

Analisis Klasifikasi Alat Tulis Kantor (ATK) Berbasis Machine Learning Menggunakan EDA

Machine Learning Based Office Stationery Classification (ATK) Analysis Using EDA

Desi Fitriana^{a,1}, Satrianansyah^{b,2}, Muhammad Nur Alamsyah^{c,3}

^{abc} Universitas Bina Insan, Jl. HM Soeharto No.Kel, Lubuk Kupang, Kec. Lubuk Linggau Sel. I, Kota Lubuklinggau, Sumatera Selatan, 31626, Indonesia
2102030007@mhs.univbinainsan.ac.id¹, satrianansyah@univbinainsan.ac.id²
muhammadnuralamsyah@univbinainsan.ac.id³

ABSTRAK

Kebutuhan akan pengelolaan alat tulis kantor (ATK) yang efisien semakin meningkat seiring dengan perkembangan dunia usaha dan digitalisasi proses bisnis. Pengelompokan ATK secara manual cenderung memakan waktu, tidak konsisten, dan rentan kesalahan. Oleh karena itu, penelitian ini bertujuan mengembangkan sistem klasifikasi ATK berbasis machine learning untuk mendukung pengambilan keputusan dalam manajemen inventaris. Penelitian dilakukan pada Toko Fotocopy Martin, dengan pendekatan kuantitatif menggunakan data penjualan ATK selama periode tertentu. Tahapan penelitian meliputi pengumpulan data, pembersihan data, analisis eksploratif menggunakan teknik Exploratory Data Analysis (EDA), serta penerapan dua algoritma machine learning, yaitu Random Forest Classifier dan Support Vector Machine (SVM). Labeling data dilakukan berdasarkan kategori produk Laris, Kurang Laris, dan Tidak Laris. Hasil menunjukkan bahwa algoritma SVM memberikan performa terbaik dengan akurasi sebesar 85%, precision 0.79, recall 0.67, dan F1-score 0.88, mengungguli Random Forest yang mencatat akurasi 80% dan F1-score 0.85. EDA juga berhasil mengungkap pola penting dalam data, seperti hubungan positif antara stok dan jumlah penjualan. Temuan ini menegaskan bahwa integrasi machine learning dan EDA mampu mengidentifikasi karakteristik produk secara lebih akurat, serta memberikan insight strategis untuk pengelolaan stok. Kesimpulannya, penggunaan metode klasifikasi berbasis machine learning sangat efektif dalam mendukung efisiensi operasional toko, menghindari kelebihan atau kekurangan stok, serta meningkatkan layanan kepada pelanggan.

Kata Kunci : Machine Learning, Klasifikasi Produk, Manajemen Inventori.

ABSTRACT

The need for efficient office stationery (ATK) management is increasing along with the development of the business world and the digitalization of business processes. Manual grouping of ATKs tends to be time-consuming, inconsistent, and error-prone. Therefore, this research aims to develop a machine learning-based ATK classification system to support decision making in inventory management. The research was conducted at the Martin Photocopy Shop, with a quantitative approach using ATK sales data over a certain period. The research stages include data collection, data cleaning, exploratory analysis using Exploratory Data Analysis (EDA) techniques, as well as the application of two machine learning algorithms, namely the Random Forest Classifier and Support Vector Machine (SVM). Data labeling is carried out based on the categories of Best-Selling, Less-Selling and Non-Selling products. The results show that the SVM algorithm provides the best performance with accuracy of 85%, precision 0.79, recall 0.67, and F1-score 0.88, ahead of Random Forest which recorded accuracy of 80% and F1-score 0.85. EDA also succeeded in uncovering important patterns in the data, such as the positive relationship between stocks and sales quantities. These findings confirm that the integration of machine learning and EDA is able to identify product characteristics more accurately, as well as provide strategic insight for stock management. In conclusion, the use of machine learning-based classification methods is very effective in supporting store operational efficiency, avoiding excess or shortage of stock, and improving service to customers.

Keywords : Machine Learning, Product Classification, Inventory Management.

Info Artikel :

Disubmit:

Direview:

Diterima :

1. PENDAHULUAN

Penjualan merupakan kegiatan ekonomi yang melibatkan proses menjual barang atau jasa kepada konsumen dengan tujuan utama untuk memperoleh keuntungan. Aktivitas penjualan menjadi salah satu aspek penting dalam keberlangsungan usaha, karena secara langsung berhubungan dengan pendapatan dan pertumbuhan bisnis. Dalam hal ini, Toko Fotocopy Martin merupakan salah satu contoh usaha yang berfokus pada kegiatan jual beli Alat Tulis Kantor (ATK) sekaligus menyediakan jasa fotokopi bagi masyarakat. Usaha ini bergerak di bidang penyediaan layanan fotokopi dokumen serta penjualan berbagai jenis ATK, seperti kertas, pulpen, map, buku tulis, dan perlengkapan kantor lainnya. Dengan menggabungkan dua layanan utama, yaitu jasa fotokopi dan penjualan ATK, Toko Fotocopy Martin memiliki peluang untuk menarik lebih banyak pelanggan, baik dari kalangan pelajar, mahasiswa, pegawai kantoran, maupun masyarakat umum yang membutuhkan layanan dokumen dan perlengkapan tulis-menulis.

Strategi penjualan yang efektif serta pengelolaan inventori yang optimal menjadi kunci dalam mempertahankan keberlanjutan usaha dan meningkatkan keuntungan toko.[11] Terletak di lokasi strategis, toko ini melayani berbagai kalangan, mulai dari pelajar, mahasiswa, hingga profesional yang membutuhkan layanan fotocopy dan perlengkapan kantor.[12] Alat Tulis Kantor (ATK) merupakan salah satu kebutuhan esensial yang memiliki peran penting dalam mendukung kelancaran operasional berbagai jenis usaha dan kegiatan. Kebutuhan terhadap ATK tidak hanya terbatas pada lingkungan perkantoran, tetapi juga sangat vital di sektor pendidikan seperti sekolah, perguruan tinggi, serta di berbagai tempat usaha jasa, termasuk jasa fotokopi. Produk ATK meliputi berbagai jenis perlengkapan, mulai dari kertas, pulpen, pensil, spidol, penghapus, hingga perlengkapan pendukung lainnya yang digunakan dalam kegiatan administrasi, pembelajaran, maupun layanan publik. Permintaan yang konsisten terhadap ATK menjadikannya sebagai salah satu komoditas yang memiliki prospek pasar berkelanjutan, meskipun tingkat penjualan setiap jenis produk dapat berbeda tergantung kebutuhan spesifik konsumen di masing-masing sektor. Oleh karena itu, pengelolaan inventori ATK yang efektif menjadi hal yang krusial agar ketersediaan produk tetap terjaga, sekaligus menghindari risiko kelebihan atau kekurangan stok yang dapat mempengaruhi kualitas layanan dan operasional usaha. [13]

