

---

## Implementasi Algoritma K-NEAREST NEIGHBOR Untuk Meningkatkan Penjualan Produk UMKM Resiko Skala Kecil

Tintin Chandra  
Fakultas Ilmu Sains dan Teknologi  
Universitas IBBI  
Jalan Sei Deli No. 18 Medan 20114  
E-mail : [tinuhnsbm@gmail.com](mailto:tinuhnsbm@gmail.com)

---

### Abstrak

Produk yang beredar di masyarakat beraneka ragam, mulai dari kebutuhan primer, sekunder sampai dengan kebutuhan tersier tersedia. Aneka ragam produk mulai dari yang umumnya digunakan sampai dengan hiasan juga tersedia. Keanekaragaman produk dan variasinya menjadikan pelaku UMKM menjadi kesulitan dalam menyediakan semua produk yang baik yang sejenis maupun yang lainnya. Produk yang tersedia dengan keanekaragaman tentu membuat pembeli menjadi banyak alternatif dalam memenuhi kebutuhan dan keperluan untuk sehari-hari sampai dengan gaya. Pelaku UMKM tentu harus melihat dengan jeli pemilihan produk yang akan dijual sehingga nantinya modal tidak tertahan pada produk terutama produk yang sepi peminat. Walaupun demikian ada produk yang banyak dicari pembeli karena harga produk tinggi sehingga banyak pelaku UMKM tidak menjualnya, sehingga perlu dipertimbangan untuk menyediakan stok. Penjualan produk tergantung kepada permintaan pasar dimana kadang harga tidak menjadi patokan. Pelaku UMKM dalam melakukan kegiatan bisnis tidak hanya mengandalkan kegiatan bisnis konvensional saja, karena pembeli lebih memilih pembelian secara online dan apabila sesuai kebutuhan, maka pembeli tidak ragu untuk datang ke tempat pelaku UMKM melakukan kegiatan bisnis. Hal ini dengan sendirinya akan meningkatkan penjualan produk berdasarkan kategori dan permintaan pasar.

**Kata Kunci:** Algoritma K-Nearest Neighbor; Penjualan; Data Mining; UMKM.

### 1. Pendahuluan

Produk yang ada di masyarakat tersedia beraneka ragam, mulai dari kebutuhan primer, sekunder sampai dengan kebutuhan tersier tersedia. Produk merupakan barang fisik yang disesuaikan dengan kebutuhan dan keperluan. Produk dapat berupa peralatan rumah tangga, peralatan dapur, part kendaraan, peralatan perbengkelan, dan lain-lain. Produk yang beredar tentunya disesuaikan dengan kebutuhan dan keperluan dari pembeli. Walaupun demikian ada beberapa produk yang dibeli sebagai koleksi karena hobi dimana produk tersebut merupakan barang tersier. Ada beberapa produk yang telah bergeser dari kebutuhan sekunder menjadi kebutuhan primer tergantung pada kebutuhan pribadi pembeli. Beberapa contoh lampu penerangan yang dulunya merupakan barang sekunder menjadi barang primer, alat komunikasi yang dulunya merupakan barang tersier menjadi barang sekunder dimana menjadi kebutuhan sekunder yang harus dimiliki di era teknologi 4.0. Pelaku UMKM yang monoton atau masih berada di zona aman pelan-pelan akan tergeser oleh pelaku UMKM yang modern yang lebih banyak memanfaatkan teknologi dalam pencapaian target penjualan produk.

Pelaku UMKM dalam melakukan transaksi jual beli tentu memiliki keterbatasan modal, apabila modal di fokuskan pada beberapa produk akan membuat pelaku UMKM sulit untuk berkembang dengan baik. Keterbatasan modal membuat pelaku UMKM harus jeli melihat produk mana yang laris di pasar dan peminat dari pembeli. Beberapa pelaku UMKM sembako juga mengalami pertumbuhan penjualan yang sangat lambat karena kurang jeli melihat peminat dan kondisi pembeli dari masyarakat. Pelaku UMKM terutama resiko kecil tentu harus mempertimbangan dalam penyediaan persediaan barang untuk dijual dan menjaga arus kas tetap aman dan bisa bertahan dalam menghadapi persaingan baik secara offline maupun secara online.

