

Perbandingan Klasifikasi Tingkat Penjualan Buah di Supermarket dengan Pendekatan Algoritma Decision Tree, Naive Bayes dan K-Nearest Neighbor

Herlambang Priatmojo¹, Feri Saputra², Muhammad Hadi Prasetyo³, Diah Puspitasari⁴, Dini Nurlaela⁵

^{1,2,3,4,5}Universitas Bina Sarana Informatika

Jl. Kramat Raya no. 98, Senen, Jakarta Pusat, Indonesia

e-mail: ¹19190052@bsi.ac.id, ²19190088@bsi.ac.id, ³19190053@bsi.ac.id, ⁴diah.puspitasari@bsi.ac.id, ⁵dini.dur@bsi.ac.id

Abstrak - Tujuan dari penelitian ini adalah membandingkan tiga metode klasifikasi, yakni Decision Tree, Naive Bayes, dan K-Nearest Neighbor untuk memberikan rekomendasi terkait penelitian ini yaitu mengurangi tingkat damage atau kerusakan barang akibat dari over order yang menimbulkan kerugian pada supermarket. Sumber data yang digunakan adalah penjualan buah dari Foodmart Supermarket. Untuk menguji performa ketiga algoritma tersebut, diukur akurasi, presisi, recall, serta nilai ROC AUC-nya. Hal ini dilakukan untuk menentukan teknik klasifikasi terbaik yang dapat memproses dataset yang diberikan. Decision Tree memiliki akurasi tertinggi sebesar 92,31% dengan nilai AUC sebesar 93,8%, K-Nearest Neighbor memiliki akurasi 88,46% dengan nilai AUC sebesar 99,4%, sementara Naive Bayes memiliki akurasi 73,08% dengan nilai AUC sebesar 79,4%. Diharapkan hasil dari penelitian ini dapat dijadikan sebagai referensi untuk penelitian teknik klasifikasi lainnya, baik dengan menggunakan data yang sama maupun teknik klasifikasi yang berbeda.

Kata Kunci : Decision Tree, Naive Bayes, K-Nearest Neighbor

Abstracts - The aim of this research is to compare three classification methods: Decision Tree, Naive Bayes, and K-Nearest Neighbor, in order to provide recommendations related to reducing the level of damage or loss of goods caused by over-ordering, which leads to losses in the supermarket. The data source used is fruit sales data from Foodmart Supermarket. To test the performance of the three algorithms, their accuracy, precision, recall, and ROC AUC values were measured. This was done to determine the best classification technique that can process the given dataset. Decision Tree achieved the highest accuracy of 92.31% with an AUC value of 93.8%, K-Nearest Neighbor had an accuracy of 88.46% with an AUC value of 99.4%, while Naive Bayes had an accuracy of 73.08% with an AUC value of 79.4%. It is expected that the results of this research can be used as a reference for other classification technique studies, both using the same data and different classification techniques.

Keywords : Decision Tree, Naive Bayes, K-Nearest Neighbor

PENDAHULUAN

Bisnis retail adalah salah satu bisnis menjanjikan yang dapat dijalankan oleh siapa saja. Industri ritel merupakan salah satu sektor yang berkembang, berubah, dan dinamis dengan cepat di seluruh dunia (Alfiah et al., 2018). Saat ini, bisnis retail mengalami perubahan dari model toko lokal independen atau toko di jalan utama menjadi pusat perbelanjaan modern berskala nasional dan internasional, supermarket, pasar swalayan, toko serba ada, dan lain sebagainya (Sunyoto & Mulyono, 2022).

Dalam industri ritel, seperti di supermarket atau swalayan, penting bagi pemilik usaha untuk memahami perilaku pembelian pelanggan dan mengidentifikasi produk-produk yang paling diminati. Dengan pemahaman yang baik tentang produk-produk yang laris, pemilik usaha dapat mengoptimalkan stok, penempatan produk di rak, serta strategi pemasaran untuk meningkatkan penjualan dan keuntungan.