Alat Tulis Kantor (ATK) merujuk pada berbagai jenis barang yang digunakan untuk mendukung aktivitas administrasi, pencatatan, dokumentasi, dan kegiatan operasional lainnya, baik di lingkungan perkantoran, sekolah, maupun tempat usaha [14]. Produk-produk ATK meliputi kertas, pulpen, pensil, penghapus, map, dan perlengkapan lain yang esensial untuk kelancaran pekerjaan sehari-hari. Dalam kegiatan usaha, khususnya di bidang perdagangan ATK, data dari transaksi penjualan memiliki peran yang sangat penting karena mengandung informasi berharga yang dapat dimanfaatkan untuk analisis kinerja usaha. Data ini mencakup rincian seperti jumlah produk yang terjual, jenis produk, waktu transaksi, dan pola permintaan konsumen.

Namun demikian, dalam praktiknya, banyak pemilik bisnis yang belum mengetahui cara yang efektif untuk mengumpulkan, mengelola, dan memanfaatkan data penjualan tersebut secara optimal. Akibatnya, potensi data yang tersedia tidak tergali sepenuhnya untuk mendukung pengambilan keputusan strategis. Oleh karena itu, penerapan metode klasifikasi yang tepat pada data ATK menjadi sangat penting, terutama dalam konteks mendukung manajemen inventaris dan pengelolaan stok yang lebih efisien. Hal ini relevan diterapkan pada usaha seperti toko fotocopy, yang selain menyediakan layanan fotokopi, juga menjual berbagai jenis ATK dengan tingkat permintaan yang bervariasi.

Dalam bidang Machine Learning (ML), klasifikasi merupakan teknik yang digunakan untuk mengelompokkan data ke dalam kategori tertentu berdasarkan karakteristik atau fitur yang dimilikinya [1]. Dengan penerapan klasifikasi, data penjualan ATK dapat diolah menjadi informasi yang lebih mudah dipahami dan diinterpretasikan, sehingga mendukung pemilik usaha dalam membuat keputusan yang lebih tepat, seperti menentukan produk prioritas, mengatur jumlah stok, serta merencanakan strategi promosi yang efektif.

Pada penelitian sebelumnya menggunakan pendekatan tradisional, seperti analisis ABC, untuk mengelompokkan produk berdasarkan nilai dan permintaan. Namun, metode ini sering tidak memadai karena mengabaikan variabel penting seperti pola permintaan dinamis dan tanggal pengiriman.[2] Dalam lingkungan bisnis yang cepat berubah, analisis tradisional sering tidak menangkap dinamika permintaan yang berkembang. Pendekatan responsif dengan memanfaatkan teknologi modern, sehingga Toko Fotocopy Martin dapat beradaptasi lebih baik terhadap perubahan pasar. Machine Learning muncul sebagai solusi untuk masalah klasifikasi, termasuk pengelompokan produk ATK. Teknologi ini memungkinkan pengusaha mengidentifikasi variabel yang mempengaruhi permintaan dan stok, seperti jenis produk, harga, dan volume penjualan.[1] Dalam pengelolaan persediaan ATK, toko fotocopy martin menghadapi tantangan dalam mengklasifikasikan berbagai jenis produk yang dimiliki.[3]

Analisis adalah proses memecah atau menguraikan suatu masalah, informasi, atau data menjadi bagian-bagian yang lebih kecil untuk memahami struktur, hubungan, dan maknanya.[4] Tujuannya adalah untuk menemukan pola, hubungan, atau kesimpulan yang dapat membantu dalam pengambilan keputusan atau pemahaman lebih mendalam.[5] Dalam analisis klasifikasi ATK, teknologi membantu mengumpulkan data dari survei, transaksi penjualan. Data yang terkumpul dianalisis dengan EDA untuk mengidentifikasi kategori produk yang diminati dan memahami preferensi konsumen.

Analisis data merupakan tahap dalam penelitian yang dilakukan setelah semua informasi yang diperlukan untuk menyelesaikan masalah yang diteliti telah dikumpulkan sepenuhnya.[6] Ketelitian dan kejelasan penggunaan alat analisis sangat mempengaruhi keakuratan kesimpulan yang dihasilkan. Oleh karena itu, proses analisis data adalah langkah yang tidak boleh diabaikan dalam penelitian.[7] Analisis data dilakukan setelah semua informasi terkumpul dan bergantung pada ketelitian serta kejelasan alat analisis.[8] Kualitas analisis yang baik berdampak pada keakuratan kesimpulan, sehingga langkah-langkah yang tidak boleh diabaikan dalam penelitian.[9] Melalui proses klasifikasi model yang telah diketahui berfungsi untuk memprediksi objek yang belum teridentifikasi.[10] Proses ekstraksi informasi berharga dari kumpulan data yang besar dan kompleks.[15]

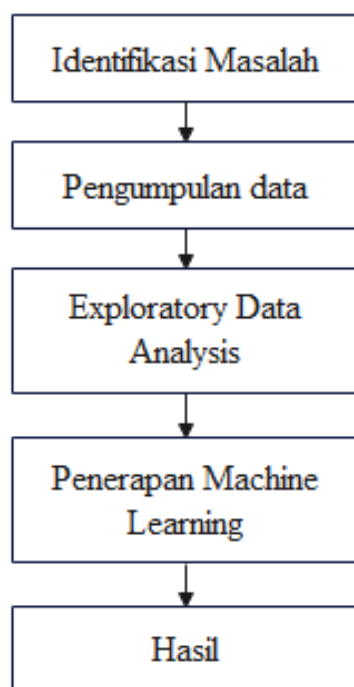
Tujuan utamanya adalah untuk menemukan pola, tren, atau hubungan yang mungkin tidak terlihat secara langsung.[16] Dengan menggunakan teknik statistik, algoritma pembelajaran mesin, dan analisis data, data mining membantu organisasi dalam pengambilan keputusan yang lebih baik dan mendukung strategi bisnis.[17] Penelitian ini bertujuan mengembangkan sistem klasifikasi berbasis machine learning untuk membantu Toko Fotocopy Martin mengelola stok ATK lebih efektif.[18] Serta dapat memenuhi permintaan pelanggan dengan lebih baik dan memberikan wawasan mendalam mengenai pola dan distribusi produk ATK yang mungkin terlewat melalui analisis manual.[19]

