar yang heterogen menjadi segmen-segmen yang lebih homogen berdasarkan karakteristik tertentu seperti demografi, perilaku, atau preferensi (Savitri et al., 2018). Salah satu pendekatan yang efektif untuk melakukan analisis data pelanggan dan segmentasi pasar adalah Model RFM (Recency, Frequency, Monetary) (Febriani & Putri., 2020). Model ini digunakan untuk mengukur nilai pelanggan berdasarkan tiga parameter utama: recency (kapan terakhir kali pelanggan melakukan pembelian), frequency (seberapa sering pelanggan melakukan pembelian), dan monetary (berapa banyak uang yang telah dihabiskan pelanggan) (Djun et al., 2024). Fokus utama adalah pada teknik segmentasi yang membantu mengelompokkan pelanggan ke dalam segmen yang berbeda berdasarkan perilaku mereka, yang pada akhirnya dapat membantu dalam merancang strategi pemasaran yang lebih tepat sasaran dan meningkatkan loyalitas pelanggan

Terakhir, Pemanfaatan media sosial dalam sektor ritel adalah penggunaan platform digital seperti Facebook, Instagram, Twitter, dan TikTok untuk mendukung kegiatan pemasaran, penjualan, dan pelayanan pelanggan. Media sosial memungkinkan perusahaan ritel untuk berinteraksi langsung dengan konsumen, mempromosikan produk, mengumpulkan umpan balik, dan membangun hubungan yang lebih erat dengan pelanggan (Kamilla & Hasanah., 2024). Melalui konten yang menarik, kampanye iklan yang ditargetkan, dan analisis data, ritel dapat meningkatkan kesadaran merek, menarik lalu lintas ke toko fisik atau online, serta mengidentifikasi tren dan preferensi konsumen (Setiawan., 2024). Media sosial juga memungkinkan ritel untuk merespons kebutuhan pelanggan secara real-time, meningkatkan pengalaman pelanggan, dan memperkuat loyalitas terhadap merek (Sukoco & Krisnaresanti., 2024). Aspek ini akan membahas bagaimana data yang diperoleh dari platform media sosial dapat digunakan untuk memahami persepsi dan preferensi konsumen, serta memprediksi perilaku pembelian. Dengan teknik seperti analisis sentimen dan analisis jaringan sosial, ritel dapat mengembangkan strategi pemasaran yang lebih efektif dan meningkatkan interaksi dengan konsumen.

Hasil dari ulasan ini, diharapkan dapat ditemukan berbagai pendekatan, metode dan teknik machine learning yang telah diterapkan serta hasil yang telah dicapai dalam penelitian sebelumnya. Selain itu, ulasan ini bertujuan untuk mendokumentasikan bagaimana teknologi machine learning telah digunakan untuk meningkatkan efisiensi operasional, merancang strategi pemasaran yang lebih efektif, dan memberikan wawasan mendalam tentang perilaku konsumen dalam konteks ritel (Aisah & Harto., 2024). Dengan demikian, melalui eksplorasi mendalam terhadap penerapan machine learning dalam tiga aspek ini diharapkan dapat mengidentifikasi praktik terbaik, potensi arah penelitian di masa depan, serta memberikan panduan yang berharga bagi pengembangan dan implementasi teknologi machine learning yang lebih efektif dalam konteks ritel. Dengan demikian, artikel ini tidak hanya menjadi kajian akademis, tetapi juga sumber informasi yang bermanfaat bagi praktisi dan pengambil keputusan di industri ritel untuk memaksimalkan nilai dari data mereka dan mencapai tujuan bisnis yang lebih besar.

State of the art dalam penerapan machine learning dalam industri ritel menunjukkan bahwa teknologi ini terus berkembang pesat. Teknik-teknik canggih seperti deep learning dan ensemble learning telah digunakan untuk meningkatkan prediksi harga, segmentasi pelanggan, dan analisis sentimen dengan tingkat akurasi yang semakin tinggi. Integrasi dengan teknologi big data dan kemampuan pengolahan data real-time juga semakin memungkinkan perusahaan ritel untuk mengambil keputusan yang lebih cepat dan tepat dalam menjawab dinamika pasar yang berubah-ubah.

Metode

Tinjauan literatur sistematis mengenai penerapan teknologi machine learning dalam industri ritel, langkah-langkah yang diambil mengikuti prosedur yang terstruktur untuk memastikan relevansi, kualitas, dan fokus dari kajian ini. Pertama, mengidentifikasi sumber pencarian dengan beberapa kata kunci yang dikembangkan terkait penerapan machine learning pada sektor retail. Kedua, mengidentifikasi berdasarkan kriteria yang akan diulas, yang merujuk pada proses seleksi dan pengelompokan sumber daya literatur yang akan dimasukkan dan dikecualikan dalam tinjauan literatur. Ketiga, melakukan penelusuran penelitian secara mendalam, diikuti dengan penilaian kritis, ekstraksi data, dan sintesis temuan-temuan sebelumnya (Mikalef et al., 2018).

Langkah pertama dalam tinjauan literatur adalah mengidentifikasi sumber pencarian yang relevan dan mengembangkan string pencarian yang sesuai dengan topik yang diteliti. Langkah awal dilakukan dengan memilih sumber informasi terpercaya seperti IEEE Xplore, ACM Digital Library, Google Scholar, PubMed, arXiv, atau Scopus dilakukan berdasarkan ketersediaan literatur terkait machine learning dalam konteks industri ritel. Selanjutnya string pencarian dikembangkan dengan menggunakan kata kunci yang relevan seperti "machine learning in retail", "price prediction", "customer data analysis", "social media utilization", dan sebagainya. Tambahan sinonim dan istilah terkait lainnya digunakan untuk memperluas cakupan pencarian. Sinonim dan istilah terkait lainnya juga dimasukkan untuk memperluas cakupan pencarian dan memastikan bahwa semua aspek penting dari topik penerapan machine learning dalam ritel dapat tercakup dalam hasil yang komprehensif.

Kriteria inklusi digunakan untuk memastikan bahwa literatur yang dipilih secara langsung relevan dengan topik penelitian yang ditetapkan. Pertama, penelitian harus berfokus pada penerapan machine learning dalam ritel dalam tiga aspek utama yaitu prediksi dan optimasi harga, analisis data pelanggan, dan pemanfaatan sosial media untuk strategi pemasaran. Kedua, rentang tahun publikasi yang relevan dibatasi hingga lima atau enam tahun terakhir untuk memastikan informasi yang diperoleh terkait dengan perkembangan terkini dalam bidang ini. Ketiga, penelitian harus dipublikasikan dalam jurnal atau konferensi yang diakui secara akademis atau diakui oleh industri yang dipertimbangkan. Hal ini memastikan bahwa informasi yang disajikan telah melalui proses peer-review yang ketat dan diakui keabsahannya oleh komunitas ilmiah dan industri. Selain itu, penelitian yang dipilih harus menggunakan metodologi yang kuat dan valid, baik itu pendekatan kuantitatif maupun kualitatif. Metodologi yang solid ini memastikan bahwa data yang dihasilkan dapat diandalkan dan hasil penelitian dapat diterapkan secara efektif dalam praktik.