Di era perdagangan modern, persaingan dalam industri ritel semakin ketat. Salah satu faktor kunci keberhasilan suatu toko ritel adalah kemampuannya untuk mengidentifikasi dan mempromosikan produk-produk yang paling diminati oleh pelanggan. Dalam hal ini, Foodmart Supermarket yang merupakan salah satu toko ritel yang menjual berbagai produk salah satunya adalah produk Fresh (Produce, Meat & Fish, Bakery dan Dairy Frozen), perlu memanfaatkan strategi yang efektif untuk mengklasifikasikan produk-produknya agar dapat mengoptimalkan penjualan dan meningkatkan kepuasan pelanggan (Sufia Maulida, Fikri Hamidy, 2020).

Foodmart Supermarket adalah sebuah perusahaan ritel terkemuka di Indonesia yang dinaungi oleh PT.

MATAHARI PUTRA PRIMA, Tbk yang termasuk dari grup Hypermart, Foodmart Supermarket, Boston, Foodmart Primo, Foodmart Express dan Hyfresh. Pada perusahaan ini terdapat permasalahan terkait dengan pendataan produk yang tidak laris sehingga mengakibatkan menumpuknya barang-barang yang mendekati kadaluarsa (expired). Hal tersebut dapat terjadi karena tidak memiliki pengambilan keputusan secara klasifikasi terhadap produk tidak laris yang bisa membantu mempermudah proses pendataan produk. Dengan banyaknya produk produk segar yang dijual selama ini hanya melihat tingkat laris atau tidak larisnya suatu produk dengan penjualan terbanyak selama periode 1 bulan. Dan akhirnya menimbulkan over order disetiap produk fresh yang tidak laris. Dari pihak toko untuk menghindarkan penyusutan barang dari produk fresh tersebut melakukan potongan harga yang menimbulkan kerugian pada supermarket. Tujuan dari penelitian ini adalah membandingkan tiga metode klasifikasi, yakni Decision Tree, Naive Bayes, dan K-Nearest Neighbor untuk memberikan rekomendasi terkait penelitian ini yaitu mengurangi tingkat damage atau kerusakan barang akibat dari over order yang menimbulkan kerugian pada supermarket. Sumber data yang digunakan adalah penjualan buah dari Foodmart Supermarket.

Penelitian yang dilakukan oleh Ahmad Zakir, Yermia Ndruru, Edrian Hadinata dan Ihsan Lubis (2020) berjudul "Penerapan Data Mining Untuk Klasifikasi Data Penjualan Makanan Terlaris Dengan Algoritma C4.5". Penelitian ini menggunakan algoritma Data Mining C4.5. Hasil penelitian ini menunjukkan bahwa Algoritma C4.5 dapat mempermudah pengambilan keputusan dengan mengonversi data ke dalam bentuk pohon keputusan. Proses ini dilakukan dengan mempertimbangkan nilai entropy dan gain dari setiap atribut data. Dan juga disarankan untuk mendapatkan hasil prediksi yang lebih akurat, diperlukan jumlah data yang lebih besar. Dengan kata lain, semakin banyak data yang digunakan, maka hasil prediksinya akan menjadi lebih akurat.

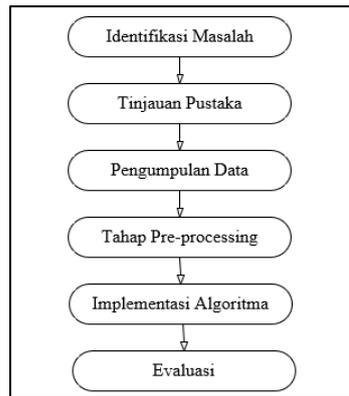
Penelitian yang dilakukan oleh Sri Puspita Dewi, Nurwati dan Elly Rahayu (2022) berjudul "Penerapan Data Mining Untuk Prediksi Produk Terlaris Menggunakan Metode K-Nearest Neighbor". Penelitian ini menggunakan algoritma Data Mining K-Nearest Neighbor. Hasil dari penelitian ini menunjukkan bahwa penggunaan teknik data mining dan algoritma K-Nearest Neighbor menghasilkan prediksi dengan tingkat akurasi yang tinggi. Dengan menerapkan metode K-Nearest Neighbor dalam sistem aplikasi dapat membantu perusahaan dalam mengatasi masalah yang dihadapi.

