

Prediksi dan Klasifikasi Transaksi Penjualan Terbaik Dalam Toko Bangunan Dengan Metode *K-Nearest Neighbors (K-NN)*

¹Adrian Chandra Kusumah, ²Nandi Adi Nugroho, ³Genta Aldora Leopriandis, ⁴Achmad Khautsar Rizaldi, ⁵Firmansyah, ⁶Maulana Fansyuri.

¹Sistem Informasi, Fakultas Ilmu Komputer, Universitas Pamulang, Kota Tangerang Selatan, Indonesia.

²Sistem Informasi, Fakultas Ilmu Komputer, Universitas Pamulang, Kota Tangerang Selatan, Indonesia.

³Sistem Informasi, Fakultas Ilmu Komputer, Universitas Pamulang, Kota Tangerang Selatan, Indonesia.

⁴Sistem Informasi, Fakultas Ilmu Komputer, Universitas Pamulang, Kota Tangerang Selatan, Indonesia.

⁵Sistem Informasi, Fakultas Ilmu Komputer, Universitas Pamulang, Kota Tangerang Selatan, Indonesia.

⁶Sistem Informasi, Fakultas Ilmu Komputer, Universitas Pamulang, Kota Tangerang Selatan, Indonesia.

¹adrianchandrakusumah@gmail.com, ²nandi.adi.nugroho@gmail.com, ³leopriandi@gmail.com,
⁴Acongkeren50@gmail.com, ⁵frmnsyah1500@gmail.com, ⁶dosen02359@unpam.ac.id.

ABSTRACT

This study uses the K-Nearest Neighbors (K-NN) method to predict and classify the best-selling products in a hardware store. With the current development of information technology, sales trend analysis and prediction have become an important part of the business decision-making process. The popular K-NN classification algorithm is used to analyze sales data from a public dataset to determine which products are most in demand by consumers. The process of data collection, selection, preprocessing, transformation, data mining, and evaluation of results are all part of the Knowledge Discovery in Database (KDD) stages. The analysis results show that products in the "active" category sell more than products in the "passive" category. Out of the total data, 56 were successfully categorized as active data, and the remaining 29 were categorized as passive data. This study is expected to provide deeper insights into consumer behavior and assist building material store management in making better decisions using the data they possess. This is anticipated to enhance the company's competitiveness and improve operational efficiency.

Keywords: *K-Nearest Neighbors (K-NN), Data Mining, KDD, Sales Prediction, Product Classification, Building Store.*

ABSTRAK

Studi ini menggunakan metode *K-Nearest Neighbors (K-NN)* untuk memprediksi dan mengklasifikasikan transaksi penjualan produk terbaik di toko bangunan. Dengan perkembangan teknologi informasi saat ini, analisis dan prediksi tren penjualan menjadi bagian penting dari proses pengambilan keputusan bisnis. Algoritma klasifikasi populer K-NN digunakan untuk menganalisis data penjualan dari *dataset* publik untuk menentukan produk mana yang paling diminati oleh konsumen. Proses pengumpulan data, pemilihan (*selection*), prapemrosesan (*preprocessing*), transformasi (*transformation*), penambangan (*data mining*), dan evaluasi hasil adalah semua bagian dari tahapan *Knowledge Discovery in Database (KDD)*. Hasil analisis menunjukkan bahwa produk dalam kategori "aktif" menjual lebih banyak daripada produk dalam kategori "pasif". Dari total data, 56 berhasil dikategorikan sebagai data aktif, dan 29 data lainnya dikategorikan sebagai data pasif. Penelitian ini diharapkan dapat memberikan wawasan yang lebih mendalam tentang perilaku konsumen dan membantu manajemen toko bangunan membuat keputusan yang lebih baik dengan data yang mereka miliki. Hal ini diharapkan akan meningkatkan daya saing perusahaan dan meningkatkan efisiensi operasional.

Kata Kunci: *K-Nearest Neighbors (K-NN), Data Mining, KDD, Prediksi Penjualan, Klasifikasi Produk, Toko Bangunan.*

1. PENDAHULUAN

Banyak industri telah beralih ke sistem digital untuk mengelola data karena kemajuan teknologi informasi, termasuk industri ritel seperti toko bangunan. Kemampuan untuk menganalisis dan memprediksi tren penjualan produk menjadi penting untuk pengambilan

keputusan bisnis di tengah persaingan yang semakin ketat. Toko bangunan memiliki berbagai jenis produk dengan tingkat permintaan yang beragam, sehingga penting bagi pelaku usaha untuk mengetahui produk apa yang memiliki tingkat penjualan tertinggi dan berpotensi menjadi produk unggulan [1].

Dalam *data mining*, metode klasifikasi adalah salah satu metode yang dapat digunakan untuk menganalisis data penjualan. Metode *K-Nearest Neighbors* (K-NN), salah satu algoritma klasifikasi yang paling populer dan efektif, dapat membuat prediksi berdasarkan kemiripan data historis [2]. Toko bangunan menghadapi tantangan dalam menentukan produk mana yang masuk dalam kategori *best-seller* karena banyaknya produk dan perubahan permintaan. Solusi strategis untuk mendukung pengambilan keputusan berbasis data adalah teknologi prediksi dan klasifikasi. Toko bangunan dapat menggunakan algoritma ini untuk mengklasifikasikan produk berdasarkan pola transaksi penjualan sebelumnya. Ini memungkinkan untuk membuat prediksi produk terbaik untuk diterapkan dalam strategi penjualan mereka di masa mendatang [3].

K-Nearest Neighbors (K-NN) adalah salah satu algoritma *data mining* yang terbukti berhasil dalam proses prediksi dan klasifikasi. Algoritma ini bekerja dengan menggunakan fitur tertentu, seperti harga, jenis produk, waktu transaksi, dan jumlah penjualan, untuk membandingkan kemiripan antara data baru dan data sebelumnya. Meskipun metode K-NN tidak parametrik dan sederhana, tetapi sangat akurat pada kumpulan data yang bersih dan terstruktur. K-NN dapat mengklasifikasikan data penjualan dengan akurasi di atas 85%, menurut beberapa penelitian sebelumnya.

Studi oleh (Guru dkk., 2023) [4] supermarket besar yang memiliki catatan penjualan yang rumit. Akurasi yang lebih tinggi dalam menangkap pola *non-linear* dan interaksi antar fitur, seperti kategori produk, lokasi *outlet*, dan visibilitas item, dapat dicapai dengan K-NN tanpa membutuhkan asumsi bentuk hubungan matematis tertentu. Sementara itu, (Safitri dan Fakhri, 2024) [5] menunjukkan bahwa metode K-NN dapat mengurangi kemungkinan kekurangan stok semen dan pasir dengan membuat prediksi berdasarkan data transaksi sebelumnya. Selain itu, apabila dibandingkan dengan algoritma *Naive Bayes* di bidang ritel lainnya, K-NN menunjukkan kinerja klasifikasi yang lebih baik dalam menentukan *produk best-seller* [6].