Penelitian ini memiliki kebaruan (novelty) dalam penerapan kombinasi metode Exploratory Data Analysis (EDA) dan algoritma Support Vector Machine (SVM) khusus untuk klasifikasi produk Alat Tulis Kantor (ATK). Sebagian besar penelitian terdahulu hanya menggunakan metode analisis tradisional seperti analisis ABC atau algoritma umum seperti Decision Tree untuk pengelompokan produk. Pendekatan ini berbeda karena mengintegrasikan EDA untuk memahami pola data secara mendalam sebelum pemodelan, kemudian menggunakan SVM yang terkenal efektif untuk data dengan dimensi tinggi dan batas klasifikasi yang kompleks. Hingga saat ini, belum ditemukan penelitian yang secara spesifik mengombinasikan SVM dan EDA dalam konteks klasifikasi ATK untuk mendukung

pengelolaan inventori. Hal ini menjadi pembeda utama penelitian ini dibandingkan studi sebelumnya yang umumnya berfokus pada optimasi stok dengan metode statistik sederhana atau algoritma tunggal.

2. METODE

Penelitian ini menerapkan metode kuantitatif dengan penekanan pada data yang diperoleh dari catatan data penjualan Alat Tulis Kantor (ATK) [20]. Data tersebut akan dianalisis menggunakan klasifikasi Random Forest Classifier dan Support Vector Machine (SVM). Metode Random Forest Classifier digunakan untuk mengidentifikasi pola kompleks dalam data penjualan ATK, sementara SVM digunakan untuk membangun model prediktif berdasarkan pola yang ditemukan. Analisis kuantitatif ini diharapkan dapat memberikan wawasan mendalam tentang faktor-faktor yang mempengaruhi penjualan ATK dan meningkatkan akurasi prediksi untuk perencanaan stok dan strategi pemasaran di masa depan. Proses penelitian dapat dilihat pada gambar berikut:



Gambar 1. Proses Penelitian

1) Identifikasi Masalah

Langkah pertama dalam penelitian ini adalah menentukan secara jelas masalah atau pertanyaan spesifik yang perlu ditangani. Identifikasi masalah dilakukan dengan menelaah kondisi aktual terkait penjualan Alat Tulis Kantor (ATK), termasuk tantangan yang dihadapi oleh perusahaan dalam memprediksi permintaan dan mengelola stok secara efektif. Dengan merumuskan masalah secara tepat, penelitian ini dapat memiliki fokus yang terarah untuk menemukan solusi melalui pendekatan analisis data.

2) Pengumpulan Data

Tahap berikutnya adalah pengumpulan data yang relevan dan akurat. Data yang dikumpulkan meliputi catatan penjualan ATK dalam periode tertentu, termasuk informasi terkait jenis produk, jumlah terjual, harga, waktu penjualan, dan faktor lain yang mempengaruhi

penjualan. Data ini diperoleh dari sistem pencatatan internal perusahaan atau sumber lain yang kredibel untuk memastikan validitas analisis selanjutnya.

3) Analisis Data Eksploratori (Exploratory Data Analysis/EDA)

Setelah data terkumpul, dilakukan analisis data eksploratori (EDA) untuk mengeksplorasi dan memahami karakteristik data. Proses ini mencakup identifikasi distribusi data, mendeteksi nilai ekstrem (outlier), mengatasi data yang hilang (missing values), serta menemukan pola, tren, dan hubungan antar variabel. EDA berperan penting sebagai dasar dalam menentukan strategi analisis lebih lanjut dan memastikan bahwa data siap digunakan untuk pemodelan.

4) Aplikasi Pembelajaran Mesin

Selanjutnya, dilakukan penerapan teknik pembelajaran mesin (machine learning) untuk membangun model prediktif. Penelitian ini menggunakan algoritma Random Forest Classifier dan Support Vector Machine (SVM) untuk menganalisis data penjualan ATK. Proses ini melibatkan pembagian data ke dalam data pelatihan dan data pengujian, pelatihan model dengan data pelatihan, dan evaluasi kinerja model menggunakan metrik evaluasi yang sesuai. Model yang dihasilkan diharapkan mampu mengklasifikasikan atau memprediksi hasil penjualan secara akurat.

5) Hasil

Tahap akhir adalah penyajian dan interpretasi hasil yang diperoleh dari model pembelajaran mesin. Hasil ini meliputi performa model berdasarkan metrik evaluasi seperti akurasi, presisi, recall, dan F1-score, serta analisis mengenai variabel-variabel yang berpengaruh signifikan terhadap penjualan ATK. Kesimpulan diambil berdasarkan interpretasi hasil, dengan memberikan rekomendasi strategis yang dapat diimplementasikan perusahaan untuk meningkatkan efektivitas pengelolaan penjualan dan stok ATK di masa mendatang.

3. HASIL DAN PEMBAHASAN

Penelitian ini dilakukan dengan tujuan untuk menemukan solusi yang efektif dalam manajemen inventori, khususnya dalam mengelola stok Alat Tulis Kantor (ATK) di Toko Fotocopy Martin. Selain itu, penelitian ini juga berfokus pada analisis klasifikasi produk ATK ke dalam dua kategori, yaitu Laris dan Tidak Laris, guna membantu pihak toko dalam mengidentifikasi produk-produk yang memiliki tingkat penjualan tinggi maupun rendah. Untuk mencapai tujuan tersebut, penelitian ini memanfaatkan pendekatan Exploratory Data Analysis (EDA) guna mengeksplorasi dan memahami pola serta tren data penjualan, yang kemudian dilanjutkan dengan penerapan teknik Machine Learning melalui algoritma Random Forest Classifier dan Support Vector Machine (SVM). Dengan penerapan algoritma ini, diharapkan dapat diperoleh model prediktif yang akurat dalam mengklasifikasikan produk, sehingga mendukung pengambilan keputusan yang lebih tepat terkait pengadaan dan pengelolaan stok ATK di Toko Fotocopy Martin.