Penelitian yang dilakukan oleh Odi Nurdiawan dan Noval Salim (2018) berjudul "Penerapan Data Mining Pada Penjualan Barang Menggunakan Metode Naive Bayes Classifier Untuk Optimasi Strategi Pemasaran". Penelitian ini menggunakan algoritma Data Mining Naive Bayes. Hasil penelitian ini menunjukkan bahwa Penggunaan metode klasifikasi Naive Bayes pada dataset penjualan barang mencapai tingkat optimalisasi lebih dari 65%. Dan juga Berdasarkan hasil penelitian, nilai p-value (sig) ditemukan sebesar 0,807, yang lebih besar dari 0,05. Oleh karena itu, hipotesis nol (H_0) diterima dan hipotesis alternatif (H_a) ditolak.

Penelitian yang dilakukan oleh Diah Puspitasari, Kresna Ramanda, Adi Supriyatna, Mochamad Wahyudi, Erma Delima Sikumbang dan Sulaeman Hadi Sukmana (2020) berjudul "Comparison of Data Mining Algorithms Using Artificial Neural Network (ANN) and Naive Bayes for Preterm Birth Prediction". Penelitian ini menggunakan algoritma Data Mining Artificial Neural Network (ANN) dan Naive Bayes. Berdasarkan penelitian yang dilakukan terhadap prediksi kelahiran prematur, dengan membandingkan akurasi dan nilai ROC pada dua algoritma yang berbeda, yaitu Artificial Neural Network (ANN) dan Naive Bayes. Hasil penelitian menunjukkan bahwa prediksi kelahiran prematur menggunakan algoritma Artificial Neural Network (ANN) memiliki tingkat akurasi sebesar 90,67% dan nilai ROC sebesar 0,954. Sedangkan algoritma Naive Bayes memiliki tingkat akurasi sebesar 84,53% dan nilai ROC sebesar 0,929. Dari hasil tersebut, dapat disimpulkan bahwa algoritma Artificial Neural Network (ANN) memiliki keunggulan dengan tingkat akurasi yang lebih tinggi sebesar 6,14% dan nilai ROC yang lebih tinggi sebesar 0,025 dalam memprediksi kelahiran prematur. Untuk penelitian selanjutnya, dapat dilakukan pemanfaatan algoritma optimisasi guna meningkatkan akurasi.

METODE PENELITIAN

Tahapan penelitian dituangkan pada gambar di bawah ini



Gambar. 1. Tahapan Penelitian

1. Identifikasi Masalah

Menentukan permasalahan penelitian Tingkat Penjualan Buah di supermarket, yaitu membandingkan performa algoritma Decision Tree, Naive Bayes dan K-Nearest Neighbor dalam memprediksi tingkat penjualan buah di supermarket. Langkah ini dilakukan dengan membaca literatur seperti buku dan jurnal untuk menemukan isu yang dapat diangkat dari penelitian ini dengan menggunakan dataset berupa log yang tercatat saat penggunaan internet melalui firewall sebagai pelindung perangkat. Dataset ini akan digunakan untuk melakukan klasifikasi dan memilih algoritma yang paling relevan untuk penelitian. Algoritma yang digunakan dalam penelitian termasuk Naive Bayes, Decision Tree, dan K-Nearest Neighbor untuk evaluasi hasil. Dengan melakukan Analisa dengan metode Decision Tree, Naive Bayes dan K-Nearest Neighbor kita bisa menentukan produk mana yang laris dan tidak laris agar tidak terjadi penyusutan buah atau busuk, dengan accuracy yang tepat atau dengan nilai accuracy yang tinggi.

