

PENERAPAN ALGORITMA *K-NEAREST NEIGHBOR* UNTUK MEMPREDIKSI PENJUALAN BAJU SERAGAM SEKOLAH TERLARIS PADA CV UNI M

Rachmat Setiabudi

*Program Studi Teknik Informatika, Fakultas Tehnologi Industri, Institut Teknologi Budi
Utomo*

raffisetiabudi@gmail.com

Abstrak

Pertumbuhan ekonomi Indonesia mengalami perlambatan pada tahun 2023, mempengaruhi banyak industri termasuk UKM seperti CVUNI M yang bergerak di bidang pembuatan dan penjualan seragam sekolah. Penurunan penjualan akibat dampak pandemi *COVID-19* memicu perlunya strategi baru berbasis teknologi informasi untuk meningkatkan penjualan. Penelitian ini bertujuan untuk memprediksi penjualan seragam sekolah terlaris di CV UNI M menggunakan algoritma *K-Nearest Neighbor* (KNN). Identifikasi masalah meliputi kurangnya analisis data untuk mengoptimalkan pasokan produk dan mempermudah pengambilan keputusan perusahaan. Penelitian ini menggunakan data penjualan tahun 2023 dari CV UNI M serta metode KNN yang diimplementasikan melalui *Software RapidMiner*. Model prediksi yang dihasilkan diharapkan dapat membantu perusahaan dalam menyusun strategi penjualan yang lebih efektif, mengurangi risiko kerugian akibat stok berlebih atau kekurangan, dan meningkatkan kepuasan pelanggan melalui penyediaan produk yang tepat. Hasil penelitian menunjukkan bahwa penerapan algoritma KNN efektif dalam memprediksi penjualan seragam sekolah terlaris, dengan akurasi yang memadai untuk digunakan sebagai alat bantu dalam perencanaan stok dan produksi. Temuan ini dapat menjadi referensi bagi UKM lain dalam mengadopsi teknologi data mining untuk meningkatkan kinerja bisnis mereka.

Kata Kunci : Penjualan, Seragam, *Data Mining*, *K-Nearest Neighbor*, *RapidMine*

1. PENDAHULUAN

Pertumbuhan ekonomi Indonesia tahun 2023 sebesar 5,05% mengalami perlambatan dibandingkan dengan tahun 2022 yang mencapai 5,31%. Perlambatan pertumbuhan ekonomi tersebut disebabkan oleh beberapa hal salah satunya yaitu berbagai industri di Indonesia yang melakukan pemutusan hubungan kerja (PHK) karena perusahaan tersebut terus menerus mengalami kerugian karena perlambatan transformasi, dampak pandemi covid-19, penurunan permintaan pasar dan lain sebagainya. Selain itu dampak pandemi covid 19 yang belum sepenuhnya pulih sepenuhnya berdasarkan data dari Kementerian Koperasi dan UKM terdapat lebih dari 164.173 pelaku usaha yang terdampak pandemi covid 19. Permasalahan tersebut juga terjadi pada CV UNI M yang merupakan industri UKM di bidang pakaian dan fashion yang terdampak hingga lebih dari 50% penjualan yang turun.

Peningkatan penjualan perlu menjadi fokus karena pada tahun 2024 pemerintah

Indonesia sedang menggalakan digitalisasi UKM & UMKM yang menargetkan 30 juta pengusaha beralih ke ranah digital.

Pasar Kemeja Seragam Sekolah adalah bagian penting pasar pakaian sekolah. CV UNI M adalah perusahaan yang bergerak di bidang pembuatan dan penjualan kemeja seragam sekolah. Agar tetap kompetitif, CV UNI M perlu memahami desain kemeja mana yang paling mungkin untuk dijual. Meningkatnya permintaan seragam sekolah telah menjadi tantangan yang signifikan bagi banyak sekolah dan pemasok seragam. Memprediksi penjualan seragam sekolah secara akurat dapat membantu pemasok

mengoptimalkan inventaris mereka, mengurangi pemborosan, dan memenuhi kebutuhan pelanggan mereka dengan lebih baik. Makalah penelitian ini mengeksplorasi penerapan algoritma *K-Nearest Neighbor* (KNN) untuk meramalkan item seragam sekolah terlaris di CV UNI M.