Metode K-NN untuk analisis penjualan di toko bangunan masih sangat terbatas, meskipun telah banyak digunakan. Tanpa menggabungkan klasifikasi produk secara bersamaan, beberapa studi lebih fokus pada prediksi jumlah penjualan [7]. Selain itu, variabel input yang digunakan masih sangat terbatas dan tidak mempertimbangkan hal-hal seperti waktu promosi, kategori produk, atau musim. Untuk memberikan hasil yang lebih relevan bagi pemilik bisnis, penelitian yang tidak hanya memprediksi volume penjualan tetapi juga mengklasifikasikan produk terbaik berdasarkan pola transaksi aktual diperlukan [8].

Analisis data penjualan dapat dilakukan menggunakan pendekatan *Knowledge Discovery in Database* (KDD), yang terdiri dari beberapa tahapan seperti *data selection*, *data preprocessing*, *data transformation*, *data mining*, dan *evaluation*. KDD

memastikan informasi data dapat diandalkan untuk pengambilan keputusan. Fokus utama penelitian adalah tahap *data mining*, di mana algoritma *K-Nearest Neighbor* (K-NN) digunakan untuk proses klasifikasi dan prediksi [9].

Penelitian ini bertujuan untuk menerapkan metode K-NN dalam memprediksi dan mengklasifikasikan transaksi penjualan produk terbaik pada toko bangunan. Hasil dari penelitian ini diharapkan dapat memberikan gambaran yang lebih jelas tentang perilaku konsumen serta membantu pihak manajemen dalam pengambilan keputusan berbasis data, guna meningkatkan efisiensi operasional dan daya saing bisnis.

2. PELAKSAAAN DAN METODE

2.1. Pengumpulan Data

Pada penelitian ini, digunakan data sekunder yang berasal dari *dataset* produk toko bangunan yang tersedia di *website* publik, yang melalui proses kurasi dan validasi awal. <https://www.kaggle.com/datasets/revgaming/data-transaksi-toko-bangunan>. Diakses pada hari Rabu tanggal 20 Juni 2025, pukul 10.09 WIB. Kaggle adalah *platform online* yang memungkinkan *data scientist* dan praktisi *machine learning* untuk belajar, berbagi, dan bekerja sama. Di *platform* ini, pengguna dapat mengakses data, membuat *notebook* (kode), mengikuti pelatihan gratis, dan membuat portofolio analisis berbasis data nyata.

2.2. Tahapan Penelitian

Pada tahapan penelitian ini menggunakan metode *Knowledge Discovery in Database* (KDD) yaitu proses menemukan pola atau pengetahuan bermanfaat dari kumpulan data besar melalui proses seperti *data selection*, *data preprocessing*, *data transformation*, *data mining*, dan *evaluation*. Tujuan KDD adalah untuk mengubah data mentah menjadi informasi yang berguna untuk pengambilan keputusan [10].



Gambar 1. Tahapan *Knowledge Discovery in Database* (KDD)

2.3. Data Selection

Data selection adalah tahap pertama proses KDD. Ini adalah proses memilih data mentah dari berbagai sumber untuk dianalisis lebih lanjut. Hanya data yang relevan dan terkait langsung dengan tujuan analisis yang diambil yang diperlukan untuk proses analisis [11]. Misalnya, untuk melakukan penelitian tentang cara terbaik untuk memprediksi penjualan produk di toko bangunan, data yang dipilih termasuk elemen penting seperti nama produk, jumlah penjualan, harga, tanggal transaksi, dan kategori produk. Untuk meningkatkan efisiensi proses selanjutnya, tujuan tahap ini adalah untuk menghasilkan kumpulan data yang terfokus dan berkualitas tinggi. Penggabungan data dari berbagai sumber atau pengambilan data historis dengan jangkauan waktu yang cukup untuk dianalisis juga sering merupakan bagian dari proses pemilihan data. Keberhasilan proses KDD secara keseluruhan sangat dipengaruhi oleh pemilihan data yang tepat [12].

2.4. Data Preprocessing

Data preprocessing adalah tahap berikutnya setelah data dipilih; ini adalah proses pembersihan data agar bebas dari kesalahan, inkonsistensi, duplikasi, dan nilai yang hilang. Proses ini sangat penting untuk menjamin kualitas data yang akan dihasilkan karena data yang tidak bersih dapat memengaruhi hasil analisis. Pada langkah ini, berbagai tindakan dilakukan, termasuk mengisi atau menghapus nilai kosong, memperbaiki format data yang tidak konsisten, seperti perbedaan format tanggal, menghapus duplikasi, dan menemukan dan mengatasi data *outlier* atau anomali. Pada tahap ini, teknik imputasi, validasi data, dan penggunaan fungsi logika atau statistik sederhana sering digunakan. Selain membersihkan data, *preprocessing* memastikan bahwa data dalam kondisi stabil dan siap untuk tahap transformasi berikutnya [13].

2.5. Data Transformation

Proses mengubah struktur atau format data agar lebih sesuai dan optimal untuk dianalisis dengan teknik *data mining* dikenal sebagai *data transformation*. Sebagai contoh, nilai kategori seperti "tinggi" dapat dinormalisasi antara 0 dan 1, nilai kategori seperti "sedang" dapat dikonversi menjadi 2, 1, dan 0 atau nilai kategori seperti "rendah" dapat dikonversi menjadi 2, 1, dan 0. Karena algoritma *data mining* seperti *K-Nearest Neighbors* sangat bergantung pada struktur dan format data yang seragam, transformasi sangat penting. Algoritma dapat mengenali pola atau jarak antar data dengan lebih baik dengan data yang telah diubah. Akibatnya, langkah ini sangat penting untuk mendukung akurasi hasil analisis [14].

2.6. Data mining

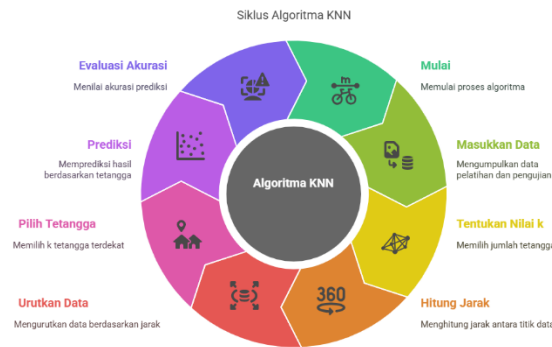
Tahap penting dari proses KDD adalah *data mining*. Di sini, analisis data digunakan untuk menemukan pola tersembunyi, hubungan antar atribut, atau prediksi berdasarkan data historis. Menurut tujuan analisis, *data mining* menggunakan berbagai metode, termasuk klasifikasi, klastering, regresi, dan asosiasi [15]. Metode klasifikasi *K-Nearest Neighbors* (K-NN) digunakan dalam penelitian ini untuk mengelompokkan produk ke dalam kategori tertentu (seperti produk laris atau tidak laris) berdasarkan kemiripan dengan data sebelumnya. *Data mining* membutuhkan pemahaman yang baik tentang algoritma yang digunakan, pemahaman tentang parameter yang harus diubah, dan validasi hasil. Model atau pola yang dihasilkan dari proses ini dapat menjelaskan perilaku data dan memberikan pengetahuan baru yang sebelumnya tidak terlihat secara langsung [16].