2. Tinjauan Pustaka

Melakukan tinjauan pustaka terkait teori-teori yang relevan, seperti data mining, algoritma *Decision Tree*, *Naive Bayes*, *K-Nearest Neighbor* dan faktor-faktor yang mempengaruhi penjualan buah di supermarket karena banyak penyusutan buah di Gudang penyimpanan di sebabkan dari over order buah dan penjualan buah tersebut ada yang tidak laris atau slow moving dikarenakan harga nilai jual buah tersebut sedang tidak promo di bulan tersebut hasil wawancara dengan supervisor Bpk Syahrullah. Meninjau penelitian terkait yang telah dilakukan untuk mendapatkan pemahaman yang baik tentang topik penelitian dan dapat memastikan produk agar tidak terjadi penyusutan atau rusak efek dari over order dan kurangnya minat pelanggan.

3. Pengumpulan Data

Data yang digunakan untuk penelitian ini berupa data penjualan buah di Foodmart Supermarket periode bulan April 2023. Dataset mencakup 13 atribut dan berisi 66 data, proses data ini harus melalui beberapa langkah, seperti tahap pre-processing yang bertujuan untuk membersihkan data yang terduplikasi, memvalidasi data, dan menormalisasi data. Oleh karena itu, tahap ini memerlukan waktu yang lebih lama.

Mengumpulkan data penjualan buah di supermarket, termasuk informasi tentang jenis buah, jumlah penjualan, harga, waktu dan tanggal penjualan, serta atribut lain yang relevan. Juga mengumpulkan data atribut yang dapat mempengaruhi tingkat penjualan buah, seperti informasi promosi, musim, ulasan pelanggan, dan lokasi supermarket.

4. Tahap Pre-processing

Pre-processing dilaksanakan guna memproses data agar menjadi bersih dan dapat dimanfaatkan secara optimal sehingga mendapatkan hasil evaluasi yang lebih baik. Pada tahap ini dilakukan beberapa prosedur seperti seleksi dan pemilihan pada setiap atribut, normalisasi data, serta menangani missing value. Pada tahap ini, langkah awal yang harus dilakukan ialah melakukan pengecekan missing value pada kumpulan data yang digunakan. Kemudian, Operator Split Data digunakan untuk menguji validitas data. Split Data merupakan salah satu Operator dalam Rapidminer yang membagi data menjadi data training dan data testing. Setelah itu, data siap untuk diproses dalam

modelling (Wirawan, 2020).

5. Implementasi Algoritma

Mengimplementasikan algoritma Decision Tree, Naive Bayes dan K-Nearest Neighbor menggunakan perangkat lunak atau platform analisis data yang menyediakan algoritma-algoritma tersebut.

Decision Tree

Pohon Keputusan (Decision Tree) terbentuk setelah melalui serangkaian perhitungan entropy dan information gain, yang diulang-ulang hingga seluruh atribut pohon memiliki kelas dan tidak dapat dilakukan perhitungan lebih lanjut (Cynthia & Ismanto, 2018). Decision tree merupakan suatu struktur yang digunakan untuk membantu proses pengambilan keputusan yang terdiri dari simpul internal, tepi dan simpul daun. Setiap node keputusan disebut tes pada atribut atau subset dari atribut, dan masing-masing sisi mewakili nilai tertentu atau rentang nilai input atribut (Mahmud Nawawi et al., 2019). Dalam data mining decision tree, teknik yang umum digunakan adalah algoritma CART (Classification and Regression Tree) atau C4.5, yang memungkinkan pembuatan pohon keputusan dengan menghitung tingkat kemurnian atau keheterogenan data pada setiap simpul. Dengan menggunakan pohon keputusan, kita dapat melakukan prediksi atau klasifikasi terhadap data baru dengan mengikuti jalur pada pohon keputusan, dimulai dari akar pohon hingga mencapai simpul yang merepresentasikan hasil prediksi atau klasifikasi.