Sementara kriteria inklusi membimbing pemilihan literatur, kriteria eksklusi diterapkan untuk memastikan bahwa literatur yang dipilih dalam tinjauan literatur

memberikan kontribusi yang signifikan dan relevan dalam memahami penerapan machine learning dalam meningkatkan efisiensi operasional dan strategi pemasaran di sektor ritel. Untuk menjaga fokus dan relevansi dalam tinjauan literatur mengenai penerapan machine learning dalam sektor ritel, beberapa kriteria eksklusi telah ditetapkan. Literatur yang tidak secara langsung mengaplikasikan machine learning dalam konteks industri ritel, hanya membahas secara umum tentang machine learning tanpa aplikasi spesifik dalam ritel, atau yang tidak relevan dengan rentang lima atau enam tahun terakhir akan dihindari. Selain itu, publikasi yang berasal dari sumber yang kurang diakui atau tidak melalui proses peer-review yang ketat serta penelitian dengan metodologi yang lemah juga akan dieksklusi.

Penilaian kualitas dalam konteks tinjauan literatur merupakan tahap penting untuk memastikan bahwa artikel-artikel yang dipilih memiliki relevansi, metodologi yang kuat, dan kontribusi yang signifikan terhadap pemahaman topik yang diteliti. Awalnya, dari sekitar 40 artikel yang diidentifikasi melalui strategi pencarian yang komprehensif, dilakukan evaluasi terhadap setiap artikel berdasarkan kriteria inklusi dan eksklusi yang telah ditetapkan sebelumnya. Setelah proses evaluasi, sebanyak 18 artikel memenuhi semua kriteria inklusi dan dipilih untuk dilanjutkan ke tahap penilaian kualitas lebih lanjut. Artikel-artikel ini kemudian dikelompokkan berdasarkan kategorinya yaitu prediksi dan optimasi harga, analisis data pelanggan menggunakan model RFM, dan pemanfaatan media sosial dalam sektor retail. Pengelompokan ini memudahkan untuk dilakukannya ekstraksi data yang sistematis, untuk mengidentifikasi kontribusi masing-masing terhadap pemahaman tentang penerapan machine learning dalam sektor ritel.

Ekstraksi data dalam konteks tinjauan literatur merupakan proses sistematis untuk mengumpulkan informasi yang relevan dari setiap artikel yang dipilih. Langkah dalam ekstraksi data yaitu mengidentifikasi dan mencatat metodologi penelitian yang digunakan dalam setiap artikel, serta pendekatan analisis data yang diterapkan. Selanjutnya, informasi penting lainnya yang diekstraksi mencakup karakteristik sampel yang digunakan dalam penelitian, seperti ukuran sampel, sumber data, dan jenis data yang dianalisis, serta temuan utama yang dihasilkan dari setiap artikel. Informasi ini kemudian disusun dalam suatu format yang memungkinkan untuk membandingkan dan mengontraskan temuan dari berbagai artikel, serta mengidentifikasi pola atau tren yang muncul. Proses ekstraksi data yang ini memastikan bahwa semua informasi yang relevan dapat disusun dengan baik dan siap untuk dianalisis lebih lanjut dalam sintesis temuan.

Sintesis temuan dalam tinjauan literatur merupakan langkah kritis untuk menyatukan dan menganalisis informasi yang telah diekstraksi dari berbagai artikel. Proses ini dimulai dengan membandingkan temuan-temuan utama dari setiap artikel, mencari pola atau kesamaan antara mereka. Tren dan tema utama dalam literatur diidentifikasi untuk memberikan pemahaman yang lebih dalam tentang bagaimana teknologi machine learning diterapkan dalam sektor ritel. Selain itu, sintesis temuan juga melibatkan penarikan kesimpulan yang mengintegrasikan hasil dari berbagai artikel untuk menjawab pertanyaan penelitian yang telah diajukan. Hal ini membantu dalam mengidentifikasi implikasi praktis dari temuan-temuan tersebut dalam konteks operasional dan strategi bisnis ritel, serta merumuskan arah penelitian yang potensial di masa depan. Dengan pendekatan sistematis ini, sintesis temuan memastikan bahwa tinjauan literatur memberikan kontribusi yang signifikan terhadap pemahaman dan aplikasi teknologi machine learning dalam industri ritel.

Hasil dan Pembahasan

Bagian ini memaparkan hasil tinjauan literatur tentang penerapan teknologi machine learning dalam sektor ritel, mencakup tiga aspek utama: prediksi dan optimasi harga, analisis data pelanggan dengan model RFM, serta pemanfaatan media sosial. Tinjauan ini didasarkan pada artikel-artikel yang telah lolos kriteria inklusi dan eksklusi serta telah melalui proses penilaian kualitas. Hasilnya memberikan gambaran mendalam tentang bagaimana machine learning dapat meningkatkan efisiensi operasional, memperbaiki strategi pemasaran, dan memberikan wawasan yang lebih baik tentang perilaku konsumen dalam industri ritel. Pada bagian ini, data disajikan dalam bentuk tabel yang merangkum metode dan temuan utama dari artikel-artikel tersebut, diikuti oleh pembahasan rinci mengenai masing-masing aspek yang ditinjau. Tabel ini dirancang untuk memudahkan pembaca dalam melihat perbandingan dan pola-pola yang muncul dari literatur yang diulas, serta memberikan pandangan yang komprehensif mengenai kontribusi teknologi machine learning dalam sektor ritel. Pembahasan ini menyoroti aplikasi praktis dan peluang untuk pengembangan lebih lanjut dalam penerapan machine learning di sektor ini.

Prediksi dan Optimasi Harga dalam Sektor Retail

Tabel 1. Prediksi dan Optimasi Harga dalam Sektor Retail

Penulis	Metode Penelitian	Hasil
Kurniati et al, 2024	Model: Jaringan Saraf Tiruan (Artificial Neural Network/ANN). Data: Observasi, wawancara, studi kepustakaan, dan analisis data penjualan (Jan 2021 - Jun 2021).	Model ANN dengan akurasi 87%.
Subbarayudu et al, 2023	Model: Regresi Random Forest dan Multiple Linear Regression. Data: Data historis penjualan.	Model random forest memiliki R-squared yang lebih tinggi dan MSE yang lebih rendah dibandingkan multiple linear regression.
Greenstein-Messica & Rokach, 2020	Model: Log-Log Demand Model, Gradient Boosting Machines (GBM), PEIL. Data: 18 bulan data dari toko department store online di Eropa. Metode: Pessimistic Prediction Interval Measure	<ul style="list-style-type: none"> PEIL meningkatkan akurasi prediksi dan profit promosi dibandingkan dengan model log-log demand. PEG memperpendek waktu pembelajaran untuk active learning.
Sarkar et al, 2023	Model: K-means Clustering, Regresi Linier dan Logistik. Data: Atribut kunjungan, detail pengunjung, riwayat pembelian, data web, dan wawasan kontekstual.	<ul style="list-style-type: none"> Kerangka kerja yang holistik untuk memprediksi pembelian online melalui harga dinamis. Penyesuaian harga berdasarkan segmen pelanggan. Potensi untuk pengembangan harga dinamis real-time dan perluasan ke sektor lain.
Harsha et al, 2019	Model: Framework pengolahan data dan pembelajaran mesin, formulasi optimasi OCP, program mixed-integer. Data: Data dari tiga retailer besar selama dua tahun.	Implementasi OCP meningkatkan keuntungan sebesar 7% sambil mempertahankan volume penjualan.
Kedia et al, 2020	Model: Model prediksi permintaan, pemrograman linear. Data: Data transaksi dan diskon, hasil eksperimen A/B untuk pengujian model.	Peningkatan pendapatan sebesar 1% dan margin kotor sebesar 0.81% dengan penerapan model dalam eksperimen A/B.