2.7. Evaluation

Evaluation adalah langkah terakhir dalam proses KDD. Ini adalah proses menilai kualitas, validitas, dan kegunaan pola atau model yang dibuat pada tahap *data mining*. Bergantung pada jenis analisisnya, *evaluation* ini dilakukan dengan menggunakan berbagai metrik. Untuk klasifikasi, digunakan metrik seperti akurasi, ketepatan, *recall*, dan skor F1. Untuk asosiasi, digunakan nilai dukungan, keyakinan, dan lift. *Evaluation* tidak hanya mencakup penilaian statistik, tetapi juga memahami bagaimana pola tersebut berdampak pada penelitian atau bisnis. Apakah pola yang ditemukan dapat diterima? Apakah ini bermanfaat untuk proses pengambilan keputusan? Untuk mengetahui apakah model konsisten dengan data uji, *evaluation* juga dapat melibatkan *cross-validation*. Jika pola ditemukan tidak valid atau tidak bermakna pada tahap sebelumnya, seperti *data selection*, *data preprocessing*, *data transformation*, dan *data mining*, maka harus diubah [15].

2.8. K-Nearest Neighbor (K-NN)

Untuk tugas regresi dan klasifikasi, algoritma *K-Nearest Neighbors* (K-NN) adalah metode pembelajaran *supervised* non-parametrik. Prinsip kerjanya sederhana: algoritma ini mengidentifikasi k data tetangga (k neighbors) dalam ruang fitur berdasarkan metrik jarak, biasanya geometri *Euclidean*. Kemudian, untuk memprediksi label data baru, kelas output ditentukan dengan memilih label mayoritas (klasifikasi) atau rata-rata (regresi) [17].



Gambar 2. Tahapan Algoritma KNN

K-NN sangat disukai karena tidak membutuhkan proses pelatihan yang rumit hanya menyimpan data pelatihan dan sangat fleksibel dalam berbagai konteks, seperti *data mining*, analisis teks, dan citra medis. Meskipun sederhana, performanya dapat ditingkatkan melalui teknik modifikasi, seperti pembobotan jarak, pemilihan jumlah k yang ideal melalui *cross-validation*, dan penggunaan algoritma pencarian terdekat atau fitur yang lebih efisien [18].

2.9. RapidMiner

RapidMiner adalah program atau aplikasi *data mining* yang memungkinkan pengolahan kumpulan data yang sangat besar untuk menghasilkan informasi baru. Ini menawarkan berbagai metode untuk evaluasi statistik, mulai dari metode sederhana untuk klasifikasi dan pengelompokan hingga analisis korelasi dan regresi. Selain itu, ada pengurangan dimensi dan optimasi parameter [19].

RapidMiner menawarkan lingkungan *drag-and-drop visual* yang memudahkan pemula dan profesional membangun alur kerja analisis seperti pembersihan, transformasi, pelatihan model, evaluasi, dan visualisasi

hasil. Algoritma seperti *K-Nearest Neighbors* (K-NN), *Decision Tree*, *Naive Bayes*, dan lainnya dapat digunakan di *platform* ini. Karena memiliki pustaka algoritma yang lengkap dan kompatibel dengan berbagai format data seperti *CSV*, *Excel*, dan *SQL*, *RapidMiner* banyak digunakan dalam bidang akademik, bisnis, dan penelitian. Kemudahan penggunaan, fleksibilitas integrasi, dan kemampuan untuk melakukan pemrosesan data secara otomatis dan efisien adalah keunggulannya [20].

3. HASIL DAN PEMBAHASAN

3.1. Data Selection

Tahap pada penelitian ini, peneliti memilih data yang bersumber dari *website* www.kaggle.com dengan judul “Data Transaksi Toko Bangunan”. Pada penelitian ini kami menggunakan *dataset* yaitu 90 data detail dengan 5 total atribut 3 kategorial dan 2 numerik. Untuk lebih jelasnya dapat dilihat dari tabel dibawah ini.