Naive Bayes

Naive Bayes adalah metode klasifikasi yang banyak digunakan berdasarkan teori Bayes (Puspitasari et al., 2020) dan merupakan metode yang dapat digunakan untuk mengklasifikasikan data (Damuri et al., 2021). Klasifikasi ini dapat memperkirakan kemungkinan keanggotaan suatu tupel ke dalam kelas tertentu seperti probabilitas. Teori Bayes menjadi dasar dari metode klasifikasi ini. Metode Naive Bayes mengasumsikan bahwa nilai atribut pada kelas tertentu tidak saling mempengaruhi dengan nilai atribut lainnya (Widaningsih, 2019).

$$P(x|y) = \frac{P(y|x) P(x)}{P(y)} \quad (1)$$

K-Nearest Neighbor

Metode K-Nearest Neighbor merupakan teknik yang digunakan untuk mengklasifikasikan objek dengan mempertimbangkan data pembelajaran yang memiliki jarak terdekat dengan objek tersebut. Tujuannya adalah untuk mengklasifikasikan objek baru dengan mempertimbangkan atribut dan sampel pelatihan. K-Nearest Neighbor dapat dijelaskan sebagai suatu teknik yang menggunakan tetangga terdekat atau data sebelumnya sebagai sampel untuk menghasilkan klasifikasi akhir (Arifin, 2019). K-Nearest Neighbor (KNN) termasuk dalam kategori pembelajaran berbasis contoh atau instance-based learning. Algoritma ini juga merupakan salah satu teknik pembelajaran yang disebut lazy learning. KNN dilakukan dengan mencari kelompok k objek dalam data pelatihan yang paling mirip dengan objek pada data uji atau data testing (Yunita, 2017).

$$d_i = \sqrt{\sum_{1=i}^p (x_{2i} - x_{1i})^2} \quad (2)$$

6. Evaluasi

Menggunakan metrik evaluasi seperti akurasi, presisi, recall atau area under the curve (AUC) untuk mengukur kinerja masing-masing algoritma pada data testing.

Precision

$$Precision = \frac{\text{True Positive}}{\text{True Positive} + \text{False positive}} \quad (3)$$

Recall

$$Recall = \frac{\text{True Positive}}{\text{True Positive} + \text{False Negative}} \quad (4)$$

Akurasi

$$Akurasi = \frac{\text{True Positive} + \text{True Negative}}{\text{False Positive} + \text{False Negative} + \text{True Positive} + \text{True Negative}} \quad (5)$$

Kemudian, ROC (Receiver Operating Characteristics) akan dipakai untuk menampilkan performa dua dimensi dari setiap klasifikasi yang diuji. Garis horizontal menunjukkan false positive, sementara garis vertikal menunjukkan true positive. AUC (Area Under Curve) merupakan area yang terletak di bawah grafik ROC. Kualitas klasifikasi dapat dikategorikan berdasarkan nilai AUC-nya, dan tabel berikut menunjukkan kualitas klasifikasi yang sesuai dengan nilai AUC.

Tabel. 1. Kriteria AUC

Nilai AUC	Penjelasan
90%-100%	Excellent
80%-90%	Good
70%-80%	Fair
60%-70%	Poor
<60%	Failure

Sumber: Hasil Penelitian (2023)

HASIL DAN PEMBAHASAN

3.1. Pengumpulan Data

Data ini didapat dari data penjualan buah di Foodmart Supermarket selama bulan April. Data tersebut terdiri dari 66 entri dan memiliki 13 atribut.

Tabel. 2. Penjelasan fitur pada dataset.