Penulis	Metode Penelitian	Hasil
Rifa'i, 2020	<p>Model: Kombinasi Fuzzy Radial Basis Function Neural Network (FRBFNN) dan optimasi menggunakan algoritma genetika.</p> <p>Data: Data time series harga CPO dari Januari 2009-Januari 2020 dan faktor-faktor ekonomi dan produksi CPO yang relevan.</p>	<ul style="list-style-type: none"> • MAPE pada data training sebelum optimasi: 11.7% • MAPE pada data testing sebelum optimasi: 9.4% • MAPE pada data training setelah optimasi: 10.2% • MAPE pada data testing setelah optimasi: 8.3%

Tabel ini menyajikan hasil dari beberapa penelitian terkait penerapan teknik machine learning untuk prediksi dan optimasi harga dalam konteks industri retail. Berbagai penelitian telah menunjukkan pendekatan inovatif dan signifikan dalam meningkatkan strategi penetapan harga. Penelitian oleh Kurniati et al. (2024) menggunakan model jaringan syaraf tiruan (ANN) untuk menentukan harga optimal produk aksesoris garmen, dengan data yang dikumpulkan melalui observasi, wawancara, studi kepustakaan, dan analisis data penjualan dari Januari hingga Juni 2021.

Model ini menunjukkan akurasi 87%, yang cukup tinggi dalam konteks prediksi harga. Hasil penelitian mengungkap bahwa harga prediksi lebih tinggi untuk produk renda dan aplikasi import, menunjukkan bahwa harga pasar untuk produk ini masih kompetitif dan ada potensi peningkatan harga untuk keuntungan lebih besar. Sebaliknya, harga prediksi lebih rendah untuk produk seperti Swarovski dan mote Cina mengindikasikan bahwa harga saat ini mungkin terlalu tinggi, yang dapat mengurangi daya tarik produk di pasar. Temuan ini menunjukkan bahwa model ANN dapat menjadi alat yang efektif dalam penetapan harga produk ritel, memungkinkan penyesuaian harga yang lebih akurat sesuai dengan dinamika pasar dan meningkatkan efisiensi serta profitabilitas bisnis.

Penelitian oleh Subbarayudu et al. (2023) berfokus pada penentuan titik harga yang paling menguntungkan untuk produk dan layanan dengan menggunakan model regresi Random Forest dan Multiple Linear Regression. Data yang digunakan berasal dari data historis penjualan, yang memungkinkan model untuk memperkirakan permintaan dan reaksi pembeli terhadap berbagai skenario harga. Hasil penelitian menunjukkan bahwa model Random Forest Regression memiliki nilai R-squared yang lebih tinggi dan Mean Squared Error (MSE) yang lebih rendah dibandingkan dengan Multiple Linear Regression, yang menunjukkan bahwa model ini lebih akurat dalam memprediksi harga yang optimal. Keunggulan model Random Forest dalam ketepatan prediksi harga menegaskan potensinya sebagai alat yang efektif untuk optimasi harga di sektor ritel, membantu perusahaan dalam menetapkan harga yang tidak hanya kompetitif tetapi juga menguntungkan. Temuan ini menyoroti pentingnya penggunaan teknik machine learning yang canggih untuk meningkatkan pengambilan keputusan terkait harga, yang dapat berdampak signifikan pada profitabilitas dan daya saing bisnis ritel.

Penelitian oleh Greenstein-Messica & Rokach. (2020) meneliti dampak elastisitas harga dalam konteks promosi produk di platform e-commerce, dengan menggunakan model Log-Log Demand, Gradient Boosting Machines (GBM), dan PEIL (Pessimistic Prediction Interval Measure). Data yang digunakan meliputi 18 bulan penjualan dari sebuah department store online di Eropa, memberikan dasar yang kuat untuk analisis prediktif tanpa memerlukan data historis elastisitas harga. Temuan menunjukkan bahwa penggunaan PEIL secara signifikan meningkatkan akurasi prediksi dan profit dari promosi, dibandingkan dengan model log-log demand tradisional. Selain itu, metode ini juga

memperpendek waktu pembelajaran melalui penerapan Pessimistic Error Growth (PEG) untuk active learning, memungkinkan respons yang lebih cepat terhadap perubahan kondisi pasar. Temuan ini menunjukkan potensi besar dalam penerapan model prediksi canggih untuk mengoptimalkan strategi harga dan promosi dalam e-commerce, dengan memberikan kemampuan bagi retailer untuk membuat keputusan harga yang lebih informatif dan tepat waktu, bahkan dalam kondisi data yang tidak lengkap atau penuh ketidakpastian.

Studi yang dilakukan oleh Sarkar et al. (2023) mengembangkan kerangka kerja untuk optimasi harga dalam e-commerce dengan fokus pada harga dinamis yang disesuaikan dengan segmen pelanggan. Mereka mengaplikasikan berbagai metode termasuk K-means Clustering, Regresi Linier dan Logistik, serta pendekatan statistik dan machine learning lainnya. Data yang digunakan mencakup atribut kunjungan, detail pengunjung, riwayat pembelian, data web, dan wawasan kontekstual, yang dianalisis menggunakan Web Mining dan Big Data. Hasilnya menunjukkan bahwa pendekatan ini berhasil menghasilkan kerangka kerja yang holistik untuk memprediksi perilaku pembelian online berdasarkan harga dinamis, dengan penyesuaian yang dipersonalisasi untuk setiap segmen pelanggan. Temuan ini tidak hanya mendukung pengembangan harga dinamis secara real-time, tetapi juga menyoroti pentingnya mempertimbangkan aspek etika dalam pengaturan harga untuk menghindari praktik yang merugikan konsumen.

Harsha et al. (2019) mengembangkan solusi untuk mengatasi tantangan penetapan harga lintas saluran dan lokasi dengan memperkenalkan Optimized Competitive Pricing (OCP). Mereka menggunakan data dari tiga retailer besar selama dua tahun untuk mengestimasi elastisitas harga yang spesifik terhadap lokasi. Metode yang mereka terapkan meliputi framework pengolahan data, pembelajaran mesin, formulasi optimasi OCP, dan program mixed-integer. Implementasi OCP berhasil meningkatkan keuntungan sebesar 7% sambil mempertahankan volume penjualan, dengan strategi menurunkan harga online dan menyesuaikan harga toko sesuai dengan lokasi. Pendekatan eksak yang mereka gunakan memberikan rekomendasi harga yang stabil dan prediktif, menghindari kesalahan yang mungkin terjadi dalam metode heuristik.