Tabel 1. *Dataset*

ID Transaksi	Tanggal Pembelian	Produk	Kategori	Satuan	Harga Satuan	Total Harga	Kuantitas
T001	1/1/2024	Semen	Material Konstruksi	Zak	63000	315000	5
T001	1/1/2024	Pasir	Material Konstruksi	Meter Kubik	447000	2235000	5
T001	1/1/2024	Cat Tembok	Cat	Kaleng	232000	1856000	8
T002	1/2/2024	Semen	Material Konstruksi	Zak	63000	630000	10
T002	1/2/2024	Besi Beton	Bahan Logam dan PVC	Batang	101000	101000	1
T002	1/2/2024	Cat Kayu	Cat	Kaleng	89000	623000	7
T003	1/2/2024	Pasir	Material Konstruksi	Meter Kubik	447000	4023000	9
T003	1/2/2024	Batu Bata	Material Konstruksi	Buah	1000	9000	9
T003	1/2/2024	Semen	Material Konstruksi	Zak	63000	567000	9
T004	1/2/2024	Cat Tembok	Cat	Kaleng	232000	232000	1
T004	1/2/2024	Cat Kayu	Cat	Kaleng	89000	534000	6
T004	1/2/2024	Kuas	Alat	Buah	18000	90000	5
T005	1/2/2024	Semen	Material Konstruksi	Zak	63000	441000	7
T005	1/2/2024	Pasir	Material Konstruksi	Meter Kubik	447000	1341000	3
T005	1/2/2024	Batu Bata	Material Konstruksi	Buah	1000	2000	2
T006	1/2/2024	Besi Beton	Bahan Logam dan PVC	Batang	101000	606000	6
T006	1/2/2024	Pipa PVC	Bahan Logam dan PVC	Batang	50000	300000	6
T006	1/2/2024	Sambungan PVC	Bahan Logam dan PVC	Buah	11000	110000	10
T007	1/3/2024	Batu Bata	Material Konstruksi	Buah	1000	3000	3
T007	1/3/2024	Pasir	Material Konstruksi	Meter Kubik	447000	447000	1
T007	1/3/2024	Semen	Material Konstruksi	Zak	63000	504000	8
T008	1/3/2024	Cat Tembok	Cat	Kaleng	232000	928000	4
T008	1/3/2024	Semen	Material Konstruksi	Zak	63000	252000	4
T008	1/3/2024	Kuas	Alat	Buah	18000	108000	6
T009	1/3/2024	Besi Beton	Bahan Logam dan PVC	Batang	101000	1010000	10
T009	1/3/2024	Batu Bata	Material Konstruksi	Buah	1000	5000	5
T009	1/3/2024	Semen	Material Konstruksi	Zak	63000	189000	3
T010	1/4/2024	Cat Kayu	Cat	Kaleng	89000	890000	10
T010	1/4/2024	Cat Tembok	Cat	Kaleng	232000	1624000	7
T010	1/4/2024	Kuas	Alat	Buah	18000	54000	3
T011	1/4/2024	Pasir	Material Konstruksi	Meter Kubik	447000	3576000	8
T011	1/4/2024	Batu Bata	Material Konstruksi	Buah	1000	6000	6
T012	1/4/2024	Sambungan PVC	Bahan Logam dan PVC	Batang	11000	110000	10
T012	1/4/2024	Cat Kayu	Cat	Kaleng	89000	623000	7
T013	1/4/2024	Kuas	Alat	Buah	18000	72000	4
T013	1/4/2024	Batu Bata	Material Konstruksi	Buah	1000	6000	6
T013	1/4/2024	Cat Tembok	Cat	Kaleng	232000	1160000	5
T014	1/4/2024	Pasir	Material Konstruksi	Meter Kubik	447000	3576000	8
T014	1/4/2024	Cat Kayu	Cat	Kaleng	89000	801000	9
T014	1/4/2024	Pipa PVC	Bahan Logam dan PVC	Batang	50000	450000	9
T014	1/4/2024	Kuas	Alat	Buah	18000	72000	4
T015	1/4/2024	Cat Kayu	Cat	Kaleng	89000	178000	2
T015	1/4/2024	Pasir	Material Konstruksi	Meter Kubik	447000	4470000	10
T015	1/4/2024	Kuas	Alat	Buah	18000	180000	10
T016	1/5/2024	Sambungan PVC	Bahan Logam dan PVC	Batang	11000	99000	9
T016	1/5/2024	Semen	Material Konstruksi	Zak	63000	189000	3
T016	1/5/2024	Kuas	Alat	Buah	18000	126000	7
T016	1/5/2024	Besi Beton	Bahan Logam dan PVC	Batang	101000	101000	1
T017	1/6/2024	Pipa PVC	Bahan Logam dan PVC	Batang	50000	300000	6
T017	1/6/2024	Semen	Material Konstruksi	Zak	63000	252000	4
T018	1/7/2024	Pipa PVC	Bahan Logam dan PVC	Batang	50000	450000	9
T018	1/7/2024	Kuas	Alat	Buah	18000	144000	8
T018	1/7/2024	Cat Tembok	Cat	Kaleng	232000	696000	3
T018	1/7/2024	Cat Kayu	Cat	Kaleng	89000	356000	4
T019	1/7/2024	Besi Beton	Bahan Logam dan PVC	Batang	101000	606000	6
T019	1/7/2024	Pipa PVC	Bahan Logam dan PVC	Batang	50000	200000	4
T020	1/7/2024	Pipa PVC	Bahan Logam dan PVC	Batang	50000	350000	7
T020	1/7/2024	Cat Kayu	Cat	Kaleng	89000	178000	2
T020	1/7/2024	Cat Tembok	Cat	Kaleng	232000	232000	1
T020	1/7/2024	Kuas	Alat	Buah	18000	108000	6
T021	1/7/2024	Cat Tembok	Cat	Kaleng	232000	1856000	8
T021	1/7/2024	Batu Bata	Material Konstruksi	Buah	1000	8000	8
T022	1/7/2024	Sambungan PVC	Bahan Logam dan PVC	Batang	11000	99000	9
T022	1/7/2024	Besi Beton	Bahan Logam dan PVC	Batang	101000	808000	8
T022	1/7/2024	Kuas	Alat	Buah	18000	144000	8

T023	1/7/2024	Kuas	Alat	Buah	18000	108000	6
T023	1/7/2024	Pipa PVC	Bahan Logam dan PVC	Batang	50000	100000	2
T023	1/7/2024	Besi Beton	Bahan Logam dan PVC	Batang	101000	1010000	10
T024	1/7/2024	Kuas	Alat	Buah	18000	108000	6
T024	1/7/2024	Pasir	Material Konstruksi	Meter Kubik	447000	1341000	3
T024	1/7/2024	Semen	Material Konstruksi	Zak	63000	567000	9
T025	1/8/2024	Besi Beton	Bahan Logam dan PVC	Batang	101000	303000	3
T025	1/8/2024	Kuas	Alat	Buah	18000	54000	3
T025	1/8/2024	Batu Bata	Material Konstruksi	Buah	1000	2000	2
T026	1/8/2024	Besi Beton	Bahan Logam dan PVC	Batang	101000	909000	9
T026	1/8/2024	Semen	Material Konstruksi	Zak	63000	63000	1
T026	1/8/2024	Cat Kayu	Cat	Kaleng	89000	534000	6
T027	1/8/2024	Sambungan PVC	Bahan Logam dan PVC	Batang	11000	88000	8
T027	1/8/2024	Besi Beton	Bahan Logam dan PVC	Batang	101000	303000	3
T027	1/8/2024	Batu Bata	Material Konstruksi	Buah	1000	5000	5
T027	1/8/2024	Pasir	Material Konstruksi	Meter Kubik	447000	4470000	10
T028	1/8/2024	Cat Tembok	Cat	Kaleng	232000	464000	2
T028	1/8/2024	Pasir	Material Konstruksi	Meter Kubik	447000	1788000	4
T028	1/8/2024	Kuas	Alat	Buah	18000	126000	7
T029	1/9/2024	Kuas	Alat	Buah	18000	108000	6
T029	1/9/2024	Sambungan PVC	Bahan Logam dan PVC	Batang	11000	66000	6
T030	1/9/2024	Batu Bata	Material Konstruksi	Buah	1000	6000	6
T030	1/9/2024	Cat Kayu	Cat	Kaleng	89000	267000	3
T030	1/9/2024	Semen	Material Konstruksi	Zak	63000	126000	2
T030	1/9/2024	Kuas	Alat	Buah	18000	108000	6