Fitur	Deskripsi
SKU	Kode Barcode
DESCRIPTION	Keterangan / Nama Produk
SUPPLIER	Nama Vendor
GRADE	Tipe Produk
STOCK ON HAND	Jumlah Stok Barang yang tersedia
ON ORDER	Barang yang sedang dalam proses pengiriman
COST	Harga beli dari supplier
RETAIL	Harga jual
QTY	Jumlah produk yang terjual
U/M	Satua (KG/PCS)
TARGET	Target penjualan produk dalam 1 bulan
TOTAL SALES	Penjualan yang tercapai dalam 1 bulan
LABEL	Kelas (LARIS / TIDAK LARIS)

Sumber: Hasil Penelitian (2023)

Tabel. 3. Pengelompokan Dataset berdasarkan kelas Label beserta jumlahnya

Jenis Kelas	Jumlah
LARIS	41
TIDAK LARIS	25
Jumlah	66

Sumber: Hasil Penelitian (2023)

3.2. Tahap pre-processing

Setelah mendapatkan dataset, pastikan bahwa tidak ada missing value (nilai yang hilang) di dalamnya. Berikut adalah hasil pemeriksaan missing value pada dataset :

Tabel. 4. hasil pemeriksaan missing value pada dataset

Name	Type	Missing
LABEL	Binominal	0
SKU	Integer	0
DESC	Nominal	0
SUPPLIER	Nominal	0
GRADE	Binominal	0
STOCK ON HAND	Real	0
ON ORDER	Integer	0
COST	Real	0
RETAIL	Integer	0
QTY	Real	0
U/M	Binominal	0
TARGET	Integer	0
TOTAL SALES	Integer	0

Sumber: Hasil Penelitian (2023)

3.3 Implementasi Algoritma

Berikut adalah tabel yang berisi hasil klasifikasi yang diperoleh:

Tabel. 5. Hasil precision dan recall pada Decision Tree

Label	Precision	Recall
LARIS	100%	87,50%
TIDAK LARIS	83,33%	100%

Sumber: Hasil Penelitian (2023)

Pada Tabel 4 terlihat bahwa hasil evaluasi Decision Tree, terdapat nilai precision 100% dan recall 87,50% pada label LARIS sedangkan pada label TIDAK LARIS memiliki nilai precision 83,33% dan recall 100%.

Tabel. 6. Hasil Akurasi pada Naive Bayes

Label	Precision	Recall
LARIS	90,91%	62,50%
TIDAK LARIS	60,00%	90,00%

Sumber: Hasil Penelitian (2023)

Pada Tabel 5 terlihat bahwa hasil evaluasi Naive Bayes, terdapat nilai precision 90,91% dan recall 62,50% pada label LARIS sedangkan pada label TIDAK LARIS memiliki nilai precision 60,00% dan recall 90,00%.

Tabel. 7. Hasil Akurasi pada K-Nearest Neighbor

Label	Precision	Recall
LARIS	100%	81,25%
TIDAK LARIS	76,92%	100%

Sumber: Hasil Penelitian (2023)

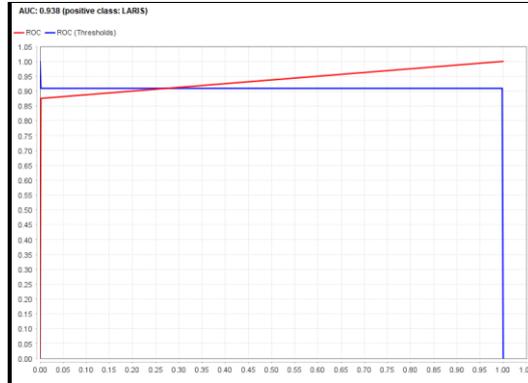
Pada Tabel 6 terlihat bahwa hasil evaluasi K-Nearest Neighbor, terdapat nilai precision 100% dan recall 81,25% pada label LARIS sedangkan pada label TIDAK LARIS memiliki nilai precision 76,92% dan recall 100%.

Tabel. 8. Perbandingan akurasi pada metode klasifikasi Decision Tree, Naive Bayes dan K-Nearest Neighbor.