Masalah umumnya adalah sulitnya memenuhi permintaan pasar dengan aneka ragam produk dengan fungsi yang hamper sama mengakibatkan persediaan tidak dapat diprediksi mana yang perlu stok lebih dan mana yang stok terbatas, selain itu waktu pengiriman dari supplier yang kadang terkendala di perjalanan. Produk yang telah sampai perlu diperiksa sebelum dijual kepada pembeli dimana kadang produk yang datang tidak sesuai atau cacat atau ada kendala pada produk tersebut. Produksi produk kadang juga mengalami terkendala di supplier yang mengakibatkan pelaku UMKM menjadi kesulitan dan memenuhi permintaan pembeli yang beraneka ragam, belum lagi ada kendala internal di supplier. Pelaku UMKM perlu melakukan perhitungan yang matang dalam persediaan produk untuk dijual agar target penjualan bisa meningkat dengan tetap menjaga ketersediaan produk, sehingga dapat mengurangi penumpukan produk yang kurang peminat di pasar. Produk yang menumpuk di gudang selain menahan modal juga mengalami penyusutan nilai dan beban sewa penyimpanan.

Pelaku UMKM resiko rendah melakukan mencatat transaksi penjualan dan menyimpan barang dalam kartu stok dan jurnal penjualan sehingga informasi tidak kurang akurat dan tidak teratur. UMKM resiko rendah masih menggunakan pencatatan secara konvensional dan pengelolaan persediaan produk masih manual melalui pengecekan konvensional, kadang-kadang menyebabkan selisih persediaan atas transaksi produk yang terjual karena produk yang terjual tidak tercatat menyebabkan pencatatan persediaan menjadi tidak akurat dan tidak tepat waktu. Produk yang tersedia di gudang kadang bisa mengalami overstock karena produk masih ada tetapi di pencatatan sudah habis. Penentuan jumlah stok

produk masih dilakukan dengan cara konvensional dari awal sehingga menimbulkan banyak kesulitan dan penumpukan produk di gudang. Overstock mengakibatkan pengelolaan arus cash menjadi terkendala dalam mempertahankan dan meningkatkan target penjualan produk, dengan menerapkan aplikasi yang ada dalam pengelolaan baik data maupun persediaan diharapkan dapat mengatasi permasalahan yang ada, dimana aplikasi dan pengelompokan data menggunakan pendekatan data mining.

Penggunaan data mining yang sering diterapkan dalam pengelompokan data penjualan adalah K-Nearest Neighbor. Data yang dikumpulkan digunakan untuk menganalisis data dalam jumlah transaksi untuk mencari informasi atau hubungan yang jelas dan menjadi pertimbangan dalam membuat keputusan apa yang dibutuhkan oleh pelaku UMKM dalam persediaan produk agar arus cash tetap terjaga. Penerapan data mining memiliki tiga tujuan utama meliputi identifikasi, validasi, dan penemuan. Untuk pengelompokan produk yang bisa menggunakan clustering tergantung dari produk yang dijual. Metode KNN merupakan algoritma dalam menganalisa data yang digunakan untuk cluster atau perhitungan data menggunakan algoritma yang tersedia, walaupun memiliki kelemahan jika nilai k kecil maka hasil kurang teliti, penggunaan metode KNN masih bisa jadi pilihan algoritma yang lebih baik, dengan perhitungan data penjualan yang direkam menjadi big data yang selanjutnya dianalisa dan nantinya bisa diterapkan untuk meningkatkan target penjualan dan menaikkan pendapatan.

Permasalahan yang telah dijelaskan, maka peneliti mencoba untuk mengimplementasi algoritma data mining untuk menyelesaikan permasalahan dengan menganalisis data produk yang terjual menggunakan metode K-Nearest Neighbor, Algoritma KNN merupakan salah satu dari banyaknya algoritma klasifikasi yang memetakan jarak terdekat, selain itu K-Nearest Neighbor merupakan metode sangat mudah diterapkan secara efisien dengan nilai akurasi yang lebih baik.

## 2. Landasan Teori

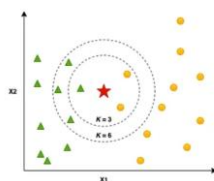
### Pengertian Penjualan

Penjualan dapat diartikan dengan proses pemenuhan kebutuhan penjual dan pembeli secara tunai maupun kredit. Hal ini menjadi salah satu tolak ukur apakah kegiatan bisnis berjalan lancar atau tidak. Aktivitas penjualan memiliki angka yang tinggi, maka pelanggan banyak yang membutuhkan barang atau jasa yang dijual apabila tidak maka ada kesalahan terhadap produk atau mungkin kegiatan bisnis memasuki pasar yang tidak tepat. Pengertian penjualan dapat diartikan kegiatan jual beli dijalankan oleh dua belah pihak atau lebih dengan alat pembayaran yang sah dengan tujuan tentunya mendatangkan keuntungan dari produk atau barang yang dijual. Pelaksanaan jual beli tidak dapat dilakukan tanpa adanya kontribusi dari pelaku yang bekerja, seperti pedagang, agen, dan tenaga pemasaran. Penjualan juga bisa menjadi tolak ukur untuk evaluasi bisnis karena dari informasi ini, bisa diketahui data yang akurat tentang kondisi produk atau jasa terhadap pasar.

