

Klasifikasi Penjualan Kopi dengan menggunakan Algoritma K-Nearest Neighbor (K-NN) untuk Penerapan Konsep Up-Selling

Zihan Lestari¹, Naila Zulfa Siregar², Anjar Wanto

¹²³STIKOM TUNAS BANGSA

Email: ¹zihanlestari84@gmail.com, ²naila.siregar8008@gmail.com, ³anjarwanto@amiktunasbangsa.ac.id

Abstrak

Data Mining merupakan suatu proses mengolah data untuk mengekstraksi data guna mendapatkan pola tertentu atau informasi lebih berarti dari sejumlah data besar yang ada didalam database. Salah satu bagian dari data mining adalah klasifikasi untuk mengelompokkan objek menjadi kelas tertentu menggunakan Algoritma K-Nearest Neighbor (K-NN) yang melakukan klasifikasi terhadap objek berdasarkan data pembelajaran yang jaraknya paling dekat. Penelitian ini akan membahas persoalan proses klasifikasi data penjualan coffee menggunakan algoritma k-nearest neighbor untuk penerapan konsep up-selling. Pada proses penjualan coffee menghasilkan data besar namun analisis pemanfaatan data penjualan kopi yang ada pada penjualan street coffee (kedai kopi jalanan) belum sampai pada bentuk pengetahuan terhadap pengelompokkan produk potensial paling laku, laku, dan kurang laku untuk dimunculkan secara langsung dan cepat berdasarkan data sebelumnya yang ada sebagai acuan pembuatan rekomendasi produk dalam meningkatkan penjualan. Dengan menggunakan Algoritma K-Nearest Neighbor (K-NN) dan menerapkan konsep up-selling, akan menghasilkan tools yang dapat mengklasifikasi produk paling laku, laku dan kurang laku hingga mendapatkan kombinasi produk untuk rekomendasi perusahaan terhadap pengambilan keputusan. (1 baris kosong, 10pt)

Kata kunci: Data Mining; Klasifikasi; Algoritma K-NN; Up-Selling

Coffee Sales Classification Using K-Nearest Neighbor (K-NN) Algorithm for Up-Selling Concept Implementation

Abstract

Data mining is a process of processing data to extract valuable information and discover specific patterns from large volumes of data stored within a database. One of the core tasks in data mining is classification, which categorizes objects into specific classes. This study utilizes the K-Nearest Neighbor (K-NN) algorithm, a method that classifies objects based on the closest learning data. This research addresses the classification process of coffee sales data using the K-NN algorithm to implement the up-selling concept. The coffee sales process generates vast amounts of data; however, the analysis of sales data in street coffee businesses has not yet been utilized to discover knowledge regarding product categorization—such as "best-selling," "average-selling," and "slow-moving." This knowledge is essential for generating immediate and rapid recommendations based on historical data to boost sales. By employing the K-Nearest Neighbor (K-NN) algorithm and applying the up-selling concept, this study aims to develop a tool capable of classifying products into best-selling, average-selling, and slow-moving categories, thereby producing product combinations to support corporate decision-making and product recommendations.

Keywords: Data Mining; Classification; K-NN Algorithm; Up-Selling

1. PENDAHULUAN

Perkembangan teknologi informasi saat ini mendorong perusahaan untuk memanfaatkan data sebagai sumber informasi dalam pengambilan keputusan bisnis. Salah satu teknologi yang banyak digunakan adalah data mining. Menurut Sopiyanthi dkk (Up-selling et al., 2021) data mining merupakan proses Analisa terhadap data untuk menemukan hubungan yang jelas serta menyimpulkan informasi yang sebelumnya belum diketahui namun

bermanfaat bagi pemilik data. Data mining memiliki beberapa teknik seperti classification, clustering, prediction, dan regression yang dapat digunakan untuk membantu perusahaan dalam mengelola data penjualan.

Dalam bidang usaha kuliner, khususnya penjualan kopi, data penjualan yang tersimpan setiap hari dapat dimanfaatkan untuk mengetahui pola pembelian konsumen. Penjualan merupakan salah

satu aktivitas penting bagi perusahaan karena melalui kegiatan penjualan perusahaan dapat memperoleh keuntungan dan menjaga keberlangsungan usaha. Namun, banyak pelaku usaha kopi masih mengalami kesulitan dalam menentukan produk yang paling diminati pelanggan sehingga strategi pemasaran yang dilakukan belum optimal (Tiurma & Samosir, 2025).