Tujuan dari penelitian ini adalah untuk mengevaluasi dan membandingkan kinerja kedua algoritma, yaitu Random Forest Classifier dan Support Vector Machine (SVM), dalam melakukan klasifikasi produk Alat Tulis Kantor (ATK) ke dalam kategori Laris dan Tidak Laris. Penelitian ini diawali dengan proses pengumpulan data penjualan ATK yang diperoleh dari Toko Fotocopy Martin selama periode bulan Oktober. Berdasarkan hasil observasi awal, tercatat sebanyak 100 item produk yang termasuk ke dalam kategori ATK, masing-masing dengan jumlah stok yang bervariasi. Data yang terkumpul ini selanjutnya dianalisis menggunakan pendekatan Exploratory Data Analysis (EDA) untuk memahami karakteristik data dan mendeteksi pola penjualan, sebelum diterapkan ke dalam model pembelajaran mesin untuk proses klasifikasi. Dengan demikian, penelitian ini diharapkan dapat memberikan

gambaran mengenai efektivitas kedua algoritma dalam memprediksi tingkat kelarisan produk, yang nantinya dapat dimanfaatkan sebagai dasar pengambilan keputusan dalam pengelolaan inventori di Toko Fotocopy Martin. Setelah data terkumpul langkah selanjutnya adalah melakukan preprocessing, hasil data dapat dilihat pada tabel dibawah ini:

Tabel 1. Hasil Preprocessing

Hasil Preprocessing	
Data ATK	100

Setelah itu melakukan proses labeling untuk setiap produk berdasarkan jumlah terjual dan stok, diperoleh hasil:

Tabel 2. Hasil Kategori

Kategori	Hasil
Laris	37
Kurang Laris	36
Tidak Laris	21

Selanjutnya, penerapan klasifikasi menggunakan algoritma Random Forest Classifier dan Support Vector Machine dengan mengimplementasikan teknik Cross Validation K-Fold 10 sebagai berikut:

Tabel 3. *Accuracy, ROC AUC, CV K-FOLD 10 SVM*

Kategori	Accuracy	ROC AUC	CV K-FOLD 10
0.0	85	96	90
0.1			

Tabel 4. *Accuracy, ROC AUC, CV K-FOLD 10 RFC*

Kategori	Accuracy	ROC AUC	CV K-FOLD 10
0.0	80	83	70
0.1			

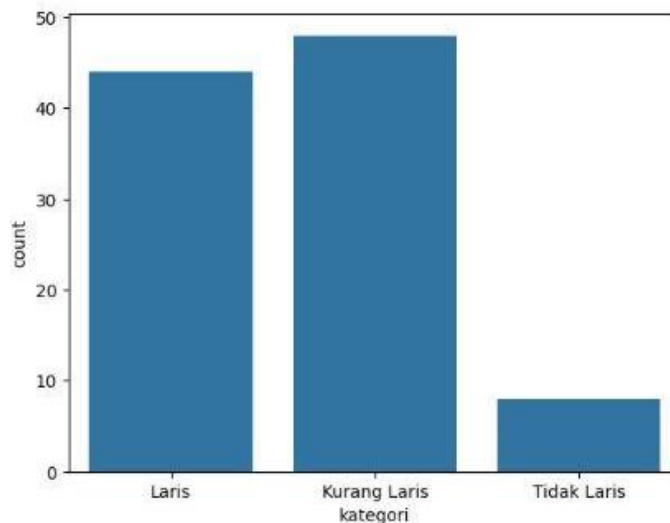
A. Visualisasi

Exploratory Data Analysis (EDA) merupakan suatu proses penting dalam tahapan machine learning yang digunakan untuk memahami struktur, pola, dan karakteristik data sebelum dilakukan pemodelan lebih lanjut. EDA berfokus pada eksplorasi awal dataset melalui penggunaan teknik statistik deskriptif dan visualisasi data, dengan tujuan untuk mengidentifikasi anomali, menemukan hubungan antar variabel, serta mendeteksi pola yang tersembunyi dalam data. Proses ini berperan penting dalam

memastikan kualitas data dan memberikan wawasan awal yang diperlukan untuk memilih metode analisis yang tepat.

Dalam klasifikasi Alat Tulis Kantor (ATK), penerapan EDA dilakukan untuk mengidentifikasi variabel-variabel utama yang mempengaruhi tingkat permintaan suatu produk. Melalui analisis eksploratori, peneliti dapat menemukan faktor-faktor penting seperti waktu pembelian (misalnya bulan atau hari tertentu dengan penjualan tertinggi), jenis produk yang paling sering dibeli oleh konsumen, serta pola pembelian musiman atau berkala. Informasi ini tidak hanya membantu dalam memahami perilaku konsumen, tetapi juga menjadi dasar dalam membangun model klasifikasi yang lebih akurat dan relevan.

Selain itu, EDA juga dapat mengungkap distribusi data penjualan, korelasi antar variabel, dan outlier yang mungkin mempengaruhi hasil analisis. Dengan memahami pola ini, pemilik usaha, seperti Toko Fotocopy Martin, dapat mengambil keputusan strategis terkait pengelolaan stok, perencanaan pengadaan, hingga penentuan strategi promosi, sehingga pengelolaan inventaris dapat dilakukan secara lebih efisien dan tepat sasaran.[18] Berikut visualisasi dalam proses Exploratory Data Analysis berupa diagram batang yang menunjukkan jumlah (count) dari tiga kategori yang berbeda terkait penjualan atau popularitas produk ATK:



Gambar 2. Diagram Penjualan

1) Laris

Kategori Laris menunjukkan produk-produk dengan tingkat penjualan yang tinggi, dengan jumlah penjualan mencapai sekitar 50 unit. Jumlah ini mencerminkan bahwa produk atau layanan dalam kategori ini memiliki permintaan yang sangat besar di pasar, sehingga termasuk ke dalam produk yang sangat diminati oleh konsumen. Produk-produk dalam kategori ini berpotensi untuk dijadikan prioritas dalam strategi pengadaan dan penataan stok, karena tingkat perputaran barang yang cepat dan kontribusi signifikan terhadap pendapatan toko.