Kegiatan penjualan biasanya memiliki garis besar meliputi melibatkan minimal dua pihak atau lebih, yaitu penjual dan pembeli, memiliki objek yang diperjualbelikan, bisa berupa barang maupun jasa, menggunakan alat pembayaran yang sah (tunai, transfer, kartu, atau instrumen pembayaran lain yang disepakati), menghasilkan kesepakatan terkait harga, jumlah, dan syarat penyerahan barang/jasa, dan menyebabkan perpindahan hak kepemilikan barang atau jasa dari penjual ke pembeli.

### Algoritma K-Nearest Neighbor

Nearest Neighbor atau K-Nearest Neighbor (kNN) merupakan salah satu algoritme klasifikasi hasil dari perhitungan data dalam data mining yang memanfaatkan data terdekat untuk melakukan prediksi pada data baru yang belum dikenal (data uji). Algoritme K-Nearest Neighbor (kNN) bekerja dengan cara mengelompokkan sejumlah tetangga terdekat dari data uji dan menentukan kelas data uji tersebut berdasarkan mayoritas kelas dari tetangga terdekat yang ditemukan. Nearest Neighbor dapat [digunakan diterapkan](#) untuk [menanganimengolah](#) berbagai jenis data, baik data numerik maupun [kategorikal-kategori](#). Pada data [kategorikal-kategori](#), perhitungan [jarak](#) jarak perbedaan atau kesamaan tidak dapat dihitung menggunakan operasi matematik karena data kategori tidak dapat diukur seperti pada data numerik. Nearest Neighbor lebih efektif pada data dengan dimensi yang rendah atau sedang. Algoritme K-Nearest Neighbor (kNN) juga lebih efektif untuk dataset yang relatif kecil hingga sedang, karena semakin besar jumlah data yang diproses maka membutuhkan waktu klasifikasi yang lebih lama. K-Nearest Neighbor (kNN) juga memiliki kekurangan antara lain sensitif terhadap nilai anomaly dan ketidakseimbangan kelas. Ilustrasi algoritma K-Nearest Neighbor (kNN) dapat dilihat pada gambar 1.



**Gambar 1.** Ilustrasi algoritma K-Nearest Neighbor (kNN)

Pada gambar 1. menggunakan dua nilai k, yaitu k=3 dan k=6. K-Nearest Neighbor (kNN) mencari sejumlah tetangga terdekat dari data uji yang belum dikenal dan menentukan batas kelasnya berdasarkan dominan kelas tetangga. Nilai k=3,

artinya data yang di target akan mencari tiga tetangga terdekat dari data yang diuji. Pada gambar 1., terlihat dua dari tiga tetangga terdekat tergolong ke dalam kelas positif dan satu tetangga terdekat tergolong ke dalam kelas negatif, karena dominan kelas tetangga terdekat adalah kelas positif, maka data targer akan diklasifikasikan sebagai kelas positif.

### Proses Algoritma K-NN

K-Nearest Neighbor (NN) merupakan salah satu algoritma machine learning data mining yang lebih sederhana namun cukup efektif dalam melakukan klasifikasi data. Algoritma K-Nearest Neighbour berdasarkan pada konsep “kesamaan” antara data, dapat dihitung dengan jarak Euclidean, Manhattan atau Minkowski distance. Algoritma K-NN mencari data latih yang paling mirip dengan data target yang akan diklasifikasikan, dan memberikan label kelas yang sama dengan data target tersebut. Cara kerja K-Nearest Neighbor adalah:

#### 1. Menentukan jumlah tetangga terdekat (K)

Menentukan nilai K untuk algoritme Nearest Neighbor. Penentuan nilai K tidak memiliki rumus yang pasti, ada beberapa cara yang dapat dipertimbangkan, contoh jika jumlah kelas adalah genap maka disarankan untuk memilih nilai K ganjil, dan jumlah kelas adalah ganjil, disarankan memilih nilai K genap. Alasannya adalah jika kita memilih nilai K yang genap pada kasus dengan dua kelas, misalnya kelas A dan B, maka ada kemungkinan hasil dari 4 tetangga terdekat terdiri dari 2 kelas A dan 2 kelas B sehingga hasilnya akan sama, yaitu A atau B. Namun, jika kita memilih nilai K yang ganjil, misalnya 3 atau 5, maka ada tetangga terdekat yang lebih banyak jumlahnya, sehingga hasilnya akan lebih variatif.