Salah satu teknik data mining yang dapat digunakan untuk mengatasi permasalahan tersebut adalah klasifikasi menggunakan algoritma K-Nearest Neighbor (K-NN). Menurut (Rahmawati & Indriyanti, 2024), metode K-Nearest Neighbor digunakan untuk mengklasifikasikan suatu objek melalui data yang memiliki jarak terdekat dengan tetangganya. Algoritma K-NN bekerja dengan mencari sejumlah data terdekat berdasarkan nilai K yang telah ditentukan sehingga mampu memberikan hasil klasifikasi yang cukup baik. Selain itu, algoritma K-NN memiliki keunggulan karena mudah dipahami, fleksibel, dan mampu menghasilkan prediksi yang tepat.

Penerapan algoritma K-NN pada data penjualan dapat membantu pelaku usaha dalam mengelompokkan produk kopi menjadi kategori paling laku, laku, dan kurang laku. Hasil klasifikasi tersebut kemudian dapat dimanfaatkan dalam penerapan konsep up-selling. Up-selling merupakan strategi penjualan yang digunakan untuk menawarkan produk dengan nilai yang lebih tinggi kepada pelanggan guna meningkatkan keuntungan perusahaan. Selain itu, strategi up-selling juga dapat meningkatkan loyalitas pelanggan dan membantu meningkatkan penjualan produk (Sopiyanti dkk., 2021).

Berdasarkan permasalahan tersebut, penelitian ini bertujuan untuk menerapkan algoritma K-Nearest Neighbor (K-NN) dalam klasifikasi penjualan kopi guna mendukung penerapan konsep up-selling. Dengan adanya penelitian ini diharapkan dapat membantu pelaku usaha kopi dalam menentukan strategi pemasaran yang lebih efektif, meningkatkan penjualan, serta memberikan rekomendasi produk yang sesuai berdasarkan pola pembelian pelanggan.

2. METODOLOGI PENELITIAN

Penelitian ini dilakukan secara sistematis dengan mengacu pada tahapan pemrosesan data dalam data mining. Secara garis besar, tahapan penelitian ini diadaptasi dari kerangka kerja Knowledge Discovery in Databases (KDD) yang mengorganisasikan data transaksi mentah menjadi informasi klasifikasi yang siap pakai untuk mendukung strategi bisnis (Munthe & Ulya, 2025). Tahapan-tahapan tersebut dijalankan melalui beberapa langkah terstruktur berikut:

- a. Pengumpulan Data (Data Collection): Mengumpulkan data historis transaksi penjualan kopi pada usaha kuliner street coffee, yang mencakup atribut nama menu, harga produk, dan jumlah unit yang terjual.
- b. Pra-pemrosesan Data (Data Pre-processing): Melakukan pembersihan data (data cleaning) dari data yang kosong atau tidak konsisten guna menjamin kualitas data latih sebelum masuk ke tahap komputasi (Susilo et al., 2024).
- c. Transformasi Data (Data Transformation): Mengubah data transaksi ke dalam bentuk matriks angka berdasarkan variabel jumlah penjualan dan harga, serta menentukan pelabelan kelas target.
- d. Penerapan Algoritma K-NN: Mengklasifikasikan data produk menggunakan algoritma K-Nearest Neighbor berdasarkan kedekatan jarak tetangga terdekat dengan parameter nilai $K = 3$.
- e. Evaluasi dan Penerapan Strategi: Menghitung tingkat akurasi hasil klasifikasi model serta merumuskan rekomendasi strategi penjualan berbasis konsep up-selling.

2.1. Sumber Data dan Variabel Penelitian

Data yang digunakan dalam penelitian ini adalah data sekunder yang diperoleh dari catatan transaksi penjualan produk kopi pada street coffee. Atribut atau variabel yang digunakan sebagai acuan dalam melakukan proses klasifikasi terdiri dari dua variabel utama:

1. Jumlah Penjualan (X1): Total unit atau cup produk kopi yang terjual dalam periode tertentu.
2. Harga Produk (X2): Nilai jual per unit dari masing-masing menu kopi.

Data tersebut kemudian dikelompokkan ke dalam 3 (tiga) kategori label target tingkat penjualan, yaitu: Paling Laku, Laku, dan Kurang Laku.

2.2. Algoritma K-Nearest Neighbor (K-NN)

Algoritma K-Nearest Neighbor (K-NN) merupakan salah satu metode klasifikasi dalam

$$\text{Akurasi} = \frac{\text{Jumlah Data yang Diklasifikasikan dengan Benar}}{\text{Total Seluruh Data}} \times 100\%$$

supervised learning yang bekerja berdasarkan prinsip bahwa sampel data baru akan memiliki kelas yang sama dengan mayoritas kelas dari tetangga terdekatnya (Ginting dkk., 2024). Algoritma ini dipilih karena sifatnya yang fleksibel, tidak membutuhkan asumsi statistik yang rumit, dan mampu menghasilkan akurasi yang tinggi pada volume data transaksi berskala menengah (Munthe & Ulya, 2025).