Studi yang dilakukan oleh Kedia et al. (2020) bertujuan untuk menentukan titik harga optimal untuk produk di e-commerce fashion dengan tujuan memaksimalkan pendapatan dan profitabilitas. Mereka mengembangkan model dinamis yang menggunakan prediksi permintaan dan elastisitas harga untuk menentukan harga harian yang optimal. Dengan menerapkan teknik pemrograman linear, penelitian ini menunjukkan bahwa implementasi model tersebut meningkatkan pendapatan sebesar 1% dan margin kotor sebesar 0.81%. Melalui eksperimen A/B, mereka berhasil menguji dan memvalidasi model ini dalam situasi nyata, yang mendukung keefektifan model harga dinamis dalam pengaturan harga produk secara lebih adaptif. Studi ini juga merencanakan pengembangan lebih lanjut untuk mengimplementasikan penetapan harga dinamis intra-hari, menunjukkan potensi untuk meningkatkan responsivitas dan fleksibilitas dalam strategi penetapan harga di masa depan.

Penelitian oleh Rifa'i (2020) bertujuan untuk memprediksi harga CPO menggunakan pendekatan kombinasi FRBFNN dan optimasi dengan algoritma genetika. Metode ini digunakan untuk mengatasi fluktuasi harga yang berpotensi merugikan negara produsen. Data yang digunakan mencakup data time series harga CPO dari Januari 2009 hingga Januari 2020, serta faktor-faktor ekonomi dan produksi CPO yang relevan. Hasil penelitian

menunjukkan bahwa setelah optimasi, model FRBFNN dapat menghasilkan prediksi harga CPO dengan tingkat error yang lebih rendah, dengan MAPE pada data testing mencapai 8.3%, menunjukkan akurasi yang baik dalam memprediksi fluktuasi harga CPO.

Secara kolektif, penelitian ini mengungkap potensi besar dari penggunaan machine learning dalam memprediksi dan mengoptimalkan harga di sektor retail. Dari berbagai pendekatan yang diterapkan, seperti penggunaan jaringan syaraf tiruan (ANN), Random Forest, dan model kombinasi untuk memprediksi elastisitas harga, hingga pengembangan solusi omnichannel dan aplikasi pada industri fashion dan komoditas, literatur ini menunjukkan bahwa integrasi teknologi ini dapat signifikan dalam meningkatkan akurasi prediksi dan efisiensi strategi harga. Hal ini tidak hanya berpotensi untuk meningkatkan pendapatan dan margin keuntungan, tetapi juga untuk mengurangi kerugian akibat fluktuasi harga yang tidak terkendali. Dengan demikian, pengembangan lebih lanjut dalam bidang ini berpotensi untuk memberikan manfaat yang lebih besar bagi industri retail secara keseluruhan, dengan memanfaatkan kemajuan teknologi untuk meningkatkan strategi dan keputusan bisnis yang lebih tepat waktu dan efektif.

Analisis Data Pelanggan dan Segmentasi Menggunakan RFM dalam Sektor Retail

Tabel 2. Analisis Data Pelanggan dan Segmentasi Menggunakan RFM dalam Sektor Retail

Penulis	Metode Penelitian	Hasil
Adiana et al, 2018	Model: Algoritma K-Means untuk clustering. Metode: Menggunakan CRISP-DM Data: Data pelanggan (RFM: Recency, Frequency, Monetary).	Terdapat 3 kelompok pelanggan: typical customer (30 pelanggan), superstar customer (8 pelanggan), dormant customer (89 pelanggan).
Wijaya et al, 2021	Model: Algoritma K-Means untuk clustering. Metode: Penggunaan metode Elbow dan Silhouette Coefficient Data: Data transaksi pelanggan (RFM: Recency, Frequency, Monetary).	<ul style="list-style-type: none"> Segmentasi pelanggan terbagi menjadi: superstar, everyday, occasional, dan dormant. Analisis demografi menunjukkan dominasi usia 16-20 tahun dan pekerjaan sebagai siswa di setiap segmen. Analisis perilaku menunjukkan hari Jumat dan Sabtu sebagai puncak aktivitas transaksi.
Monalisa, 2018	Model: Algoritma K-Means Metode: Silhouette Index Data: Data historis pembelian pelanggan.	Terdapat 3 kluster dengan Silhouette Index sebesar 0,589. Segmentasi pelanggan: - Kluster 1: 57 pelanggan ($R\uparrow F\downarrow M\downarrow$). - Kluster 2: 257 pelanggan. - Kluster 3: 8 pelanggan ($R\downarrow F\uparrow M\uparrow$).
Putra et al, 2020	Model: Algoritma Fuzzy C-Means, FP-Growth. Data: Data transaksi pelanggan (RFM: Recency, Frequency, Monetary).	Terdapat 3 kluster: bolu durian, brownies, dan bolu topping untuk klaster 1, talam untuk klaster 2, dan bolu durian untuk klaster 3.
Nisa & Heikal, 2022	Model: Clustering Metode: Analisis deskriptif Data: Data primer pelanggan.	Segmentasi pelanggan berdasarkan analisis RFM: - Kategori Paling Rendah: 2 pelanggan. - Kategori Sederhana: 12 pelanggan. - Kategori Sangat Loyalis: 16 pelanggan.
Agustino et al, 2023	Model: DBSCAN, Affinity Propagation (AP) Metode: Silhouette Scores dan Davies-Bouldin Index (DBI). Data: Dataset transaksi pelanggan dari tenant inkubator bisnis STIKOM Bali.	<ul style="list-style-type: none"> AP algorithm memiliki kinerja lebih baik dibandingkan DBSCAN dengan Silhouette Scores 0.699 dan DBI 0.429. AP merekomendasikan 4 segmen pelanggan dengan karakteristik yang jelas terkait variabel RFM.

Tabel ini menguraikan beberapa penelitian yang menggunakan model RFM (Recency, Frequency, Monetary) dan teknik clustering untuk melakukan analisis data pelanggan serta segmentasi pasar. Setiap penelitian bertujuan untuk memahami perilaku pembelian pelanggan, mengukur kesetiaan, dan merancang strategi pemasaran yang lebih tepat sasaran berdasarkan karakteristik pelanggan. Adiana et al. (2018) yang menggunakan segmentasi pelanggan menggunakan model RFM dan teknik clustering: Penelitian ini bertujuan untuk mengukur kesetiaan pelanggan terhadap produk UKM dengan menerapkan model RFM (Recency, Frequency, Monetary) dan algoritma clustering K-Means. Model RFM digunakan untuk mengevaluasi perilaku pelanggan berdasarkan tiga dimensi utama: seberapa baru pelanggan melakukan pembelian terakhir (Recency), seberapa sering mereka melakukan pembelian (Frequency), dan jumlah uang yang mereka habiskan (Monetary). Algoritma K-Means kemudian digunakan untuk mengelompokkan pelanggan menjadi tiga kelompok utama: typical customer (30 pelanggan), superstar customer (8 pelanggan), dan dormant customer (89 pelanggan).

Penggunaan Davies Bouldin Index digunakan untuk menentukan jumlah cluster optimal, sedangkan evaluasi menggunakan global silhouette dan variance menunjukkan kesesuaian yang baik antara data dalam klaster. Hasil dari segmentasi ini memberikan manfaat dalam merancang strategi pemasaran yang tepat untuk setiap kelompok pelanggan. Pendekatan ini tidak hanya memungkinkan identifikasi pelanggan yang paling berharga (superstar customer) dan yang berisiko mengalami churn (dormant customer), tetapi juga memberikan wawasan yang mendalam tentang perilaku konsumen dalam konteks UKM.