Tabel 2. *Dataset* setelah diseleksi

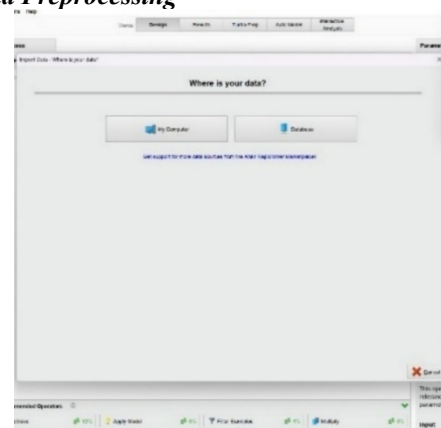
ID Transaksi	Produk	Kategori	Satuan	Kuantitas	Hasil
T001	Semen	Material Konstruksi	Zak	5	Aktif
T001	Pasir	Material Konstruksi	Meter Kubik	5	Aktif
T001	Cat Tembok	Cat	Kaleng	8	Aktif
T002	Semen	Material Konstruksi	Zak	10	Aktif
T002	Besi Beton	Bahan Logam dan PVC	Batang	1	Pasif
T002	Cat Kayu	Cat	Kaleng	7	Aktif
T003	Pasir	Material Konstruksi	Meter Kubik	9	Aktif
T003	Batu Bata	Material Konstruksi	Buah	9	Aktif
T003	Semen	Material Konstruksi	Zak	9	Aktif
T004	Cat Tembok	Cat	Kaleng	1	Pasif
T004	Cat Kayu	Cat	Kaleng	6	Aktif
T004	Kuas	Alat	Buah	5	Pasif
T005	Semen	Material Konstruksi	Zak	7	Aktif
T005	Pasir	Material Konstruksi	Meter Kubik	3	Pasif
T005	Batu Bata	Material Konstruksi	Buah	2	Pasif
T006	Besi Beton	Bahan Logam dan PVC	Batang	6	Aktif
T006	Pipa PVC	Bahan Logam dan PVC	Batang	6	Aktif
T006	Sambungan PVC	Bahan Logam dan PVC	Buah	10	Aktif
T007	Batu Bata	Material Konstruksi	Buah	3	Pasif
T007	Pasir	Material Konstruksi	Meter Kubik	1	Pasif
T007	Semen	Material Konstruksi	Zak	8	Aktif
T008	Cat Tembok	Cat	Kaleng	4	Pasif
T008	Semen	Material Konstruksi	Zak	4	Aktif
T008	Kuas	Alat	Buah	6	Aktif
T009	Besi Beton	Bahan Logam dan PVC	Batang	10	Aktif
T009	Batu Bata	Material Konstruksi	Buah	5	Pasif
T009	Semen	Material Konstruksi	Zak	3	Pasif
T010	Cat Kayu	Cat	Kaleng	10	Aktif
T010	Cat Tembok	Cat	Kaleng	7	Aktif
T010	Kuas	Alat	Buah	3	Pasif
T011	Pasir	Material Konstruksi	Meter Kubik	8	Aktif
T011	Batu Bata	Material Konstruksi	Buah	6	Aktif
T012	Sambungan PVC	Bahan Logam dan PVC	Buah	10	Aktif
T012	Cat Kayu	Cat	Kaleng	7	Aktif
T013	Kuas	Alat	Buah	4	Pasif
T013	Batu Bata	Material Konstruksi	Buah	6	Aktif
T013	Cat Tembok	Cat	Kaleng	5	Pasif
T014	Pasir	Material Konstruksi	Meter Kubik	8	Aktif
T014	Cat Kayu	Cat	Kaleng	9	Aktif
T014	Pipa PVC	Bahan Logam dan PVC	Batang	9	Aktif
T014	Kuas	Alat	Buah	4	Pasif
T015	Cat Kayu	Cat	Kaleng	2	Pasif
T015	Pasir	Material Konstruksi	Meter Kubik	10	Aktif
T015	Kuas	Alat	Buah	10	Aktif
T016	Sambungan PVC	Bahan Logam dan PVC	Buah	9	Aktif
T016	Semen	Material Konstruksi	Zak	3	Pasif
T016	Kuas	Alat	Buah	7	Aktif
T016	Besi Beton	Bahan Logam dan PVC	Batang	1	Pasif
T017	Pipa PVC	Bahan Logam dan PVC	Batang	6	Aktif
T017	Semen	Material Konstruksi	Zak	4	Aktif
T018	Pipa PVC	Bahan Logam dan PVC	Batang	9	Aktif

T018	Kuas	Alat	Buah	8	Aktif
T018	Cat Tembok	Cat	Kaleng	3	Pasif
T018	Cat Kayu	Cat	Kaleng	4	Pasif
T019	Besi Beton	Bahan Logam dan PVC	Batang	6	Aktif
T019	Pipa PVC	Bahan Logam dan PVC	Batang	4	Pasif
T020	Pipa PVC	Bahan Logam dan PVC	Batang	7	Aktif
T020	Cat Kayu	Cat	Kaleng	2	Pasif
T020	Cat Tembok	Cat	Kaleng	1	Pasif
T020	Kuas	Alat	Buah	6	Aktif
T021	Cat Tembok	Cat	Kaleng	8	Aktif
T021	Batu Bata	Material Konstruksi	Buah	8	Aktif
T022	Sambungan PVC	Bahan Logam dan PVC	Buah	9	Aktif
T022	Besi Beton	Bahan Logam dan PVC	Batang	8	Aktif
T022	Kuas	Alat	Buah	8	Aktif
T023	Kuas	Alat	Buah	6	Aktif
T023	Pipa PVC	Bahan Logam dan PVC	Batang	2	Pasif
T023	Besi Beton	Bahan Logam dan PVC	Batang	10	Aktif
T024	Kuas	Alat	Buah	6	Aktif
T024	Pasir	Material Konstruksi	Meter Kubik	3	Pasif
T024	Semen	Material Konstruksi	Zak	9	Aktif
T025	Besi Beton	Bahan Logam dan PVC	Batang	3	Pasif
T025	Kuas	Alat	Buah	3	Pasif
T025	Batu Bata	Material Konstruksi	Buah	2	Pasif
T026	Besi Beton	Bahan Logam dan PVC	Batang	9	Aktif
T026	Semen	Material Konstruksi	Zak	1	Pasif
T026	Cat Kayu	Cat	Kaleng	6	Aktif
T027	Sambungan PVC	Bahan Logam dan PVC	Buah	8	Aktif
T027	Besi Beton	Bahan Logam dan PVC	Batang	3	Pasif
T027	Batu Bata	Material Konstruksi	Buah	5	Pasif
T027	Pasir	Material Konstruksi	Meter Kubik	10	Aktif
T028	Cat Tembok	Cat	Kaleng	2	Pasif
T028	Pasir	Material Konstruksi	Meter Kubik	4	Aktif
T028	Kuas	Alat	Buah	7	Aktif
T029	Kuas	Alat	Buah	6	Aktif
T029	Sambungan PVC	Bahan Logam dan PVC	Buah	6	Aktif
T030	Batu Bata	Material Konstruksi	Buah	6	Aktif
T030	Cat Kayu	Cat	Kaleng	3	Pasif
T030	Semen	Material Konstruksi	Zak	2	Pasif
T030	Kuas	Alat	Buah	6	Aktif