#### 2. Hitung jarak antara data uji dengan semua data latih

Data nilai K, selanjutnya dalam algoritm Nearest Neighbor adalah menghitung jarak antara data fokus dengan semua data latih. Jarak latih dapat dihitung menggunakan berbagai metode, seperti Euclidean distance, Manhattan distance, Minkowski distance, atau metode lainnya. Perhitungan jarak, setiap fitur pada data tetap diperhitungkan dan diakumulasikan menjadi satu skor jarak. Semakin kecil jarak skor, maka kedua data makin memiliki kemiripan. Proses ini diulang untuk semua data latih yang tersedia.

#### 3. Urutkan data latih berdasarkan jarak dari yang terdekat ke yang terjauh dengan data latih

Jarak antara data latih dengan data fokus menggunakan suatu metrik, langkah berikutnya adalah mengurutkan data latih berdasarkan jarak dari yang terdekat ke yang terjauh dengan data fokus. Data latih yang paling dekat dengan data fokus akan ditempatkan di bagian atas daftar, sedangkan data latih yang paling jauh akan ditempatkan di bagian bawah daftar. Untuk memudahkan dalam mencari tetangga terdekat dari data fokus pada langkah berikutnya. Kelebihan mengurutkan data latih berdasarkan jarak, dapat memastikan bahwa k-nearest neighbor yang dipilih benar-benar merupakan tetangga terdekat dari data uji.

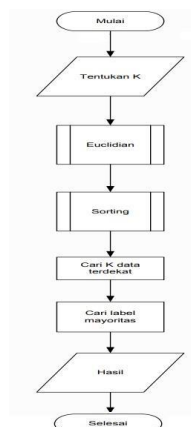
#### 4. Pilih K data dengan jarak terdekat dari data uji

Data latih diurutkan berdasarkan jarak terdekat dengan data fokus, maka langkah berikutnya pilih K data dengan jarak terkecil. Banyaknya K di sini berdasarkan jumlah K yang telah ditentukan di awal.

#### 5. Hitung frekuensi masing-masing kategori pada K data yang telah dipilih

Perhitungan frekuensi setiap kategori akan dihitung berapa kali muncul pada K data tersebut dan kemudian dicatat jumlahnya. Untuk mengetahui frekuensi masing-masing kategori, dapat menentukan kategori mana yang paling banyak muncul dan kemudian menjadikannya sebagai hasil prediksi untuk data yang sedang dicari kategorinya.

Dimana matriks *distance* adalah jarak skala dari kedua vektor  $x$  dan  $y$  dari matriks dengan ukuran  $n$  dimensi. Pada fase training, algoritma ini hanya melakukan penyimpanan vektor-vektor fitur dan klasifikasi data. Pada fase klasifikasi, fitur-fitur yang sama dihitung untuk testing data (yang klasifikasinya tidak diketahui). Jarak dari vektor baru yang ini terhadap seluruh vector sample dihitung dan sejumlah  $k$  buah yang paling dekat diambil. Alur dalam K-NN dapat dilihat pada gambar 2.



**Gambar 2** Alur K-Nearest Neighbour**Pengujian Menggunakan Cross Validation**

Cross validation bekerja dengan memecah dataset menjadi beberapa bagian, kemudian dilakukan pengujian pada bagian yang tidak digunakan untuk dataset yang tidak dikenal. Proses dilakukan beberapa kali dengan kombinasi subset yang berbeda, kemudian lakukan performa dirata-ratakan untuk mendapatkan estimasi yang lebih stabil dan mengurangi risiko overfitting. Ada beberapa metode dalam cross validation antara lain:

## 1. Hold-out validation

Hold-out adalah bentuk paling sederhana. Dataset dibagi sekali menjadi train dan test (dengan 70% pelatihan dan 30% pengujian). Metode hold-out validation cepat dan mudah serta cocok untuk dataset besar, kurang efektif untuk pembagian dataset karena hanya menggunakan satu pembagian acak sehingga estimasi performa bisa memiliki performa yang tidak stabil.

## 2. K-fold cross validation

Pada k-fold cross validation, data dibagi menjadi kelipatan k (k-fold) berukuran hampir sama, kemudian dilakukan proses pelatihan dan pengujian berulang kali, setiap kali satu fold menjadi test dan sisanya train. Nilai k=5 atau k=10 sering dijadikan acuan karena memberi titik temu yang lebih baik antara bias dan varians estimasi performa tanpa beban komputasi berlebihan. Estimasi performanya lebih stabil dan tidak terlalu bergantung pada pembagian data.