2) Kurang Laris

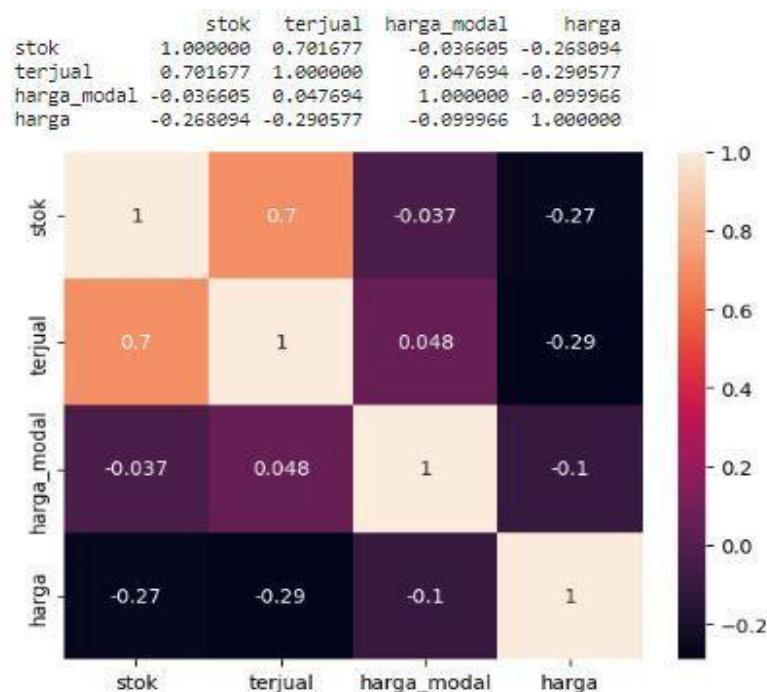
Kategori Kurang Laris mencakup produk-produk dengan jumlah penjualan yang lebih rendah dibandingkan kategori Laris, namun masih tergolong cukup signifikan. Produk dalam kategori ini menunjukkan adanya permintaan dari konsumen, meskipun tidak sebesar produk dalam kategori Laris. Hal ini mengindikasikan bahwa produk masih memiliki pasar, tetapi memerlukan strategi pemasaran atau promosi tambahan untuk meningkatkan tingkat

penjualannya. Produk-produk ini tetap perlu dikelola dalam stok dengan jumlah moderat agar tidak terjadi kelebihan atau kekurangan persediaan.

3) Tidak Laris

Kategori Tidak Laris berisi produk-produk dengan jumlah penjualan yang sangat rendah, menunjukkan bahwa produk atau layanan tersebut kurang diminati atau bahkan tidak populer di pasar. Jumlah penjualan yang minim mengindikasikan bahwa permintaan konsumen terhadap produk ini sangat terbatas, sehingga perlu dilakukan evaluasi terkait alasan rendahnya penjualan. Produk dalam kategori ini berisiko menjadi barang tidak terjual (dead stock) jika tetap dipertahankan dalam inventori, sehingga diperlukan keputusan strategis, seperti pengurangan stok, promosi khusus, atau bahkan penghapusan dari daftar produk untuk menghindari kerugian lebih lanjut.

Berikut matrik korelasi yang dapat memberikan informasi mengenai hubungan antar variabel:



Gambar 3. Matrik Korelasi

Berdasarkan gambar di atas, terdapat beberapa informasi penting yang dapat diinterpretasikan dari hasil analisis korelasi antar variabel. Pertama, nilai korelasi antara Stok dan Terjual sebesar 0.701 menunjukkan adanya hubungan positif yang kuat. Hal ini mengindikasikan bahwa semakin banyak stok yang tersedia, maka semakin besar pula jumlah produk yang terjual. Dengan hal ini, ketersediaan stok yang memadai berperan penting dalam memenuhi permintaan pasar dan mendorong peningkatan penjualan.

Kedua, nilai korelasi antara Stok dan Harga sebesar -0.267 menunjukkan adanya hubungan negatif yang lemah. Artinya, peningkatan jumlah stok tidak selalu diikuti dengan penurunan atau kenaikan harga yang signifikan. Hubungan negatif ini dapat diartikan bahwa harga produk relatif tidak terlalu dipengaruhi oleh jumlah stok yang tersedia, atau terdapat faktor lain yang lebih dominan dalam menentukan harga, seperti kualitas produk, merek, atau strategi pemasaran.

Ketiga, nilai korelasi antara Terjual dan Harga Modal sebesar -0.036 menunjukkan adanya hubungan yang sangat lemah atau hampir tidak ada hubungan antara jumlah produk yang terjual dengan harga modal. Hal ini mengindikasikan bahwa besar kecilnya harga modal yang dikeluarkan tidak memiliki pengaruh yang signifikan terhadap jumlah produk yang terjual. Dengan kata lain, faktor harga modal tidak menjadi penentu utama dalam tingkat penjualan, sehingga perlu dipertimbangkan variabel lain seperti harga jual, promosi, atau preferensi konsumen dalam menganalisis pola penjualan. Informasi ini penting sebagai dasar dalam pengambilan keputusan strategis terkait manajemen inventori dan penentuan harga, sehingga Toko Fotocopy Martin dapat mengoptimalkan pengelolaan stok dan meningkatkan efektivitas penjualan.

B. Klasifikasi

Klasifikasi data berbasis Machine Learning (ML) memberikan kontribusi signifikan dalam mendukung pengelolaan stok yang lebih optimal, terutama bagi usaha yang bergerak di bidang penjualan produk seperti Alat Tulis Kantor (ATK). Dengan memanfaatkan teknik klasifikasi, pemilik usaha dapat mengurangi potensi kesalahan dalam estimasi kebutuhan stok, sehingga dapat menghindari risiko kelebihan atau kekurangan persediaan. Selain itu, penerapan klasifikasi juga berperan dalam meningkatkan efisiensi operasional, karena proses pengambilan keputusan terkait pengadaan barang, penataan stok, hingga strategi promosi dapat dilakukan secara lebih terstruktur dan berbasis data [19].