Pada Tabel 2 menunjukkan bahwa *dataset* dari *kaggle* ada yang dihapus karena tidak dibutuhkan, kemudian data juga ditambah hasil dengan atribut “Aktif” dan “Pasif”. Dengan rumus *Excel* seperti:

=IF(OR(AND(OR(D2="Zak",D2="Meter Kubik"),E2>=4), AND(OR(D2="Kaleng",D2="Buah",D2="Batang"),E2>=6)), "Aktif", "Pasif")

3.2. Data Preprocessing



Gambar 3. Import Data

The screenshot shows the AWS Glue console interface during the setup of a crawler. The 'Select the data location' step is active. The 'Data location' dropdown is set to 'S3'. The 'Data source' is 'Amazon S3'. The 'Data location' field is populated with 's3://glue-logs-123456789012-us-east-1/'. The 'Data source' field is populated with 'Amazon S3'. The 'Data location' field is populated with 's3://glue-logs-123456789012-us-east-1/'. The 'Data source' field is populated with 'Amazon S3'.

Kemudian dilanjutkan dengan tahap pencarian lokasi penyimpanan data.

Name	Type	Missing	Statistics	Filter (3 attributes)
ID Transaksi	Polynomial	0	TC34 (1) TC07 (1)	TC07 (1), TC31 (1), ... [more]
hasil	Polynomial	5	Good Bad Very	Very
prediction(hasil)	Polynomial	0	Pelanggan Pasif (1) Pelanggan Aktif (4)	Pelanggan Aktif (4), Pelanggan Pasif (1)
Confidence Pelanggan dari confidence(Pelanggan Aktif)	Real	0	0 1	0.850
Confidence Pelanggan dari confidence(Pelanggan Pasif)	Real	0	0 1	0.310
Produk	Polynomial	0	Semen (1) Batu Bata (1)	Batu Bata (1), Besi Beton (1), ... [more]
Kategori	Polynomial	0	Cat (1) Material Konstruksi (2)	Material Konstruksi (3), Bahan Logam dan PVC (1)
Satuan	Polynomial	0	meter kubik (1) Zak (1)	Zak (1), batang (1), ... [more]
Kuantitas	Integer	0	1 9	0.600

Hasil proses *cleansing* pada *data testing* menunjukkan bahwa data yang digunakan telah dibersihkan dari nilai kosong.

Process

Input Data - Format your columns.

Process

Format your columns.

☐ Replace errors with missing values

ID Transaksi	Produk	Kategori	Satuan	Kuantitas	Hasil	
polynomial	polynomial	polynomial	polynomial	integer	polynomial	
1	T001	Materai Konstruksi	Zak	5	Pelangan Asst	
2	T001	Materai Konstruksi	Meter Kubik	5	Pelangan Asst	
3	T001	Cat	Kaleng	8	Pelangan Asst	
4	T002	Semen	Materai Konstruksi	Zak	10	Pelangan Asst
5	T002	Besi Beton	Bahan Logam dan PVC	Batang	1	Pelangan Pasif
6	T002	Cat Kayu	Cat	Kaleng	7	Pelangan Asst
7	T003	Pasir	Materai Konstruksi	Meter Kubik	9	Pelangan Asst
8	T003	Batu Batu	Materai Konstruksi	Buah	9	Pelangan Asst
9	T003	Semen	Materai Konstruksi	Zak	9	Pelangan Asst
10	T004	Cat Tembok	Cat	Kaleng	11	Pelangan Pasif
11	T004	Cat Kayu	Cat	Kaleng	6	Pelangan Asst
12	T004	Kuas	Alat	Buah	0	Pelangan Asst
13	T005	Semen	Materai Konstruksi	Zak	7	Pelangan Asst
14	T005	Pasir	Materai Konstruksi	Meter Kubik	3	Pelangan Pasif
15	T005	Batu Batu	Materai Konstruksi	Buah	2	Pelangan Pasif
16	T006	Besi Beton	Bahan Logam dan PVC	Batang	6	Pelangan Asst
17	T006	Pipa PVC	Bahan Logam dan PVC	Batang	6	Pelangan Asst
18	T006	Sambungan PVC	Bahan Logam dan PVC	Buah	10	Pelangan Asst
19	T007	Batu Batu	Materai Konstruksi	Buah	3	Pelangan Pasif

no problems.

← Previous Erase Jump to Cancel

Report Data - Format your columns.

Format your columns.

☐ Replace errors with missing values (1)

ID Transaksi	Produk	Kategori	Satuan	Kuantitas	hasil
polynomial	polynomial	polynomial	polynomial	Integer	polynomial
1 T001	Semen	Material Konstruksi	Zak	5	Pelanggan Aktif
2 T001	Pasir	Material Konstruksi	Meter Kubik	5	Pelanggan Aktif
3 T001	Cal Te				Pelanggan Aktif
4 T002	Semen				Pelanggan Aktif
5 T002	Besi B				Pelanggan Pasif
6 T002	Cal Ka				Pelanggan Aktif
7 T002	Pasir				Pelanggan Aktif
8 T003	Batu B				Pelanggan Aktif
9 T003	Semen				Pelanggan Aktif
10 T004	Cal Te				Pelanggan Pasif
11 T004	Cal Klay				Pelanggan Aktif
12 T004	Kues	Alat	Buah	5	Pelanggan Pasif
13 T005	Semen	Material Konstruksi	Zak	7	Pelanggan Aktif
14 T005	Pasir	Material Konstruksi	Meter Kubik	3	Pelanggan Pasif
15 T005	Batu Beta	Material Konstruksi	Buah	2	Pelanggan Pasif
16 T006	Besi Beton	Bahan Logam dan PVC	Batang	6	Pelanggan Aktif
17 T006	Pipa PVC	Bahan Logam dan PVC	Batang	6	Pelanggan Aktif
18 T006	Sambungan PVC	Bahan Logam dan PVC	Buah	10	Pelanggan Aktif
19 T007	Batu Beta	Material Konstruksi	Buah	3	Pelancoran Pasif

Change role

Please enter the new role:

label

OK Cancel

no problem

Proses pengubahan *role* atribut dilakukan untuk menetapkan fungsi masing-masing atribut seperti *label* dan *id*.