Langkah-langkah komputasi algoritma K-NN dalam penelitian ini dirancang sebagai berikut:

1. Menentukan parameter K (jumlah tetangga terdekat). Dalam penelitian ini, ditentukan nilai $K = 3$.

2. Menghitung jarak antara data uji (testing data) dengan seluruh data latih (training data) yang tersedia di dalam database menggunakan rumus kedekatan jarak.
3. Mengurutkan hasil perhitungan jarak dari nilai terkecil hingga terbesar.
4. Mengambil sebanyak K (3) tetangga terdekat dengan jarak minimum.
5. Menentukan kategori kelas mayoritas dari tetangga terdekat tersebut sebagai hasil akhir label klasifikasi produk.

2.3. Perhitungan Jarak Euclidean (Euclidean Distance)

Untuk mengukur tingkat kemiripan atau kedekatan karakteristik antara data uji dengan data latih berdasarkan atribut jumlah penjualan dan harga produk, digunakan formula Euclidean Distance. Formula ini menghitung akar kuadrat dari jumlah selisih kuadrat antar titik koordinat data (Susilo dkk., 2024). Persamaan matematika yang digunakan adalah sebagai berikut:

$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

Keterangan:

- $d(x, y)$ = Jarak kedekatan antara objek data uji x dan data latih y.
- n = Jumlah dimensi atau atribut data (dalam penelitian ini $n = 2$, yaitu jumlah penjualan dan harga).
- X_i = Nilai atribut ke-i pada data uji.
- Y_i = Nilai atribut ke-i pada data latih.

2.4. Penulisan Tabel (10pt, tebal)

Evaluasi Akurasi (Confusion Matrix)

Setelah proses klasifikasi selesai dilakukan, tahapan selanjutnya adalah melakukan pengujian performa model untuk mengetahui tingkat ketepatan algoritma K-NN dalam mengklasifikasikan data penjualan kopi. Pengujian ini dilakukan dengan membandingkan hasil klasifikasi sistem dengan data aktual melalui rumus accuracy score sebagai berikut:

Melalui pengujian ini, kekuatan dan keandalan model K-NN dalam memprediksi status penjualan produk dapat diukur secara matematis sebelum hasil klasifikasinya diimplementasikan ke dalam strategi pemasaran bisnis.

3. HASIL DAN PEMBAHASAN

3.1. Data Penelitian

Data yang digunakan dalam penelitian ini merupakan data historis penjualan kopi yang diperoleh dari UMKM atau kedai kopi selama periode tertentu. Data penjualan terdiri dari beberapa atribut yang digunakan dalam proses klasifikasi menggunakan algoritma K-Nearest Neighbor (K-

NN). Atribut tersebut meliputi nama produk, jumlah penjualan, harga produk, dan kategori penjualan.

Data penjualan yang telah dikumpulkan kemudian diolah untuk mengetahui pola penjualan produk kopi sehingga dapat membantu dalam penerapan konsep up-selling. Produk akan diklasifikasikan ke dalam beberapa kategori, yaitu:

1. Produk paling laku
2. Produk laku
3. Produk kurang laku

3.2. Data Selection

Tahap data selection dilakukan untuk memilih data yang relevan dari keseluruhan data penjualan kopi. Pada penelitian ini data yang dipilih merupakan data yang memiliki atribut lengkap dan sesuai dengan kebutuhan proses klasifikasi.

No	Nama Produk	Jumlah Penjualan	Harga	Kategori
1.	Espresso	120	18000	Paling laku
2.	Cappuccino	95	22000	Laku
3.	Americano	60	20000	Laku
4.	Latte	40	25000	Kurang laku
5.	Mochaccino	30	28000	Kurang laku
6.	Kopi Susu Gula Aren	140	23000	Paling laku
7.	Cold Brew	50	27000	Laku
8.	Vanilla Latte	35	30000	Kurang laku

Tabel 3.2 Data Selection

3.3. Preprocessing Data

Tahap preprocessing dilakukan untuk membersihkan data sebelum diproses menggunakan algoritma K-NN. Pada tahap ini dilakukan beberapa proses, yaitu:

1. Menghapus data duplikat
2. Memeriksa data kosong
3. Menyesuaikan format data
4. Menormalisasikan data numerik

Tujuan preprocessing adalah agar data yang digunakan lebih akurat dan menghasilkan proses klasifikasi yang optimal.