Metode klasifikasi	Nilai akurasi
Decision Tree	92,31%
Naive Bayes	73,08%
K-Nearest Neighbor	88,46%

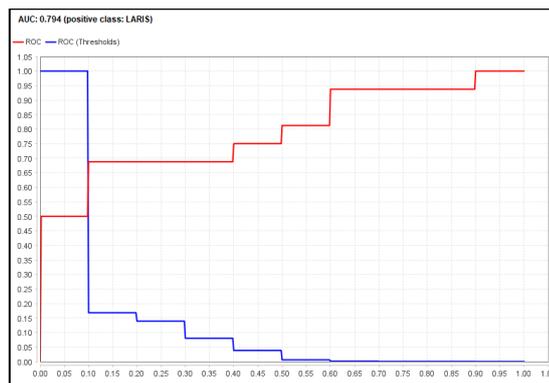
Sumber: Hasil Penelitian (2023)

Setelah membandingkan nilai presisi dan recall, nilai akurasi akan digunakan untuk mengevaluasi seberapa akurat klasifikasi yang digunakan. Berdasarkan hasil perhitungan akurasi pada ketiga metode klasifikasi, diperoleh hasil sebagai berikut: Decision Tree memiliki nilai akurasi tertinggi, yaitu 92,31%, diikuti oleh K-Nearest Neighbor dengan akurasi sebesar 88,46%, dan Naive Bayes dengan persentase akurasi sebesar 73,08%.



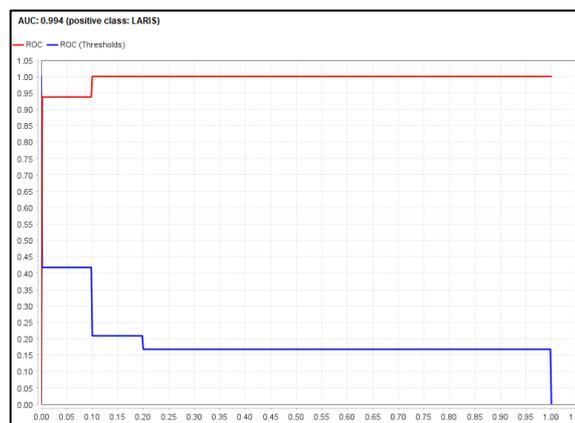
Sumber: Hasil Penelitian (2023)

Gambar. 2. Kurva AUC pada Decision Tree



Sumber: Hasil Penelitian (2023)

Gambar. 3. Kurva AUC pada Naive Bayes



Sumber: Hasil Penelitian (2023)

Gambar. 4. Kurva AUC pada K-Nearest Neighbor

Gambar grafik AUC diatas menunjukkan bahwa Decision Tree termasuk pada kategori Excellent dengan nilai AUC sebesar 93,8% pada kriteria AUC yang telah ditetapkan diatas. Sama halnya dengan K-Nearest

Neighbor, termasuk pada kategori Excellent dengan nilai AUC yang dihasilkan yaitu sebesar 99,4%. Sedangkan Naive Bayes memiliki nilai 79,4% pada AUC, dan termasuk pada kategori fair. Hal ini menunjukkan bahwa, Decision Tree, Naive Bayes dan K-Nearest Neighbor bisa digunakan untuk pengklasikasian dikarenakan memiliki nilai akurasi dan nilai AUC yang cukup tinggi.

KESIMPULAN

Tujuan dari penelitian ini adalah membandingkan tiga model algoritma dalam mengklasifikasikan data penjualan buah, yaitu Decision Tree, Naive Bayes, dan K-Nearest Neighbor. Setelah melakukan tahap preprocessing dan klasifikasi dataset, Decision Tree menunjukkan performa yang sangat baik dengan akurasi sebesar 92,31% dan AUC sebesar 93,8%, yang dikategorikan sebagai Excellent. Sementara itu, K-Nearest Neighbor mencapai akurasi sebesar 88,46% dan AUC sebesar 99,4%, juga termasuk dalam kategori Excellent. Namun, metode klasifikasi Naive Bayes memiliki akurasi sebesar 73,08% dan AUC sebesar 79,4%, yang masuk dalam kategori fair menurut kriteria AUC. Berdasarkan penjelasan tersebut, dapat disimpulkan bahwa Decision Tree merupakan algoritma terbaik untuk mengolah data penjualan buah di Foodmart Supermarket.