## 3. Stratified k-fold cross validation

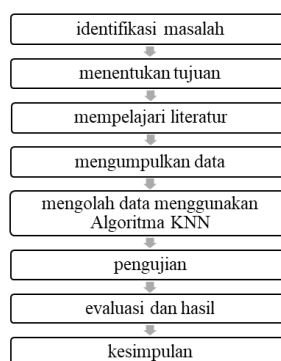
Stratified k-fold digunakan untuk memastikan bahwa proporsi kelas di setiap fold sama dengan proporsi kelas di dataset awal, sehingga sangat penting untuk dataset yang imbalanced. Keseimbangan data membuat model dievaluasi secara lebih adil, terutama untuk kelas minoritas yang rentan terabaikan jika pembagian dilakukan secara acak biasa.

## 4. Leave-one-out cross validation (LOOCV)

Setiap iterasi menggunakan satu sampel untuk testing dan sisanya digunakan untuk training. Jika dataset punya 100 sampel, maka model akan dilatih sebanyak 100 kali. LOOCV memanfaatkan hampir semua data secara maksimal untuk pelatihan sehingga diperoleh hasil bias rendah.

**3. Metode Penelitian**

Pelaksanaan penelitian memerlukan kerangka kerja yang terencana dan terukur agar pelaksanaan penelitian dilaksanakan secara terurut dan terukur serta membantu peneliti dalam menyelesaikan penelitian Implementasi Algoritma K-NEAREST NEIGHBOR sesuai alur yang telah dibuat. Kerangka penelitian pada penelitian implementasi algoritma KNN yaitu identifikasi masalah dari penelitian, menemukan permasalahan penelitian, menentukan tujuan, mengumpulkan dan menganalisis literatur baik dari buku data mining maupun dari jurnal data mining, pengumpulan data yang diperoleh dari UMKM resiko rendah, mengolah data menggunakan metode KNN, pengujian dan evaluasi serta membuat Kesimpulan dari penelitian implementasi algoritma K-Nearest Neighbour seperti pada gambar 3.

**Gambar 3.** Kerangka Penelitian Implmentasi K-Nearest Neighbour

Tahap awal peneliti melakukan observasi ke UMKM resiko skala kecil dan melakukan wawancara untuk menemukan masalah yang terjadi dalam kegiatan transaksi jual beli. Data yang difokuskan peneliti terutama pada transaksi penjualan. Data penjualan yang dikumpulkan nantinya ditindaklanjuti dengan melakukan analisis lebih lanjut untuk menentukan produk atau barang yang lebih diminati di pasar atau pembeli. Tujuan penelitian yang dilakukan agar dapat memberikan solusi terbaik kepada UMKM resiko skala kecil dalam pengelompokkan produk atau barang yang laku atau tidak sehingga dapat meminimalkan persediaan produk di gudang sehingga tidak membebani arus kas.

Setelah peneliti mendapatkan sumber informasi primer dari pelaku UMKM dengan tujuan membantu pelaku UMKM, maka peneliti mengumpulkan kajian kepustakaan dengan membaca beberapa buku teks data mining serta beberapa jurnal data mining yang memiliki keterkaitan dengan algoritma KNN. Kajian kepustakaan dimaksud agar hasil penelitian memiliki dasar untuk menganalisis agar target penjualan pada usaha UMKM resiko skala kecil dapat tercapai. Peneliti melakukan pengumpulan data dengan melihat transaksi penjualan karena pelaku UMKM menjual aneka produk maka peneliti mengumpulkan data pada beberapa produk dari kumpulan transaksi secara satu periode.

Data yang terkumpul di-rekam menggunakan alat bantu Microsoft Excel dan disimpan dalam format CSV yang nantinya diolah menggunakan aplikasi WEKA. Untuk proses perhitungan menggunakan rumus  $\sqrt{((x_i-A)^2+(x_i-A)^2)+(x_i-A)^2}$ ). Data yang terekam dilakukan seleksi anomoli agar hasil analisis dapat memiliki hasil yang lebih baik. Hasil perhitungan dengan algoritma KNN untuk mengelompokkan kategori produk yang laku dan kurang laku. Hasil perhitungan yang diperoleh dari algoritma KNN dilakukan menggunakan untuk memvalidasi untuk mengetahui sejauhmana hasil berfungsi sebagaimana mestinya dengan menggunakan bantuan aplikasi WEKA.