Penelitian yang dilakukan oleh Wijaya et al. (2021) bertujuan untuk mengoptimalkan strategi retensi pelanggan di Ova Gaming E-sports Arena menggunakan segmentasi pelanggan berdasarkan model RFM dan algoritma K-Means. Metode ini digunakan untuk memahami pola transaksi pelanggan berdasarkan tiga dimensi RFM: Recency (seberapa baru pelanggan melakukan transaksi), Frequency (frekuensi transaksi), dan Monetary (nilai transaksi). Data yang digunakan adalah data transaksi pelanggan yang mencakup informasi RFM tersebut. Hasil dari penelitian ini menunjukkan bahwa dengan menggunakan metode Elbow dan Silhouette Coefficient, peneliti berhasil mengidentifikasi empat segmen pelanggan yang optimal: superstar, everyday, occasional, dan dormant. Setiap segmen memiliki karakteristik yang berbeda, seperti tingkat aktivitas transaksi, nilai transaksi rata-rata, dan tingkat kegiatan yang berbeda dalam arena e-sports. Strategi retensi yang disarankan berdasarkan nilai Customer Lifetime Value (CLV) juga disesuaikan dengan karakteristik masing-masing segmen. program loyalitas untuk superstar, reward untuk everyday, publisitas kesehatan untuk occasional, dan informasi produk baru untuk dormant. Analisis demografi menunjukkan dominasi pelanggan berusia 16-20 tahun dengan pekerjaan sebagai siswa, yang merupakan wawasan penting untuk mengarahkan strategi pemasaran lebih lanjut. Selain itu, pola perilaku transaksi menunjukkan bahwa aktivitas transaksi mencapai puncaknya pada hari Jumat dan Sabtu, yang dapat digunakan untuk merancang promosi atau kegiatan khusus pada hari-hari tersebut guna meningkatkan volume penjualan dan keterlibatan pelanggan.

Penelitian yang dilakukan oleh Monalisa (2018) menggunakan model RFM dan algoritma K-Means untuk mengukur perilaku pembelian pelanggan berdasarkan data historis pembelian. Metode Silhouette Index digunakan untuk mengevaluasi kualitas klaster yang dihasilkan. Hasil analisis menunjukkan bahwa terdapat tiga kluster pelanggan

yang optimal dengan nilai Silhouette Index sebesar 0,589. Kluster 1 terdiri dari 57 pelanggan dengan karakteristik regency tinggi, frekuensi transaksi rendah, dan nilai moneter rendah ($R \uparrow F \downarrow M \downarrow$). Kluster 2 memiliki 257 pelanggan dengan karakteristik yang menengah. Sedangkan Kluster 3, yang terdiri dari 8 pelanggan, menunjukkan nilai pembelian terbaik dengan frekuensi dan nilai moneter tinggi serta regency rendah ($R \downarrow F \uparrow M \uparrow$), menunjukkan potensi sebagai pelanggan masa depan yang bernalih. Temuan ini dapat digunakan untuk merancang strategi pemasaran yang lebih terfokus untuk masing-masing kluster, meningkatkan retensi pelanggan, dan mengoptimalkan pengalaman pelanggan secara keseluruhan.

Penelitian yang dilakukan oleh Putra et al. (2020) fokus pada pengukuran Customer Lifetime Value (CLV) dan segmentasi pelanggan pada Feandra Cake menggunakan model RFM dan Algoritma Fuzzy C-Means. Mereka menggunakan data transaksi pelanggan yang meliputi Recency, Frequency, dan Monetary (RFM) untuk mengidentifikasi tiga kluster pelanggan yang berbeda. Klaster 1 menunjukkan nilai CLV tertinggi, menandakan loyalitas tinggi dengan karakteristik Recency rendah, Frequency tinggi, dan Monetary tinggi. Produk cross-selling yang direkomendasikan juga bervariasi sesuai dengan preferensi setiap klaster, seperti bolu durian, brownies, dan bolu topping untuk klaster 1, talam untuk klaster 2, dan bolu durian kembali untuk klaster 3. Hasil ini memberikan insight penting bagi Feandra Cake untuk mengimplementasikan strategi pemasaran yang lebih terfokus dan meningkatkan retensi pelanggan, terutama dalam mempertahankan pelanggan klaster 1 dengan pelayanan terbaik dan pendekatan pemasaran yang lebih spesifik.

Studi yang dilakukan oleh Nisa & Heikal (2022) bertujuan untuk menganalisis tingkat Recency, Frequency, dan Monetary (RFM) pelanggan pada Manja Beauty Skincare guna meningkatkan kesetiaan pelanggan melalui segmentasi dan strategi pemasaran yang tepat. Menggunakan data primer pelanggan, penelitian ini menerapkan analisis deskriptif dengan teknik pengumpulan data observasi dan wawancara. Hasil analisis menunjukkan segmentasi pelanggan berdasarkan RFM menjadi tiga kategori utama: Kategori Paling Rendah dengan 2 pelanggan, Kategori Sederhana dengan 12 pelanggan, dan Kategori Sangat Loyalis dengan 16 pelanggan. Rekomendasi strategi yang diusulkan adalah mempertahankan pelanggan dengan predikat Sangat Loyalis melalui layanan terbaik serta memberikan stimulus seperti diskon atau penawaran menarik kepada pemegang loyal card, yang diharapkan dapat meningkatkan retensi dan loyalitas pelanggan secara efektif.

Penelitian yang dilakukan oleh Agustino et al. (2023) membandingkan performa dua algoritma clustering, yaitu DBSCAN dan Affinity Propagation (AP), dalam segmentasi pelanggan berdasarkan analisis RFM menggunakan dataset transaksi pelanggan dari tenant inkubator bisnis STIKOM Bali. Metode evaluasi yang digunakan adalah Silhouette Scores dan Davies-Bouldin Index (DBI). Hasil penelitian menunjukkan bahwa AP algorithm memiliki kinerja lebih baik dibandingkan DBSCAN, dengan nilai Silhouette Scores mencapai 0.699 dan DBI sebesar 0.429. AP merekomendasikan 4 segmen pelanggan yang memiliki karakteristik yang jelas terkait variabel RFM, menunjukkan kemampuan AP dalam memberikan segmentasi yang lebih akurat dan mudah diinterpretasikan. Rekomendasi dari penelitian ini adalah untuk menggunakan AP dalam analisis CRM untuk mendapatkan pemahaman yang lebih baik mengenai perilaku dan preferensi pelanggan.