Dalam penelitian ini, proses klasifikasi dilakukan dengan menerapkan dua algoritma machine learning, yaitu Random Forest Classifier dan Support Vector Machine (SVM). Kedua algoritma ini dipilih karena memiliki performa yang baik dalam menangani masalah klasifikasi dengan data yang bersifat kompleks dan multidimensional. Untuk memastikan keakuratan dan keandalan model, data yang digunakan dalam penelitian ini dibagi menjadi dua bagian, yaitu data latih dan data uji, dengan menggunakan teknik Cross-Validation k-fold. Secara khusus, Cross-Validation k-fold 10 diterapkan, di mana dataset dibagi menjadi 10 subset (fold) yang berukuran sama.

Dalam setiap iterasi, satu subset digunakan sebagai data uji, sementara sembilan subset lainnya digunakan sebagai data latih. Proses ini dilakukan secara berulang sebanyak 10 kali, sehingga setiap subset memiliki kesempatan yang sama untuk menjadi data uji. Hasil dari setiap iterasi kemudian dirata-ratakan untuk memperoleh evaluasi model yang lebih akurat, stabil, dan konsisten, serta meminimalkan kemungkinan bias akibat pemisahan data yang bersifat acak. Penerapan metode ini diharapkan dapat memberikan gambaran yang lebih komprehensif mengenai kinerja kedua algoritma dalam mengklasifikasikan data penjualan ATK, sekaligus mendukung pengambilan keputusan berbasis data dalam pengelolaan inventori di Toko Fotocopy Martin.

C. Evaluasi

Evaluasi hasil klasifikasi dalam penelitian ini mencakup empat metrik utama, yaitu Accuracy, Precision, Recall, dan F1 Score, yang digunakan untuk menilai kinerja algoritma Random Forest Classifier dan Support Vector Machine (SVM). Accuracy mengukur seberapa tepat model dalam memprediksi data secara keseluruhan, sementara Precision menunjukkan sejauh mana prediksi positif yang dilakukan oleh model benar adanya. Recall mengukur kemampuan model dalam mengidentifikasi semua data positif yang sebenarnya ada, dan F1 Score memberikan gambaran keseimbangan antara precision dan recall dengan menghitung rata-rata harmonik keduanya. Keempat metrik ini saling melengkapi dan memberikan gambaran yang lebih holistik mengenai kemampuan kedua algoritma dalam melakukan klasifikasi, yang pada akhirnya mendukung pengambilan keputusan dalam manajemen stok ATK di Toko Fotocopy Martin, dapat dilihat pada gambar dibawah ini:

Classification Report for Random Forest:				
	precision	recall	f1-score	support
0.0	0.73	1.00	0.85	11
1.0	1.00	0.56	0.71	9
accuracy			0.80	20
macro avg	0.87	0.78	0.78	20
weighted avg	0.85	0.80	0.79	20

Classification Report for SVM:				
	precision	recall	f1-score	support
0.0	0.79	1.00	0.88	11
1.0	1.00	0.67	0.80	9
accuracy			0.85	20
macro avg	0.89	0.83	0.84	20
weighted avg	0.88	0.85	0.84	20

Gambar 4. Hasil Evaluasi

Laporan klasifikasi untuk dua model machine learning, Random Forest Classifier dan Support Vector Machine (SVM), menunjukkan hasil yang berbeda dalam hal Precision dan Recall. Precision mengukur seberapa banyak prediksi positif yang benar dari semua prediksi yang dikategorikan sebagai positif. Untuk Random Forest Classifier, precision tercatat sebesar 0.73, yang berarti 73% dari semua yang diprediksi sebagai positif benar-benar positif. Sementara itu, untuk SVM, precision mencapai 0.79, yang berarti 79% dari semua prediksi positif adalah benar. Dalam hal Recall, yang mengukur seberapa baik model menemukan semua contoh positif, Random Forest Classifier memiliki recall sebesar 0.56, yang menunjukkan bahwa hanya 56% dari semua contoh positif yang berhasil ditemukan oleh model. Sementara untuk SVM, recall tercatat 0.67, yang berarti model ini dapat menemukan 67% dari contoh positif yang ada. Dengan demikian, SVM menunjukkan kinerja yang lebih baik dalam hal recall, sementara Random Forest Classifier unggul dalam hal precision.

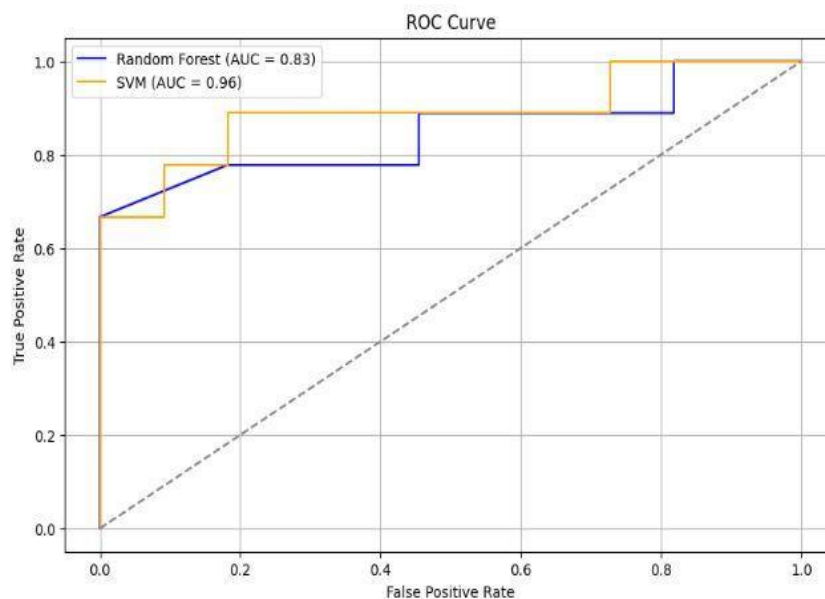
F1-Score memberikan gambaran yang lebih seimbang tentang kinerja model dengan menggabungkan precision dan recall dalam satu metrik. F1-Score sangat berguna ketika ada ketidakseimbangan antara precision dan recall, karena memberikan bobot yang lebih adil untuk keduanya. Dalam hal ini, Support menunjukkan distribusi kelas dalam data pengujian, di mana ada 11 kelas positif dan 9 kelas negatif pada kedua model Random Forest Classifier dan SVM. Akurasi mengukur total prediksi yang benar dari keseluruhan prediksi yang dibuat oleh model. Untuk Random Forest Classifier, akurasi mencapai 0.80, yang berarti 80% dari semua prediksi yang dibuat oleh model adalah benar. Sementara itu, untuk SVM, akurasi mencapai 0.85, yang menunjukkan bahwa 85% dari prediksi yang dibuat model adalah benar. Meskipun SVM memiliki akurasi yang lebih tinggi, evaluasi lebih mendalam dengan metrik seperti precision, recall, dan F1-Score akan memberikan gambaran yang lebih komprehensif tentang kekuatan dan kelemahan masing-masing model dalam tugas klasifikasi ini.