#### 4. Hasil Penelitian

Data yang dikumpulkan peneliti direkam per bulan berdasarkan kategori item dengan kuantitas terjual periode caturwulan (4 bulan) dimulai dari bulan Nopember, Desember tahun 2025 dan bulan Januari, Februari 2026. Data yang dikumpulkan kemudian di kelompokkan 2 kategori Laku dan Kurang Laku. Peneliti menggunakan perhitungan dengan algoritma K-Neighbour Nearest yaitu:

- a. Menentukan nilai konstanta k
- b. Data yang telah divalidasi kemudian dilakukan tahap normalisasi untuk semua data, tujuannya untuk menyamakan batasan nilai yang sama pada data yang direkam antara 0 sampai 1.
- c. Menggunakan rumus Euclidean untuk menghitung jarak terdekat dengan m.
- d. Mengurutkan data dari nilai terkecil.
- e. Menentukan data dari nilai terbananyak.

Peneliti melakukan analisis menggunakan algoritma K-means dalam bagian dari proses pembuatan produk. Hasil analisis dari hasil penelitian Algoritma KNN akan digunakan sebagai informasi dalam tahapan pembuatan produk. Tahapan dalam melakukan analisis menggunakan algoritma KNN.

##### 1. Analisa data survei

Peneliti telah melakukan survey di UMKM resiko skala kecil dan memperoleh informasi bahwa belum terdapat sistem pengelompokkan produk yang sering terjual/laku. Peneliti melakukan analisis berdasarkan permasalahan yang dialami pelaku UMKM terhadap objek penelitian dengan observasi dan wawancara dengan pelaku UMKM resiko skalo kecil untuk mendapatkan informasi lebih akurat mengenai masalah yang yang timbul pada usaha pelaku UMKM.

##### 2. Analisa hasil wawancara

Hasil yang diperoleh peneliti dari observasi dan wawancara terhadap pelaku UMKM resiko skala kecil, peneliti memperoleh berbagai macam informasi tentang masalah dalam pengelompokan produk yang beraneka ragam pada usaha UMKM resiko skala kecil. Peneliti mengajukan beberapa pertanyaan yang berhubungan dengan masalah penjualan dan pengelompokan produk yang ada pada usaha UMKM resiko skala kecil.

##### 3. Analisa data mining

Data yang diperoleh dari wawancara dan peninjuaun dikumpulkan menjadi satu kesatuan kemudian direkap dan disimpan pada Microsft Excel. Peneliti melakukan analisis dari data yang diperoleh berupa daftar penjualan produk di UMKM resiko skala kecil. Peneliti menggunakan daftar penjualan produk untuk dipilah menggunakan kriteria yaitu: total penjualan produk dalam 5 bulan > 30 item memiliki kriteria laris dan total penjualan produk dalam 5 bulan < 30 item memiliki kriteria kurang laris.

Analisa KNN klasifikasi

Data yang telah direkap kemudian dihitung menggunakan Perhitungan Euclidean distance dengan cara yaitu:

- a. Menentukan nilai K yang di digunakan adalah 8.
- b. Perhitungan Euclidean distance memanfaatkan fungsi Microsoft Excel yaitu  $=SQRT((D2-SD\$12)^2+(E2-SD\$12)^2+(F2-SD\$12)^2+(G2-SD\$12)^2+(H2-SD\$12)^2)$ . Nilai D2, E2, F2, G2 dan H2 merupakan nilai data yang diuji dan SD\$12 adalah angka dari data data training yaitu 30.
- c. Data yang digunakan pada perhitungan Euclidean distance dapat dilihat pada tabel 1.
- d. Hasil yang diperoleh pada rumus bagian b yaitu

T1 = 21.72556098	T2 = 17.97220076
T3 = 12.88409873	T4 = 20.1246118
T5 = 13.19090596	T6 = 57.5934024
T7 = 75.40557009	T8 = 62.77738446
T9 = 26.68332813	T10 = 72.6704892

**Tabel 1.** Data Set Awal

No	Jenis Item	Nov	Des	Jan	Feb	Mar	Kategori
1	Peralatan Listik	38	20	18	22	20	Kurang Laku
2	Peralatan Dapur	25	32	17	20	25	Kurang Laku
3	Obat Serangga	35	30	34	25	20	Kurang Laku

4	Peralatan Pertukangan	17	24	20	22	24	Kurang Laku
5	Perlengkapan Rumah Tangga	36	26	23	27	22	Kurang Laku
6	Peralatan bengkel	68	42	57	48	56	Laku
7	Lampu	66	54	61	72	63	Laku
8	Pewangi	46	44	68	56	67	Laku
9	Cat	16	14	18	20	26	Kurang Laku
10	Aksesoris	56	68	59	66	62	Laku