Secara kolektif, penelitian-penelitian ini menunjukkan bahwa penerapan machine learning dalam analisis data pelanggan dan segmentasi RFM memiliki potensi besar dalam meningkatkan pemahaman terhadap perilaku pelanggan serta merancang strategi pemasaran dan retensi yang lebih efektif di berbagai sektor industri, termasuk retail. Berbagai algoritma clustering seperti K-Means, Fuzzy C-Means, DBSCAN, dan Affinity Propagation telah berhasil digunakan untuk mengidentifikasi segmen pelanggan berdasarkan karakteristik RFM mereka, memungkinkan perusahaan untuk mengoptimalkan layanan dan menyesuaikan pendekatan pemasaran sesuai dengan kebutuhan setiap segmen. Dengan memanfaatkan teknik-teknik ini, bisnis dapat lebih baik dalam mempertahankan pelanggan yang sudah ada, menarik pelanggan baru, serta meningkatkan profitabilitas jangka panjang dengan strategi yang lebih terarah dan efisien.

Pemanfaatan Sosial Media dalam Sektor Retail

Tabel 3. Pemanfaatan Sosial Media dalam Sektor Retail

Penulis Tahun	Metode Penelitian	Hasil
Singgalen, 2024	Model: CRISP-DM dan SVM (Support Vector Machine). Metode: Sentiment Analysis, Social Network Analysis (SNA). Data: Data sentiment dan jaringan dari komunitas XR.	Sentiment Analysis: SVM tanpa SMOTE: Akurasi 84.33%, Precision 84.67%, Recall 99.28%, F-measure 91.39%. SNA Metrics: Kepadatan 0.000124, sentralisasi 0.001331, modularitas 0.999000.
Yi & Liu, 2020	Model: Hybrid Recommendation System (HRS) berbasis model regresi. Metode: Machine Learning, evaluasi menggunakan MAE, MSE, dan MAPE. Data: Data produk dan ulasan pelanggan dari sistem Unified Computing System (UCS).	HRS memiliki nilai MAPE hampir 98%, dengan MAE (Mean Absolute Error) sekitar 0.6, yang lebih rendah dibandingkan metode lain.
Ramadhan et al, 2023	Model: Naïve Bayes dan Support Vector Machine (SVM). Metode: pre-processing data teks Data: 4500 ulasan produk earphone nirkabel dari Tokopedia.	Akurasi: Naïve Bayes memberikan akurasi 77.03%, recall 73.03%, dan F-measure 74.41%. SVM menghasilkan presisi terbaik dengan nilai 76.05%.
Anwar, 2022	Model: Gaussian Naive Bayes dan Multinomial Naive Bayes. Metode: Scraping data, labelisasi, preprocessing. Data: Komentar pada video Instagram yang berisi review smartphone Vivo dan Oppo.	Akurasi: Gaussian Naive Bayes menghasilkan akurasi 73% dengan F-score 73%. Multinomial Naive Bayes menghasilkan akurasi sebesar 81% dengan F-score 83%,
Arasu et al, 2020	Model: Machine learning menggunakan WEKA. Metode: Pengumpulan data, analisis data menggunakan WEKA dan algoritma lain, dan evaluasi. Data: Data media sosial terkait perilaku dan preferensi pembelian konsumen.	Alat WEKA menunjukkan kinerja yang lebih baik dibandingkan dengan algoritma lain dalam hal presisi, recall, dan F-measure. Dapat diterapkan pada kampanye pemasaran di berbagai domain.

Tabel ini menggambarkan hasil dari beberapa penelitian terkait penerapan teknik machine learning untuk menganalisis dan memanfaatkan data dari platform sosial media dalam konteks strategi pemasaran dan interaksi konsumen. Berbagai penelitian telah menyoroti penggunaan teknik seperti analisis sentimen dan sistem rekomendasi. Dalam penelitian Singgalen (2024), fokus utama adalah untuk memahami persepsi konsumen terhadap produk Extended Reality (XR), khususnya Apple Vision Pro, melalui analisis sentimen dan jaringan sosial menggunakan metode CRISP-DM dan SVM. Metode

Sentiment Analysis yang diterapkan menggunakan SVM tanpa SMOTE menunjukkan hasil yang cukup baik, dengan akurasi mencapai 84.33%, precision sebesar 84.67%, recall mencapai 99.28%, dan F-measure sebesar 91.39%. Hasil ini menunjukkan bahwa SVM efektif dalam mengklasifikasikan sentimen dari data yang diperoleh dari komunitas XR terkait produk tersebut. Selain itu, analisis menggunakan Social Network Analysis (SNA) juga dilakukan untuk memahami struktur jaringan sosial komunitas XR. Metrik SNA seperti kepadatan (density), sentralisasi (centrality), dan modularitas (modularity) digunakan untuk mengukur karakteristik jaringan. Dalam konteks penelitian ini, nilai kepadatan jaringan sebesar 0.000124 menunjukkan tingkat keterhubungan yang rendah antara anggota komunitas dalam berbicara tentang produk XR. Sentralisasi sebesar 0.001331 menunjukkan tingkat pentingnya beberapa anggota dalam jaringan, sedangkan modularitas yang tinggi (0.999000) menunjukkan adanya struktur komunitas yang terpisah secara fungsional dalam topik pembicaraan terkait XR. Analisis juga mengidentifikasi kata kunci yang paling sering muncul dalam percakapan, seperti "apple," "tech," dan "headset." Kata kunci ini mencerminkan fokus utama komunitas XR terhadap merek Apple dan teknologi headset dalam diskusi mereka. Dengan demikian, penelitian ini memberikan wawasan yang dalam tentang bagaimana konsumen mempersepsi produk XR tertentu, serta dinamika sosial yang terlibat dalam komunitas tersebut melalui pemanfaatan analisis sentimen dan jaringan sosial.

Penelitian yang dilakukan oleh Yi & Liu. (2020), mereka menerapkan algoritma Machine Learning untuk mempelajari, menganalisis, dan mengklasifikasikan informasi produk dan informasi toko berdasarkan pengalaman pelanggan. Model yang dikembangkan adalah Hybrid Recommendation System (HRS) berbasis regresi, yang dievaluasi menggunakan metrik MAE (Mean Absolute Error), MSE (Mean Squared Error), dan MAPE (Mean Absolute Percentage Error). Data yang digunakan berasal dari produk dan ulasan pelanggan yang terdapat dalam sistem Unified Computing System (UCS). Hasil penelitian menunjukkan bahwa HRS mencapai nilai MAPE hampir 98%, menunjukkan tingkat akurasi yang sangat tinggi dalam prediksi preferensi pelanggan terhadap toko berdasarkan produk yang dibeli. Lebih lanjut, sistem ini memiliki MAE sekitar 0.6, yang menunjukkan tingkat kesalahan prediksi yang rendah dibandingkan dengan metode lain yang dievaluasi. Secara keseluruhan, kinerja HRS dalam penelitian ini menunjukkan performa yang lebih baik dibandingkan pendekatan lain dalam memprediksi preferensi pelanggan terhadap toko berdasarkan analisis produk dan ulasan pelanggan yang ada dalam sistem UCS. Hal ini menegaskan bahwa penggunaan algoritma Machine Learning, seperti HRS, dapat memberikan solusi yang efektif dalam mengoptimalkan pengalaman pelanggan dan strategi pemasaran berbasis data.