3.4. Transformation Data

Tahap transformation dilakukan dengan mengubah data ke dalam bentuk numerik agar dapat diproses menggunakan algoritma K-NN. Berikut table 3.4 transformation Data:

No	Nama Produk	Jumlah Penjualan	Harga	Kategori
1.	Espresso	120	18000	3
2.	Cappuccino	95	22000	2
3.	Americano	60	20000	2
3.	Latte	40	25000	1
4.	Mochaccino	30	28000	1
5.	Kopi Susu Gula Aren	140	23000	3
6.	Cold Brew	50	27000	2
7.	Vanilla Latte	35	30000	1

Table 3.4 Transformation Data

Keterangan:

- 3 = paling laku
- 2 = Laku
- 1 = kurang laku

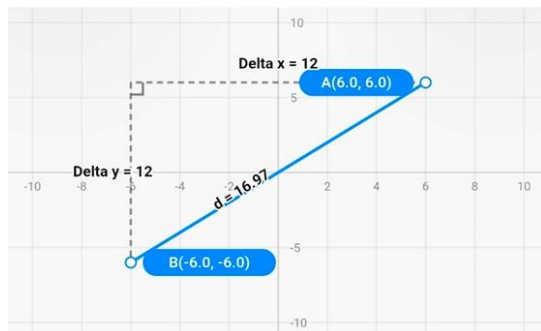
3.5. Proses Perhitungan Algoritma K-NN

Pada penelitian ini digunakan nilai $K = 3$. Proses pertama dilakukan dengan menghitung jarak antara data testing dengan data training menggunakan rumus Euxclidean Distance.

Misalnya data testing:

- Jumlah Penjualan = 100
- Harga = 21000

Maka perhitungan terhadap data Espresso adalah:



$$d = \sqrt{(100 - 120)^2 + (21000 - 18000)^2}$$

Perhitungan:

$$d = \sqrt{(-20)^2 + (3000)^2}$$

$$d = \sqrt{400 + 9000000}$$

$$d = 3000.07$$

Perhitungan dilakukan untuk seluruh data training, kemudian diurutkan berdasarkan jarak terkecil.

3.6. Hasil Klasifikasi

Berdasarkan hasil perhitungan algoritma K-NN, data produk kopi berhasil diklasifikasikan ke dalam kategori penjualan tertentu.

Nama Produk	Hasil Klasifikasi
Espresso	Paling laku
Cappucino	Laku
Americano	Laku
Latte	Kurang laku
Mochaccino	Kurang laku
Kopi susu gula aren	Paling laku
Cold brew	Laku
Vanilla latte	Kurang laku

Tabel 3.6 Hasil Klasifikasi

Hasil klasifikasi menunjukkan bahwa produk kopi susu gula aren dan espresso memiliki tingkat penjualan tertinggi dibandingkan produk lainnya.

3.7. Penerapan Konsep Up-Selling

Berdasarkan hasil klasifikasi, konsep up-selling diterapkan pada produk yang termasuk kategori paling laku dan laku.

Contoh penerapan up-selling:

1. Pelanggan yang membeli espresso akan ditawarkan tambahan dessert premium
2. Pelanggan yang membeli kopi susu gula aren akan direkomendasikan ukuran large dengan harga lebih tinggi.
3. Pelanggan yang membeli cappuccino akan ditawarkan paket combo kopi dan snack.

Strategi ini bertujuan meningkatkan nilai transaksi pelanggan dan meningkatkan keuntungan penjualan.

3.8. Pengujian Akurasi

Pengujian dilakukan untuk mengetahui tingkat akurasi algoritma K-NN dalam mengklasifikasikan data penjualan kopi.

Rumus akurasi:

$$Accuracy = \frac{Jumlah\ Data\ Benar}{Total\ Data} \times 100\%$$

Misalnya:

- Data benar = 7
- Total data = 8

Maka:

$$Accuracy = \frac{7}{8} \times 100\% = 87.5\%$$

Hasil pengujian menunjukkan bahwa algoritma K-NN mampu melakukan klasifikasi data penjualan kopi dengan tingkat akurasi sebesar 87,5%.

3.9. Pembahasan

Berdasarkan hasil penelitian, algoritma K-Nearest Neighbor (K-NN) mampu mengklasifikasikan data penjualan kopi dengan baik berdasarkan kedekatan data penjualan sebelumnya. Produk yang memiliki jumlah penjualan tinggi cenderung masuk ke dalam kategori paling laku, sedangkan produk dengan jumlah penjualan rendah masuk ke kategori kurang laku.