Tabel 1. merupakan data set awal yang telah dikumpulkan dan direkap berdasarkan kelompok seperti peralatan listrik (stop kontak, kabel, lakban, dan lain-lain), peralatan dapur (panci, gelas, sendok, kual, dan lain-lain) serta produk yang di kelompokkan atau disatukan menjadi jenis item

**Tabel 2. Menghitung Jarak Eucludian Distance**

No	Jenis Item	Jarak
T1	1 Peralatan Listrik	21.7256
T2	2 Peralatan Dapur	17.9722
T3	3 Obat Serangga	12.8841
T4	4 Peralatan Pertukangan	20.1246
T5	5 Perlengkapan Rumah Tangga	13.1909
T6	6 Peralatan bengkel	57.5934
T7	7 Lampu	75.4056
T8	8 Pewangi	62.7774
T9	9 Cat	26.6833
T10	10 Aksesoris	72.6705

Perhitungan jarak Euclidean dari data hasil uji terbanyak dari nilai k tetangga terdekat yang ditentukan sebesar 8. Hasil dari perhitungan dengan nilai k yang telah ditentukan yaitu 8 maka hasil yang diperoleh dengan terdekat ke K yaitu T3, T5, T2, T4, T1, T9, T6 dan T8

Perhitungan yang diterapkan dengan algoritma K-Neighbour Nearest, maka dapat diprediksikan produk yang laku, dan menjadi patokan kepada pelaku UMKM untuk mengambil keputusan dalam menentukan stok produk pada periode berikutnya.

Pengolahan Data lebih lanjut menggunakan aplikasi WEKA untuk mendapatkan informasi lebih lanjut dengan pemilihan klasifikasi. Informasi yang dihasil dari pengolahan data menggunakan aplikasi WEKA dapat dilihat pada gambar 4. Informasi yang diperoleh dapat dijadikan acuan awal dalam analisis sebelum pelaku UMKM mengambil keputusan.

Informasi diperoleh dapat di proses lebih lanjut dengan menampilkan grafik/kurva dengan pendekatan nilai k yang ditentukan yaitu 8. Hasil grafik/kurva dapat dilihat pada gambar 5. Titik-titik pada gambar 5., menunjukkan data produk yang diolah dengan pemilihan fitur yang tersedia pada aplikasi WEKA.

```

Classifier output
====
Best first.
Search met: no attributes
Search direction: Forward
Stops search after 5 node expansions
Total number of subsets evaluated: 25
Merit of best subset found: 100
Evaluation (for feature selection): CV (leave one out)
Feature set: 3,5

Time taken to build model: 0.01 seconds

==== Stratified cross-validation ====
see Summary view

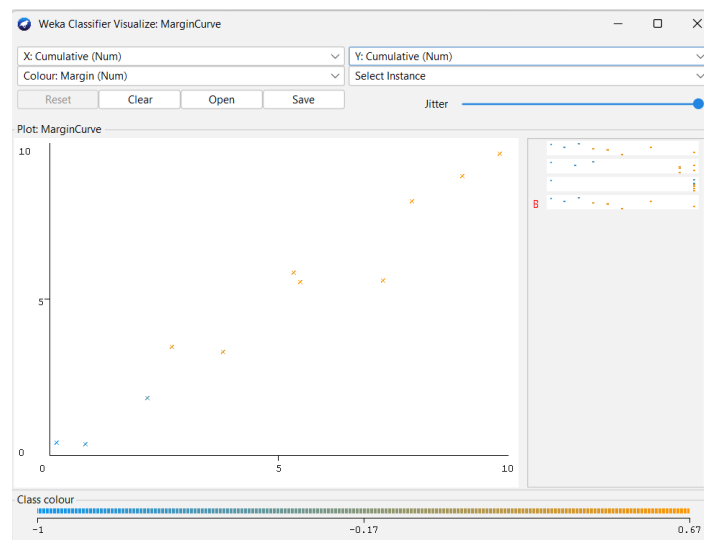
Currently classified instances  0      80  %
Incorrectly classified instances  2      20  %
Gappa statistic  0.5933
Mean absolute error  0.2129
Root mean squared error  0.4529
Relative absolute error  62.4701 %
Root relative squared error  71.8989 %
Total Number of Instances  10

==== Detailed Accuracy By Class ====
TP Rate  FP Rate  Precision  Recall  F-Measure  MCC  ROC Area  PRC Area  Class
0.833  0.250  0.833  0.833  0.833  0.583  0.668  0.794  Kurang Laku
0.750  0.187  0.750  0.750  0.750  0.583  0.468  0.643  Laku
Weighted Avg.  0.800  0.217  0.800  0.800  0.800  0.583  0.668  0.742

==== Confusion Matrix ====
a b  --- classified as
1 1 | a = Kurang Laku
1 2 | b = Laku
    
```

**Gambar 4. Hasil Keluaran Klasifikasi**

Gambar 4. Menunjukkan ringkasan dari data, tingkat keakuratan dari hasil pengolahan data menggunakan aplikasi WEKA.