Studi yang dilakukan oleh Ramadhan et al. (2023) membandingkan berbagai metode analisis sentimen untuk memahami pola pengeluaran konsumen dan meningkatkan interaksi konten, terutama untuk produk earphone nirkabel di Tokopedia. Mereka menggunakan data dari 4500 ulasan produk earphone nirkabel. Dalam penelitian ini, dua model utama yang dibandingkan adalah Naïve Bayes dan Support Vector Machine (SVM). Metode yang digunakan mencakup tokenization, case folding, normalization, filtering, stopword removal, stemming, TF-IDF, hyperparameter tuning, dan confusion matrix untuk evaluasi. Hasil penelitian menunjukkan bahwa Naïve Bayes dengan hyperparameter tuning mencapai akurasi sebesar 77.03%, recall sebesar 73.03%, dan F-measure sebesar 74.41%. Sedangkan SVM dengan hyperparameter tuning memberikan hasil presisi terbaik

dengan nilai 76.05%. Studi ini mengindikasikan bahwa penggunaan metode analisis sentimen seperti Naïve Bayes dan SVM dapat efektif dalam memahami persepsi konsumen terhadap produk tertentu, seperti earphone nirkabel di platform e-commerce, dengan memberikan evaluasi yang lebih mendalam terhadap interaksi konten dan pola pengeluaran konsumen.

Studi yang dilakukan oleh Anwar (2022) menggunakan analisis sentimen untuk mengklasifikasikan komentar di Instagram terkait review smartphone Vivo dan Oppo, serta membandingkan kinerja dua algoritma Naive Bayes. Penelitian ini menggunakan data dari komentar pada video Instagram yang berisi review tentang smartphone Vivo dan Oppo. Dua model Naive Bayes yang dibandingkan adalah Gaussian Naive Bayes dan Multinomial Naive Bayes. Metode yang digunakan meliputi scraping data, labelisasi, preprocessing (termasuk case folding, tokenisasi, dan filtering), perhitungan term frequency (tf), serta klasifikasi sentimen. Hasil penelitian menunjukkan bahwa Gaussian Naive Bayes mencapai akurasi sebesar 73% dengan F-score 73%. Sementara itu, Multinomial Naive Bayes mencatat akurasi sebesar 81% dengan F-score 83%, menunjukkan kinerja yang lebih baik dibandingkan dengan Gaussian Naive Bayes dalam mengklasifikasikan sentimen dari komentar terkait smartphone Vivo dan Oppo di Instagram. Studi ini memberikan pemahaman yang lebih dalam tentang persepsi publik terhadap produk smartphone melalui platform media sosial, dengan menunjukkan bahwa penggunaan Multinomial Naive Bayes dapat memberikan hasil klasifikasi yang lebih akurat dalam konteks ini.

Studi yang dilakukan oleh Arasu et al. (2020) menggunakan analitik data media sosial (WEKA) untuk memprediksi perilaku konsumen online dan mengembangkan strategi pemasaran yang efektif. Penelitian ini membandingkan kinerja WEKA dengan beberapa algoritma machine learning lainnya dalam analisis data media sosial terkait perilaku dan preferensi pembelian konsumen. Metode yang digunakan mencakup pengumpulan data media sosial, analisis data menggunakan WEKA dan algoritma lain, serta evaluasi kinerja menggunakan metrik seperti presisi, recall, dan F-measure. Hasil penelitian menunjukkan bahwa WEKA menunjukkan kinerja yang lebih baik dibandingkan dengan algoritma lain dalam hal presisi, recall, dan F-measure. Alat WEKA juga terbukti lebih unggul dalam penerapan berbagai teknik penambangan data, aplikasi bisnis, dan metode analisis data. Kemampuannya untuk memprediksi perilaku konsumen dapat diterapkan secara efektif dalam kampanye pemasaran di berbagai domain seperti pendidikan online, kesehatan, dan musik, meningkatkan efisiensi strategi pemasaran berbasis data di era digital saat ini.

Secara kolektif, penelitian ini menyoroti kemajuan signifikan dalam penerapan machine learning dalam konteks retail, khususnya dalam memanfaatkan data dari sosial media untuk memahami perilaku konsumen, meningkatkan strategi pemasaran, dan mengoptimalkan pengalaman belanja. Studi-studi yang telah dipaparkan memberikan bukti bahwa teknik seperti analisis sentimen, sistem rekomendasi, dan prediksi perilaku konsumen dapat secara efektif diterapkan menggunakan algoritma seperti Naïve Bayes, SVM, dan WEKA. Hasil-hasil ini tidak hanya memberikan wawasan mendalam tentang preferensi dan persepsi konsumen terhadap produk, tetapi juga menunjukkan potensi untuk meningkatkan efisiensi dan efektivitas strategi bisnis di berbagai platform e-commerce dan sektor ritel secara lebih luas.

Meskipun terdapat kemajuan yang signifikan, masih terdapat beberapa gap penelitian yang perlu diisi dalam konteks penerapan machine learning dalam industri ritel.

Pertama, integrasi data multi-channel dari saluran penjualan online dan offline membutuhkan pendekatan yang lebih holistik untuk mendapatkan pemahaman yang lebih baik terhadap perilaku konsumen secara menyeluruh. Kedua, pengembangan teknik personalisasi yang lebih mendalam diperlukan untuk memahami preferensi individual konsumen dengan lebih baik, termasuk faktor-faktor psikologis dan kontekstual yang mempengaruhi keputusan pembelian. Ketiga, aspek keamanan dan etika dalam penggunaan data konsumen untuk machine learning perlu mendapat perhatian serius, dengan perluasan penelitian untuk mengembangkan kerangka kerja yang sesuai untuk melindungi privasi konsumen sambil memungkinkan inovasi dalam penerapan teknologi ini.

Kesimpulan

Dari hasil literature review, dapat disimpulkan bahwa penerapan machine learning dalam industri ritel memberikan kontribusi signifikan dalam beberapa aspek kunci. Pertama, dalam aspek Prediksi dan Optimasi Harga, teknik seperti Artificial Neural Network (ANN) dan Random Forest telah terbukti efektif dalam meningkatkan akurasi prediksi harga produk. Ini membantu perusahaan ritel untuk menetapkan harga yang kompetitif, mengoptimalkan strategi promosi, serta merespons dinamika pasar dengan lebih cepat dan tepat. Kedua, dalam aspek Analisis Data Pelanggan dan Segmentasi Pasar RFM, penggunaan algoritma clustering seperti K-Means dan Fuzzy C-Means telah berhasil dalam membagi pelanggan berdasarkan pola pembelian (RFM) mereka. Hal ini memungkinkan perusahaan untuk merancang strategi retensi yang lebih efektif, memaksimalkan nilai pelanggan sepanjang masa (CLV), dan mengoptimalkan pengalaman pelanggan secara personal. Ketiga, dalam aspek Pemanfaatan Sosial Media, analisis sentimen dengan menggunakan metode seperti Support Vector Machine (SVM) dan Naïve Bayes memberikan wawasan yang berharga mengenai persepsi konsumen terhadap produk dan merek, yang dapat digunakan untuk mengarahkan strategi pemasaran yang lebih tepat sasaran. Secara keseluruhan, meskipun telah terjadi kemajuan yang signifikan dalam penerapan machine learning dalam industri ritel, ada banyak potensi untuk penelitian lebih lanjut yang dapat mengisi gap-gap ini dan membawa manfaat lebih lanjut bagi perusahaan-perusahaan ritel dalam menghadapi tantangan pasar yang semakin kompleks dan dinamis di masa depan.