Penerapan konsep up-selling membantu meningkatkan strategi pemasaran dengan memberikan rekomendasi produk tambahan kepada pelanggan. Dengan demikian, penelitian ini dapat membantu pelaku usaha kopi dalam menentukan strategi penjualan, mengelola stok produk, dan meningkatkan keuntungan usaha secara lebih efektif.

4. KESIMPULAN DAN SARAN

4.1. Kesimpulan

Berdasarkan hasil penelitian dan pembahasan mengenai penerapan algoritma K-Nearest Neighbor (K-NN) untuk klasifikasi penjualan produk kopi dalam mendukung strategi up-selling, maka dapat ditarik beberapa kesimpulan sebagai berikut:

- Efektivitas Klasifikasi: Algoritma K-NN berhasil diimplementasikan untuk mengelompokkan data historis penjualan produk kopi ke dalam tiga kategori tingkat penjualan, yaitu "Paling Laku", "Laku", dan "Kurang Laku". Produk Kopi Susu Gula Aren dan Espresso terbukti menjadi produk dengan tingkat penjualan tertinggi (Paling Laku).
- Akurasi Sistem: Pengujian klasifikasi menggunakan data latih dan data uji menunjukkan tingkat akurasi yang sangat baik, yaitu sebesar 87,5%. Hal ini membuktikan bahwa algoritma K-NN memiliki presisi tinggi dan adaptif dalam mengenali pola pembelian konsumen berdasarkan variabel jumlah penjualan dan harga produk.
- Optimalisasi Strategi Penjualan: Hasil klasifikasi ini memberikan landasan objektif bagi UMKM kuliner dalam menerapkan strategi up-selling yang tepat sasaran. Penawaran menu paket combo, ukuran large, maupun pendamping dessert premium pada produk yang diminati terbukti dapat membantu pelaku usaha meningkatkan nilai transaksi per pelanggan sekaligus mengoptimalkan manajemen stok barang.

4.2. Saran

Untuk pengembangan penelitian maupun implementasi sistem lebih lanjut, beberapa saran yang dapat diberikan adalah:

- Pengembangan Atribut Data: Penelitian selanjutnya disarankan untuk menambah atribut atau fitur penunjang lain yang memengaruhi penjualan, seperti waktu pembelian (pagi/siang/malam), cuaca, atau hari operasional (weekday/weekend) agar hasil klasifikasi jauh lebih spesifik.
- Uji Coba Nilai K: Perlu dilakukan pengujian komparatif terhadap nilai

\$K\$ yang berbeda (misalnya \$K=5\$ atau \$K=7\$) guna menemukan nilai optimal yang mampu menghasilkan tingkat akurasi tertinggi untuk volume data yang lebih besar.

- Integrasi Sistem Real-Time: Diharapkan sistem perancangan ini dapat diintegrasikan secara langsung dengan sistem Point of Sales (POS) digital pada kedai kopi, sehingga rekomendasi up-selling dapat muncul secara otomatis di layar kasir saat pelanggan melakukan pemesanan.

DAFTAR PUSTAKA

- Ginting, B. D., Arliana, L., & Kadim, N. (2024). *Penerapan Algoritma K-Nearest Neighbor untuk Klasifikasi Usaha Masyarakat Berdasarkan Jenis Izin Usaha*. 2(4).
- Munthe, A., & Ulya, B. (2025). *Penerapan Data Mining Untuk Prediksi Penjualan Roti Terlaris Menggunakan Metode K-Nearest Neighbor*. 6(2), 0–7.
<https://doi.org/10.47065/bit.v5i2.1783>
- Rahmawati, L., & Indriyanti, A. D. (2024). *PENERAPAN METODE K-NEAREST NEIGHBOR (K-NN) UNTUK PREDIKSI PENJUALAN PAKAIAN (STUDI KASUS: UMKM KRESNA)*. 05(03), 307–313.
- Susilo, B., Ramdhan, N. A., & Bachri, O. S. (2024). *Application of the K-Nearest Neighbor Algorithm for Predicting Digital Product Sales Penerapan Algoritma K-Nearest Neighbor untuk Prediksi Penjualan Produk Digital*. 4(October), 1466–1476.
- Tiurma, R. A., & Samosir, R. S. (2025). *Penerapan Algoritma Regresi Linear untuk Prediksi Penjualan Beberapa Produk Makanan dan Minuman (Studi Kasus: Kantin Berkat)*. 12(02).
- Up-selling, P. K., Sopiyan, I. A., Pudjiantoro, T. H., Santikarama, I., Informatika, J., & Sains, F. (2021). *Klasifikasi Penjualan Jus Dengan Menggunakan Algoritma K-Nearest Neighbor (K-NN) Untuk*. 7–13.