D. Validasi

Validasi dengan ROC Curve (Receiver Operating Characteristic Curve) digunakan untuk mengevaluasi kinerja model klasifikasi, seperti Random Forest Classifier dan SVM (Support Vector Machine), dengan cara yang lebih mendalam dan visual. Pada ROC Curve, sumbu X menggambarkan False Positive Rate (FPR), yang menunjukkan proporsi prediksi yang salah dikategorikan sebagai positif. Semakin rendah nilai FPR, semakin baik model dalam menghindari kesalahan dalam mengklasifikasikan

contoh negatif sebagai positif. Sementara itu, sumbu Y menggambarkan True Positive Rate (TPR), atau sering disebut juga sebagai Recall, yang menunjukkan proporsi prediksi yang benar dikategorikan sebagai positif. Semakin tinggi nilai TPR, semakin baik model dalam mengidentifikasi data positif dengan akurat.

Pada ROC Curve, terdapat garis putus-putus yang bertindak sebagai garis referensi, yang menunjukkan kinerja acak, dengan nilai 0.5. Garis ini menunjukkan kinerja model yang tidak lebih baik dari prediksi acak. Oleh karena itu, semakin jauh garis ROC dari garis putus-putus ini, semakin baik kinerja model dalam membedakan antara dua kelas, yaitu kelas positif dan kelas negatif. Pada model Random Forest Classifier, garis berwarna kuning mewakili AUC (Area Under the Curve) yang bernilai 0.83. Nilai AUC ini menunjukkan bahwa model Random Forest cukup baik dalam membedakan antara dua kelas, yaitu ATK Laris dan ATK Tidak Laris, dengan kemampuan untuk mengklasifikasikan dengan benar sebagian besar contoh positif dan negatif. Meskipun demikian, AUC sebesar 0.83 juga mengindikasikan bahwa masih ada ruang untuk perbaikan. Model ini tidak sempurna dan dapat lebih ditingkatkan, misalnya dengan penyempurnaan dalam pemilihan fitur atau pemrosesan data lebih lanjut, untuk lebih mengoptimalkan kemampuan dalam membedakan kedua kelas tersebut. Secara keseluruhan, ROC Curve dan AUC memberikan wawasan penting mengenai trade-off antara True Positive Rate dan False Positive Rate serta membantu dalam memilih model terbaik berdasarkan seberapa efektif model dapat memisahkan kelas-kelas yang berbeda. Berikut gambar ROC AUC:



Gambar 5. ROC AUC

Dari hasil evaluasi, SVM menunjukkan performa lebih unggul dibandingkan Random Forest dalam hal akurasi dan F1-Score. Faktor dominan yang memengaruhi keunggulan SVM adalah kemampuannya membentuk hyperplane optimal yang memisahkan kelas dengan margin maksimum. Hal ini sangat efektif ketika data memiliki distribusi non-linear yang dapat ditangani melalui kernel trick. Selain itu, SVM relatif lebih stabil terhadap overfitting pada dataset dengan jumlah fitur yang cukup besar tetapi observasi yang terbatas, seperti kasus ATK ini.

4. KESIMPULAN

Berdasarkan hasil penelitian yang dilakukan pada manajemen inventori stok alat tulis kantor (ATK) di Toko Fotocopy Martin, dapat disimpulkan bahwa penerapan metode analisis dan klasifikasi produk dengan menggunakan algoritma machine learning, khususnya Random Forest Classifier dan Support Vector Machine, memberikan hasil yang signifikan dalam mengelola stok.

Namun, penelitian ini memiliki keterbatasan, yaitu evaluasi hanya dilakukan pada data historis (simulasi) tanpa melibatkan pengguna nyata seperti manajer toko atau uji coba dalam operasional langsung. Untuk pengembangan ke depan, perlu dilakukan validasi lapangan agar dapat memastikan kepraktisan dan keandalan sistem dalam skenario nyata. Setelah melakukan pengumpulan dan preprocessing data, sebanyak 100 produk ATK berhasil dikategorikan, dengan hasil labeling menunjukkan 37 produk tergolong Laris, 36 produk Kurang Laris, dan 21 produk Tidak Laris. Penggunaan teknik Exploratory Data Analysis (EDA) membantu dalam memahami karakteristik data sebelum diterapkan algoritma klasifikasi.

Dari hasil klasifikasi, algoritma Random Forest Classifier menunjukkan performa dengan nilai akurasi 80% dan F1 Score sebesar 85% untuk kategori Laris. Sementara itu, Support Vector Machine menunjukkan akurasi yang lebih baik dengan nilai 85% dan F1 Score 88% untuk kategori yang sama. Secara keseluruhan, kedua algoritma menunjukkan hasil yang memuaskan, dengan SVM memberikan kinerja yang sedikit lebih baik dalam hal akurasi dan nilai F1 Score. Hal ini mengindikasikan bahwa penerapan machine learning dan EDA dapat menjadi solusi yang efektif untuk meningkatkan manajemen inventori di Toko Fotocopy Martin, dengan demikian mengurangi risiko kelebihan atau kekurangan stok dan meningkatkan efisiensi operasional.