Acknowledgment

Daftar Pustaka

- Adiana, B. E., Soesanti, I., & Permanasari, A. E. (2018). Analisis segmentasi pelanggan menggunakan kombinasi RFM model dan teknik clustering. *Jurnal Terapan Teknologi Informasi*, 2(1), 23-32.
- Agustino, D. P., Budaya, I. G. B. A., Harsemadi, I. G., Dharmendra, I. K., & Pande, I. M. S. A. (2023). Comparison of the DBSCAN algorithm and affinity propagation on business incubator tenant customer segmentation. *Jurnal Sisfokom (Sistem Informasi dan Komputer)*, 12(2), 315-321.
- Aisah, S., & Harto, B. (2024). Analisis customer journey dalam meningkatkan customer experience pelanggan di Fore Coffee Shop yang berbasis teknologi. *EKOMA: Jurnal Ekonomi, Manajemen, Akuntansi*, 3(3), 971-986.

- Anwar, K. (2022). Analisa sentimen pengguna Instagram di Indonesia pada review smartphone menggunakan Naive Bayes. *KLIK: Kajian Ilmiah Informatika dan Komputer*, 2(4), 148-155.
- Arasu, B. S., Seelan, B. J. B., & Thamaraiselvan, N. (2020). A machine learning-based approach to enhancing social media marketing. *Computers & Electrical Engineering*, 86, 106723.
- Djun, S. F., Gunadi, I. G. A., & Sariyasa, S. (2024). Analisis segmentasi pelanggan pada bisnis dengan menggunakan metode K-Means clustering pada model data RFM. *JTIM: Jurnal Teknologi Informasi dan Multimedia*, 5(4), 354-364.
- Dong, J., Rao, B., Liu, Y., Jiang, L., Lu, W., & Guo, Q. (2019). Strategi penetapan harga untuk periode berbeda selama musim penjualan berikutnya untuk produk musiman. *Akses IEEE*, 8, 39479-39490.
- Febriani, A., & Putri, S. A. (2020). Segmentasi konsumen berdasarkan model recency, frequency, monetary dengan metode K-Means. *JIEMS (Journal of Industrial Engineering and Management Systems)*, 13(2).
- Greenstein-Messica, A., & Rokach, L. (2020). Machine learning and operation research-based method for promotion optimization of products with no price elasticity history. *Electronic Commerce Research and Applications*, 40, 100914.
- Harsha, P., Subramanian, S., & Ettl, M. (2019). A practical price optimization approach for omnichannel retailing. *INFORMS Journal on Optimization*, 1(3), 241-264.
- Kamilla, A. M., & Hasanah, R. U. (2024). Pemanfaatan aplikasi media sosial Instagram pada aktivitas digital public relations dalam meningkatkan customer trust online shop @Lessa. *Id. Journal of Comprehensive Science (JCS)*, 3(2), 308-317.
- Kedia, S., Jain, S., & Sharma, A. (2020). Price optimization in fashion e-commerce. *arXiv preprint arXiv:2007.05216*.
- Kurniati, I., Trilaksono, A. R., & Sulistiwyat, S. (2024). Pemanfaatan machine learning untuk optimasi harga dengan menggunakan model artificial neural network pada PT Asia Garment Accessories. *JEIS: Jurnal Elektro dan Informatika Swadharma*, 4(1), 59-71.
- Mikalef, P., Pappas, I. O., Krogstie, J., & Giannakos, M. (2018). Big data analytics capabilities: A systematic literature review and research agenda. *Information Systems and E-Business Management*, 16, 547-578.
- Mohamed, M. A., El-Henawy, I. M., & Salah, A. (2022). Price prediction of seasonal items using machine learning and statistical methods. *Computers, Materials & Continua*, 70(2).
- Monalisa, S. (2018). Segmentasi perilaku pembelian pelanggan berdasarkan model RFM dengan metode K-Means. *Query: Journal of Information Systems*, 2(1).
- Nisa, K., & Heikal, J. (2022). Strategi segmentasi pelanggan Manja Beauty Skincare dengan menggunakan analisa RFM model. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 6(1), 348-351.
- Putra, M. A. S., Monalisa, S., Julhandri, J., & Khoiru, I. (2020). Penerapan algoritma fuzzy C-means menggunakan model RFM dalam klasterisasi pelanggan pada Toko Kue Feandra Cake. *Jurnal Ilmiah Rekayasa dan Manajemen Sistem Informasi*, 6(1), 64-69.
- Ramadhan, F. A., Ruslan, R. R. P., & Zahra, A. (2023). Sentiment analysis of e-commerce product reviews for content interaction using machine learning. *Cakrawala Repository IMWI*, 6(1), 207-220.

- Rifa'i, A. (2020). Optimasi fuzzy artificial neural network dengan algoritma genetika untuk prediksi harga crude palm oil. *Jurnal Teknik Informatika dan Sistem Informasi*, 6(2).
- Sarkar, M., Wang, T., Li, H., Gupta, A., & Smith, J. (2023). Optimizing e-commerce profits: A comprehensive machine learning framework for dynamic pricing and predicting online purchases. *Journal of Computer Science and Technology Studies*, 5(4), 186-193.
- Savitri, A. D., Bachtiar, F. A., & Setyawan, N. Y. (2018). Segmentasi pelanggan menggunakan metode K-Means clustering berdasarkan model RFM pada klinik kecantikan (studi kasus: Belle Crown Malang). *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, 2(9), 2957-2966.
- Setiawan, H. (2024). Application of data mining using the support vector machine (SVM) method to analyze fashion retail products to determine trends. *Academia Open*, 9(1), 10-21070.
- Singgalen, Y. A. (2024). Social network analysis and sentiment classification of extended reality product content. *KLIK: Kajian Ilmiah Informatika dan Komputer*, 4(4), 2197-2208.
- Subbarayudu, Y., Reddy, G. V., Raj, M. V. K., Uday, K., Fasiuddin, M. D., & Vishal, P. (2023). An efficient novel approach to e-commerce retail price optimization through machine learning. In *E3S Web of Conferences* (Vol. 391, p. 01104). EDP Sciences.
- Sudirjo, F., Purwati, T., Widyastuti, W., Budiman, Y. U., & Manuhutu, M. (2023). Analisis dampak strategi pemasaran digital dalam meningkatkan loyalitas pelanggan: Perspektif industri e-commerce. *Jurnal Pendidikan Tambusai*, 7(2), 7524-7532.
- Sukoco, H., & Krisnaresanti, A. (2024). Penggunaan media sosial sebagai adaptasi strategi pemasaran UMKM saat pandemi Covid-19: Tinjauan literatur. *Jurnal Manajemen, Akuntansi, Ekonomi*, 3(1), 32-40.
- Wijaya, K. Z., Djunaidi, A., & Mahananto, F. (2021). Segmentasi pelanggan menggunakan algoritma K-Means dan analisis RFM di Ova Gaming E-Sports Arena Kediri. *Jurnal Teknik ITS*, 10(2), A300-A237.
- Yi, S., & Liu, X. (2020). Machine learning-based customer sentiment analysis for recommending shoppers, shops based on customers' reviews. *Complex & Intelligent Systems*, 6(3), 621-634.