The screenshot shows the Orange3 data mining environment. The workflow consists of the following widgets and connections:

- Read Excel (1)**: Connected to **k-NN** (input 1).
- Read Excel (2)**: Connected to **k-NN** (input 2) and **Apply Model** (input 1).
- k-NN**: Connected to **Apply Model** (input 2).
- Apply Model**: Connected to **Performance** (input 1).
- Performance**: Connected to **File** (input 1).
- File**: Configured to save the results as **res.csv**.

The **Apply Model** widget shows a green checkmark, indicating that the model was successfully applied. The **Performance** widget shows a yellow lightbulb icon, indicating that the performance metrics were calculated.

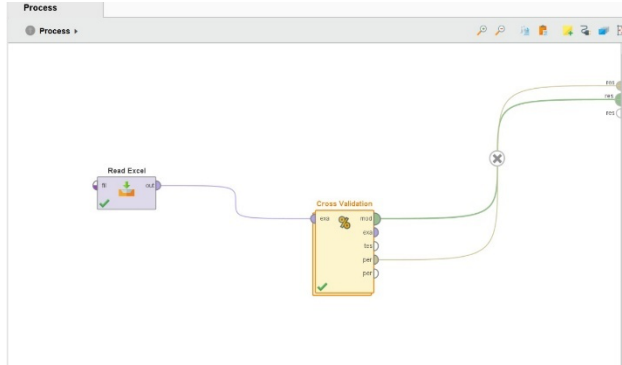
Dalam proses pengolahan data menggunakan *RapidMiner*, dilakukan penambahan beberapa *operator* utama untuk membentuk alur analisis. *Operator 'Read Excel'* digunakan untuk membaca dan mengimpor data dari file *Excel* ke dalam lingkungan kerja *RapidMiner*. Selanjutnya, *operator 'K-NN' (K-Nearest Neighbors)* digunakan sebagai algoritma klasifikasi untuk membangun model prediksi berdasarkan kedekatan antar data. Setelah model dibentuk, *operator 'Apply Model'* diterapkan untuk menguji model tersebut terhadap *data testing*. Terakhir, *operator 'Performance'* digunakan untuk mengevaluasi kinerja model berdasarkan metrik tertentu seperti *accuracy*, *precision*, dan *recall*, guna menilai seberapa baik model dalam melakukan klasifikasi.

Row No.	ID Transaksi	hasil	prediction(h...	confidence(...	confidence(...	Produk	Kategori	Satuan	Kuantitas
1	T031	?	Pelanggan Aktif	0.823	0.177	Semen	Material Kon...	Zak	5
2	T032	?	Pelanggan Aktif	0.625	0.375	Pasir	Material Kon...	meter kubik	5
3	T033	?	Pelanggan Aktif	1	0	Cat Tembok	Cat	kaleng	8
4	T034	?	Pelanggan P...	0	1	Besi Beton	Bahan Loga...	batang	1
5	T007	?	Pelanggan Aktif	1	0	Batu Bata	Material Kon...	buah	9

Lisensi: Creative Commons Attribution 4.0 International (CC BY 4.0)

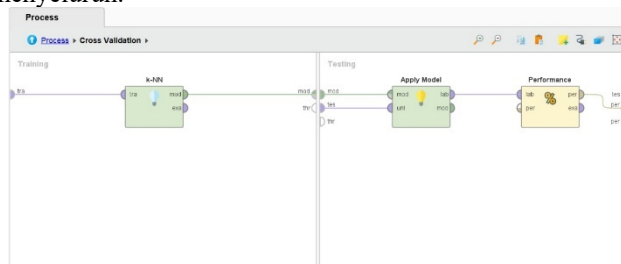
Hasil seleksi data testing menunjukkan bahwa data yang digunakan telah dipilih berdasarkan atribut yang relevan dan sesuai dengan tujuan analisis, sehingga siap digunakan untuk proses pengujian model.

3.5. Evaluation



Gambar 10. Menambahkan Operator Read Excel dan Cross Validation

Operator 'Read Excel' digunakan untuk mengimpor data dari file Excel ke dalam RapidMiner, sedangkan operator 'Cross Validation' digunakan untuk membagi data ke dalam beberapa subset guna melakukan pelatihan dan pengujian model secara bergantian, sehingga menghasilkan evaluasi kinerja model yang lebih akurat dan menyeluruh.



Gambar 11. Menambahkan Operator K-NN, Apply Model dan Performance

Operator 'K-NN' digunakan untuk membangun model klasifikasi, 'Apply Model' untuk menerapkan model pada data uji, dan 'Performance' untuk mengevaluasi hasil prediksi model.

	True Pelanggan Aktif	True Pelanggan Pasif	class precision
pred. Pelanggan Aktif	56	4	93.33%
pred. Pelanggan Pasif	1	29	96.67%
class recall	96.25%	87.88%	

Gambar 12. Performance Recall

Metrik evaluasi 'Recall' mengukur kemampuan model dalam menangkap seluruh data positif yang benar.

	True Pelanggan Aktif	True Pelanggan Pasif	class precision
pred. Pelanggan Aktif	56	4	93.33%
pred. Pelanggan Pasif	1	29	96.67%
class recall	96.25%	87.88%	

Gambar 13. Performance Precision

Metrik evaluasi 'Precision' menunjukkan ketepatan model dalam memprediksi data positif.

	True Pelanggan Aktif	True Pelanggan Pasif	class precision
pred. Pelanggan Aktif	56	4	93.33%
pred. Pelanggan Pasif	1	29	96.67%
class recall	96.25%	87.88%	

Gambar 14. Performance Accuracy

Metrik evaluasi 'Accuracy' adalah rasio keseluruhan prediksi yang benar terhadap total data.

4. PENUTUP

Kesimpulan

Hasil penelitian menunjukkan bahwa kategori aktif memiliki tingkat penjualan terbaik, seperti yang ditunjukkan oleh hasil analisis data, di mana ada 56 data yang diprediksi untuk kategori aktif dan hanya 29 data untuk kategori pasif. Kesimpulan ini menunjukkan bahwa kategori aktif berkontribusi lebih banyak terhadap penjualan daripada kategori pasif.

Saran

1. Untuk meningkatkan akurasi model prediksi, penelitian lebih lanjut diperlukan dengan menggunakan dataset yang lebih besar dan bervariasi. Selain itu, dipertimbangkan untuk memasukkan fitur tambahan seperti waktu promosi, musim, lokasi pembelian, dan jenis pelanggan.
2. Untuk membandingkan kinerja metode K-NN dan mendapatkan hasil klasifikasi yang lebih optimal, disarankan untuk menguji kombinasi metode K-NN dengan algoritma lain seperti Decision Tree, Random Forest, atau SVM.
3. Selanjutnya, model klasifikasi dapat ditambahkan ke dalam sistem informasi penjualan toko yang terintegrasi. Ini akan memungkinkan hasil prediksi untuk digunakan dalam proses pengambilan keputusan bisnis secara real-time.