UCAPAN TERIMA KASIH

Penulis ingin mengungkapkan rasa terima kasih yang tulus kepada Universitas Bina Insan atas fasilitas dan bantuan yang telah diberikan dalam menyelesaikan penelitian ini. Dukungan dan bimbingan dari para dosen dan rekan-rekan peneliti sangat berharga dalam proses penelitian ini. Semoga karya ilmiah ini dapat memberikan manfaat dan kontribusi positif dalam pengembangan ilmu pengetahuan. Terima kasih atas segala support dan dorongan yang diberikan.

REFERENSI

- [1] P. Mesin and T. Komprehensif, “Kemajuan dan Tantangan dalam Pembelajaran Mesin: Tinjauan Komprehensif tentang Model, Pustaka, Aplikasi, dan Algoritma,” 2023.
- [2] A. Alamanda and W. Azlia, “KLASIFIKASI ABC,” vol. 02, no. 06, pp. 610–618, 2024.
- [3] M. Basket, A. Tren, and H. Menggunakan, “INTI NUSA MANDIRI MARKET BASKET ANALYSIS TREND HIJAB MENGGUNAKAN ALGORITMA,” vol. 16, no. 2, pp. 1–6, 2022.
- [4] Y. Septiani, E. Arribe, and R. Diansyah, “(Studi Kasus : Mahasiswa Universitas Abdurrah Pekanbaru),” vol. 3, no. 1, pp. 131–143, 2020.
- [5] F. A. Pambudi, A. P. Windarto, M. Fauzan, J. T. Hardinata, and R. Winanjaya, “Analisis Klasifikasi C4 . 5 Pada Pola Pembayaran Sepeda Motor Adira Cabang Pematangsiantar,” vol. 1, no. 4, pp. 160–169, 2021.
- [6] A. H. Wibowo *et al.*, “Analisa Dan Visualisasi Data Penjualan Menggunakan Exploratory Data Analysis Pada PT. Telkominfra,” vol. 9, no. 3, pp. 2292–2304, 2022.

-
- [7] A. S. Millah, D. Arobiah, E. S. Febriani, and E. Ramdhani, "Analisis Data dalam Penelitian Tindakan Kelas," vol. 1, no. 2, pp. 140–153, 2023.
- [8] A. Hermawan, B. Wicaksono, T. Ahmadjayadi, and B. S. Prakasa, "Implementasi Algoritma Apriori pada Market Basket Analysis terhadap Data Penjualan Produk Supermarket," vol. 2, no. 5, 2024.
- [9] R. Artikel, A. Widjaja, and O. Karnalim, "Analisis Perbandingan Algoritma Machine Learning Untuk Forecasting Persediaan Produk Barang Pokok Comparative Analysis Of Machine Learning Algorithms For Forecasting Of Basic Goods Products," vol. 10, pp. 361–378, 2024.
- [10] M. H. Tinambunan, A. Hasibuan, S. Wahyuni, and A. S. Wibowo, "Jurnal Bisnis Net Volume : 6 No . 1 Juni 2023 | ISSN : 2621 -3982 KLASIFIKASI TINGKAT KEPUASAN MAHASISWA TERHADAP FASILITAS PADA FTIK UNIVERSITAS DHARMAWANGSA Jurnal Bisnis Net Volume : 6 No . 1 Juni 2023 | ISSN : 2621 -3982 EISSN : 2722- 3574," no. 1, pp. 208–215, 2023.
- [11] P. Issn, "Jurnal Ekonomi dan Bisnis , Vol . 10 No . 2 Mei 2022 E - ISSN," vol. 10, no. 2, pp. 313–318, 2022.
- [12] P. Studi, S. Informasi, F. Teknik, D. A. N. Komputer, and U. P. Batam, "Analisis pola pembelian konsumen dengan metode asosiasi menggunakan algoritma apriori pada minimarket," 2021.
- [13] L. Nirawati and T. M. Seibinna, "Optimalisasi Proses Stock Opname Alat Tulis Kantor (ATK) Melalui Digitalisasi Quick Response Code (Qr Code) Pada PT Surabaya Industrial Estate Rungkut (SIER)," vol. 1, pp. 475–479, 2024.
- [14] H. Nurfauziah and A. A. Chaniago, "SISTEM INFORMASI PENGAJUAN KEBUTUHAN ALAT TULIS KANTOR (ATK) BERBASIS WEBSITE DI CV . BHINEKA JAYA BOGA BEKASI," vol. 10, no. 1, pp. 15–29, 2024.
- [15] A. Lewis, M. Zarlis, and Z. Situmorang, "Penerapan Data Mining Menggunakan Task Market Basket Analysis Pada Transaksi Penjualan Barang di AB Mart dengan Algoritma," vol. 5, no. April, pp. 676–681, 2021, doi: 10.30865/mib.v5i2.2934.
- [16] Ö. Albayrak Ünal, B. Erkeyman, and B. Usanmaz, "Applications of Artificial Intelligence in Inventory Management: A Systematic Review of the Literature," *Arch. Comput. Methods Eng.*, vol. 30, no. 4, pp. 2605–2625, 2023, doi: 10.1007/s11831-022-09879-5.
- [17] A. Wasik *et al.*, "Implementasi data mining untuk memprediksi penjualan accessoris handphone dan handphone terlaris menggunakan metode k-nearest neighbor (k-nn) 1," vol. 1, no. 2, pp. 469–479, 2024.
- [18] M. Putri and A. Vaoutama, "Analisis Dan Visualisasi Data Untuk Meningkatkan Penjualan Menggunakan Exploratory Data Analysis Dan Looker Studio (Studi Kasus : Nies Collection)," vol. 12, no. 4, pp. 899–914, 2024.
- [19] M. Mittal, V. Jain, J. T. Pandey, M. Jain, and H. Dem, "Optimizing Inventory Management : A Comprehensive Analysis of Models Integrating Diverse Fuzzy Demand Functions," pp. 1–18, 2024.
- [20] N. Utami, K. A. Baihaqi, E. E. Awal, and D. Wahiddin, "Analisis Kinerja Algoritma Decision Tree Dan Random Forest Dalam Klasifikasi Penyakit Kardiovaskular," vol. 6, no. 2, pp. 970–980, 2024, doi: 10.47065/bits.v6i2.5722.