REFERENSI

- Alfiah, F., Pandhito, B. W., Sunarni, A. T., Muharam, D., & Matusin, P. R. (2018). Data Mining Systems to Determine Sales Trends and Quantity Forecast Using Association Rule and CRISP-DM Method. *International Journal of Engineering and Techniques*, 4(1), 186–192. <http://www.ijetjournal.org>
- Arifin, Z. (2019). Penerapan Metode Knn (K-Nearest Neighbor) Dalam Sistem Pendukung Keputusan Penerimaan Kip (Kartu Indonesia Pintar) Di Desa Pandean Berbasis Web Dan Mysql. *NJCA (Nusantara Journal of Computers and Its Applications)*, 4(1). <https://doi.org/10.36564/njca.v4i1.101>
- Cynthia, E. P., & Ismanto, E. (2018). Metode Decision Tree Algoritma C.45 Dalam Mengklasifikasi Data Penjualan Bisnis Gerai Makanan Cepat Saji. *Jurasik (Jurnal Riset Sistem Informasi Dan Teknik Informatika)*, 3(July), 1. <https://doi.org/10.30645/jurasik.v3i0.60>
- Damuri, A., Riyanto, U., Rusdianto, H., & Aminudin, M. (2021). Implementasi Data Mining dengan Algoritma Naive Bayes Untuk Klasifikasi Kelayakan Penerima Bantuan Sembako. *JURIKOM (Jurnal Riset Komputer)*, 8(6), 219. <https://doi.org/10.30865/jurikom.v8i6.3655>
- Mahmud Nawawi, H., Rahayu, S., Ja, M., Shidiq, far, Jaya Purnama, J., & Studi Ilmu Komputer STMIK Nusa Mandiri Jakarta www.nusamandiri.ac.id. (2019). Algoritma C4.5 Untuk Memprediksi Pengambilan Keputusan Memilih Deposito Berjangka. *Maret*, 16(1), 65. www.nusamandiri.ac.id
- Puspitasari, D., Ramanda, K., Supriyatna, A., Wahyudi, M., Sikumbang, E. D., & Sukmana, S. H. (2020). Comparison of Data Mining Algorithms Using Artificial Neural Networks (ANN) and Naive Bayes for Preterm Birth Prediction. *Journal of Physics: Conference Series*, 1641(1). <https://doi.org/10.1088/1742-6596/1641/1/012068>
- Sufia Maulida, Fikri Hamidy, A. D. W. (2020). Monitoring Aplikasi Menggunakan Dashboard Untuk Sistem Informasi Akuntansi Pembelian Dan Penjualan (Studi Kasus : Ud Apung). *Jurnal Tekno Kompak*, 14(1), 47. <https://doi.org/10.33365/jtk.v14i1.503>
- Sunyoto, D., & Mulyono, A. (2022). Manajemen Bisnis Retail. *Suparyanto Dan Rosad*, 5(3), 248–253.
- Widaningsih, S. (2019). Perbandingan Metode Data Mining Untuk Prediksi Nilai Dan Waktu Kelulusan Mahasiswa Prodi Teknik Informatika Dengan Algoritma C4.5, Naive Bayes, Knn Dan Svm. *Jurnal Tekno Insentif*, 13(1), 16–25. <https://doi.org/10.36787/jti.v13i1.78>
- Wirawan, C. (2020). Teknik Data Mining Menggunakan Algoritma Decision Tree C4.5 untuk Memprediksi Tingkat Kelulusan Tepat Waktu. *Applied Information System and Management (AISM)*, 3(1), 47–52. <https://doi.org/10.15408/aism.v3i1.13033>
- Yunita, D. (2017). Perbandingan Algoritma K-Nearest Neighbor dan Decision Tree untuk Penentuan Risiko Kredit Kepemilikan Mobil. *Jurnal Informatika Universitas Pamulang*, 2(2), 103. <https://doi.org/10.32493/informatika.v2i2.1512>