5. DAFTAR PUSTAKA

- [1] N. W. Mardiyah, N. Rahaningsih, and I. Ali, "Penerapan Data mining Menggunakan Algoritma K-Nearest Neighbor Pada Prediksi Pemberian Kredit di Sektor Finansial," Feb. 2024. [Online]. Available: www.kaggle.com
- [2] M. N. Maskuri, K. Sukerti, and R. M. Herdian Bhakti, "Penerapan Algoritma K-Nearest Neighbor (KNN) untuk Memprediksi Penyakit Stroke Stroke Disease Predict Using KNN Algorithm," *Jurnal Ilmiah Intech: Information Technology Journal of UMUS*, vol. 4, no. 1, pp. 130–140, May 2022.
- [3] V. P. Sabandar and R. Ahmad, "Sistem Pendukung Keputusan Penentuan Produk Terbaik Menggunakan Weighted Product Method," *Jurnal Ilmiah Computer Science*, vol. 1, no. 2, pp. 58–68, Jan. 2023, doi: 10.58602/jics.v1i2.7.
- [4] P. Guru, J. Sathyapriya, K. V. R. Rajandran, J. Bhuvaneswari, and C. Parimala, "International Journal of INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING Product Sales Forecasting and Prediction Using Machine Learning Algorithm," 2023. [Online]. Available: www.ijisae.org

- [5] D. Safitri and M. Fakhriza, "K-Nearest Neighbor (KNN) Algorithm to Determine the Stock of Building Material Store Materials," *Journal of Computer System and Informatics (JoSYC)*, vol. 5, no. 4, pp. 850–857, Aug. 2024, doi: 10.47065/josyc.v5i4.5731.
- [6] A. S. Miha Djami, N. W. Utami, and A. A. I. I. Paramitha, "The Prediction Of Product Sales Level Using K-Nearest Neighbor and Naive Bayes Algorithms (Case Study: PT Kotamas Bali)," *Jurnal Pilar Nusa Mandiri*, vol. 19, no. 2, pp. 77–84, Sep. 2023, doi: 10.33480/pilar.v19i2.4420.
- [7] R. Syabania and N. Rosmawarni, "Perancangan Aplikasi Customer Relationship Management (CRM) Pada Penjualan Barang Pre-Order Berbasis Website," 2021.
- [8] F. D. Elyana and S. Arifin, "Analisis Strategi Pemasaran Secara Online Melalui TikTok dan Instagram Terhadap Volume Penjualan Skincare Justmine Beauty Pada Agen Yuyun di Lamongan," *Neraca Jurnal Ekonomi, Manajemen dan Akuntansi*, vol. 1, no. Penjualan, pp. 344–352, 2023, [Online]. Available: <http://jurnal.kolibi.org/index.php/neraca>
- [9] V. Immanuel Sunarko, D. Lizard Sambawo Dimara, P. Sandya Etniko Siagian, D. Manalu, and F. Tri Anggraeny, "Implementasi K-Fold Dalam Prediksi Hasil Produksi Agrikultur Pada Algoritma K-Nearest Neighbor (KNN)," Mar. 2025. [Online]. Available: <https://satudata.pertanian.go.id/>
- [10] Y. Zidane, N. Betha, I. Sari, A. Maulana, and G. Primaya, "Penerapan *Data mining* Dalam Klasifikasi Data Transaksi Produksi Koperasi di SMK PGRI 2 KARAWANG," 2025.
- [11] L. I. Keksi, D. A. Lestari, and B. Prasetyo, "Application of K-Nearest Neighbor Algoritihm in Classification of Engine Performance in Car Companies Using *RapidMiner*," *Journal of Student Research Exploration*, vol. 2, pp. 120–130, Jun. 2024.
- [12] V. Ariyani, P. Putri, A. B. Prasetijo, and D. Eridani, "Perbandingan Kinerja Algoritme Naive Bayes dan K-Nearest Neighbor (KNN) Untuk Prediksi Harga Rumah," *Transmisi: Jurnal Ilmiah Teknik Elektro*, vol. 4, pp. 162–171, Oct. 2022, doi: 10.14710/transmisi.24.4.162-171.
- [13] A. Prayoga Permana, K. Ainayah, and K. Fahmi Hayati Holle, "Analisis Perbandingan Algoritma Decision Tree, kNN, dan Naive Bayes untuk Prediksi Kesuksesan Start-up," 2021. [Online]. Available: <https://www.kaggle.com/manishkc06/startup-success-prediction>.
- [14] N. Sitompul, "Rapid Miner Testing With The Knn Algorithm," 2023.
- [15] E. Rosida, A. Firmansyah, and U. Pelita Bangsa Corresponding Author, "Comparative Analysis of Classification of K-Nearest Neighbor (KNN) Algorithm and Decision Tree in Breast Cancer Using *RapidMiner*," *International Journal of Applied Research and Sustainable Sciences (IJARSS)*, vol. 2, no. 12, pp. 1039–1060, 2024, doi: 10.59890/ijarss.v2i12.48.
- [16] L. H. Jufri and D. Dasari, "Evaluation of the implementation of machine learning algorithm K-Nearest Neighbors (KNN) using rapid miner on junior high school student learning outcomes," *Jurnal Gantang*, vol. 8, no. 2, pp. 193–197, Dec. 2023, doi: 10.31629/jg.v8i2.6590.
- [17] N. T. Ujianto, Gunawan, H. Fadillah, A. P. Fanti, A. D. Saputra, and I. G. Ramadhan, "Penerapan algoritma K-Nearest Neighbors (KNN) untuk klasifikasi citra medis," *IT-Explore: Jurnal Penerapan Teknologi Informasi dan Komunikasi*, vol. 4, no. 1, pp. 33–43, Feb. 2025, doi: 10.24246/itexplore.v4i1.2025.pp33-43.
- [18] A. A. Amer, S. D. Ravana, and R. A. A. Habeeb, "Effective k-nearest neighbor models for data classification enhancement," *J Big Data*, vol. 12, no. 1, Dec. 2025, doi: 10.1186/s40537-025-01137-2.
- [19] D. Damayanti, "Perbandingan Akurasi Software *RapidMiner* dan Weka Menggunakan Algoritma K-Nearest Neighbor (K-NN)," *Jurnal Syntax Admiration*, vol. 2, no. 6, pp. 994–1006, Jun. 2021, doi: 10.46799/jsa.v2i6.247.
- [20] T. Hidayat, Y. Handayani, and A. Syaifudin, "Implementasi Algoritma K-Nearest Neighbor Untuk Meningkatkan Penjualan Produk Meuble dan Furniture," *Jurnal Ilmiah Rekayasa dan Manajemen Sistem Informasi*, vol. 9, no. 2, pp. 118–124, Feb. 2023.