**Gambar 5.** Hasil Visualisasi MarginCurve

Titik-titik pada gambar 5., menunjukkan distribusi penjualan item yang diproses dengan nilai  $k$  yang ditentukan sebesar 8. Titik dengan warna kuning menunjukkan nilai  $k$  yang ditetapkan, sedangkan warna biru dapat diartikan sebagai jarak terjauh dari nilai tetangga yang telah ditentukan sebagai nilai training yaitu sebesar 30.

## 5. Kesimpulan

Data yang dikumpulkan perlu dilakukan proses anomali data agar hasil analisis lebih baik, karena data anomali dapat menyebabkan hasil pengolahan data kurang efektif. Data yang telah melalui proses seleksi kemudian dilakukan perhitungan dengan algoritma K-Nearest Neighbour dan bantuan aplikasi untuk mengetahui sejauh mana produk mana yang laku dan tidak laku dengan pendekatan yang terdapat pada algoritma KNN. Hasil yang diperoleh dari hasil penelitian perlu dikelompokkan lebih detail untuk memperoleh detail produk yang laku atau tidak berdasarkan periode, musiman, atau kategori tergantung pada produk yang tersedia. Analisis data dari aplikasi WEKA dapat digunakan sebagai tolak ukur produk mana saja yang perlu dilakukan stok, atau produk mana yang apabila sudah laku terjual tidak perlu lagi stok. Keputusan yang nantinya diambil pelaku UMKM resiko skala kecil diharapkan bisa membuat arus kas menjadi lebih baik sehingga tidak terbebani pada stok yang di gudang.

## Daftar Pustaka

- [1] Ardiani, G.T., Apip Supriadi, Andi Rustandi, D. H. L. K. (2018). Analytical Hierarchy Process (AHP) Teknik Penentuan Strategi Daya Saing Kerajinan Bordir. Grup Penerbitan CV Budi Utama
- [2] Pujawan, I Nyoman dan Mahendrawathi. (2017). Supply Chain Management. ANDI, Yogyakarta
- [3] Azmi, T. U., Haryanto, H., & Sutojo, T. (2018). Prediksi Jumlah Produksi Jenang di PT Menara Jenang Kudus Menggunakan Metode Logika Fuzzy Tsukamoto. *Sisfotenika*, 8(1), 23. <https://doi.org/10.30700/jst.v8i1.176>
- [4] Fathoni, M. Y., Prabowo, D. A., Wijayanto, S., Fernandez, S., & Susanto, A. (2022). Analisis Kinerja Rantai Pasok Produk Kedelai Menggunakan Metode Supply Chain Operation Reference. *Jurnal Informatika: Jurnal Pengembangan IT*, 7(2), 74–79. <https://doi.org/10.30591/jpit.v7i2.3740>
- [5] S. P. Dewi, N. Nurwati, and E. Rahayu, “Penerapan Data Mining Untuk Prediksi Penjualan Produk Terlaris Menggunakan Metode K-Nearest Neighbor,” *Build. Informatics, Technol. Sci.*, vol. 3, no. 4, pp. 639–648, 2022.
- [6] Imaroh, T. S., Widiyani, K., & Muttaqien, F. (2023). Value Chain and Supply Chain Management of Products from Women Farmer Groups in South Tangerang. *Wiga : Jurnal Penelitian Ilmu Ekonomi*, 13(2), 190–203. <https://doi.org/10.30741/wiga.v13i2.1088>
- [7] Rainanto, B.H, 2019, Analisis Permasalahan Yang Dihadapi Oleh Pelaku Usaha Mikro Agar Berkembang Menjadi Usaha Kecil (Scalling Up) Pada Umkm Di 14 Kecamatan Di Kabupaten Bogor, *Jurnal Ilmiah Manajemen Kesatuan*, 7(1), 201 –210,
- [8] M. Faisal, W. S. Utami, and S. Parmica, “Implementasi Algoritma K-Nearest Neighbor (KNN) Dalam Memprediksi Indeks Kemiskinan,” *J. Sensi*, vol. 9, no. 1, pp. 11–23, 2023.
- [9] S. Hendrian, “Algoritma Klasifikasi Data Mining Untuk Memprediksi Siswa Dalam Memperoleh Bantuan Dana Pendidikan,” *Fakt. Exacta*, vol. 11, no. 3, pp. 266–274, 2018.