Masalah yang umum terjadi di kalangan pemilik bisnis adalah kemampuan untuk meramalkan penjualan di masa depan berdasarkan data historis sebuah masalah yang sulit dipecahkan oleh banyak orang. Prediksi ini mempunyai risiko besar bagi pemilik karena menentukan keputusan mengenai jumlah barang yang akan disediakan oleh perusahaan.

2. METODOLOGI

2.1. K-Nearest Neighbor (K-NN)

K-Nearest Neighbor (K-NN) adalah salah satu algoritma yang digunakan dalam machine learning untuk masalah klasifikasi dan regresi. Algoritma ini beroperasi dengan cara mencari tetangga terdekat dari suatu titik data berdasarkan ukuran jarak, seperti jarak *Euclidean* atau jarak Manhattan. Berikut ini poin-poin utama tentang K-NN:

1. Prinsip Kerja: K-NN mengklasifikasikan titik data baru berdasarkan mayoritas label dari K tetangga terdekatnya dalam dataset pelatihan. Dalam kasus regresi, K-NN menghitung nilai rata-rata dari K tetangga terdekat untuk memprediksi nilai kontinu
2. Parameter K: Parameter K adalah jumlah tetangga terdekat yang akan dipertimbangkan dalam proses klasifikasi atau regresi. Nilai K harus ditentukan sebelumnya oleh pengguna. Pemilihan nilai K yang tepat dapat mempengaruhi kinerja dan akurasi algoritma. Metrik Jarak: Algoritma K-NN menggunakan metrik jarak, seperti jarak Euclidean, jarak Manhattan, atau metrik jarak lainnya untuk mengukur seberapa dekat atau jauhnya titik data yang baru dengan titik data lain dalam dataset.
3. Aplikasi: K-NN dapat diterapkan dalam berbagai bidang seperti pengenalan pola, sistem rekomendasi, analisis teks, dan pengolahan citra. Kelebihannya termasuk sederhana dalam konsep dan mudah diimplementasikan. Di bawah ini adalah rumus *Euclidean distance*

Gambar 1 Rumus *Ecliden distance* di mana:

- 1) $d(x, y)$ adalah antara jarak data x dan data y .
- 2) x_i yaitu data testing ke- i
- 3) y_i yaitu data training ke- i

Berikut ini adalah langkah-langkah algoritma K-NN:

- 1) menentukan parameter nilai k (dimana nilai k akan dipilih dengan manual).
- 2) untuk menghitung antara jarak data *testing* dan data *training*
- 3) mengurutkan data training dengan berdasarkan nilai jarak yang terkecil.
- 4) untuk menetapkan kelas, di mana data testing akan dipilih berdasarkan nilai k dengan jumlah terbanyak.

2.2. Algoritma C4.5

Dapat melakukan pengembangan diantaranya mengatasi *continue data*, *running* serta mengatasi *missing value*. Algoritma ini adalah salah satu teknik dari pohon keputusan atau *decision tree* yang dapat digunakan untuk menghasilkan beberapa aturan. sebuah *decision tree* atau pohon keputusan memiliki tujuan untuk meningkatkan nilai prediksi dan keakuratan data yang dilakukan.

2.3. Algoritma K-Means

Algoritma ini adalah salah satu algoritma yang mengelompokkan data non hirarki yang membagi data dalam bentuk 2 atau lebih kelompok. metode atau cara ini membagi data ke dalam kelompok agar data yang berkarakteristik sama dapat dimasukkan ke dalam satu kelompok. Data dengan karakteristik yang berbeda akan dikelompokkan dengan kelompok data yang lain. pengelompokkan data ini bertujuan untuk memaksimalkan variasi antar kelompok dan juga meminimalkan fungsi objek yang diatur ke dalam suatu kelompok.

Dalam penelitian ini menggunakan penelitian deskriptif kualitatif, yaitu penelitian yang mengelola dan menggambarkan data serta informasi berdasarkan fakta-fakta yang tampak untuk kemudian dianalisis lebih lanjut. Metode ini tidak terbatas sampai pada pengumpulan (

$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad 71$$

analisis. Penyampaian data dan informasi digambarkan dalam bentuk tampilan kalimat yang lebih bermakna dan mudah dipahami. Penelitian deskriptif kualitatif yang bertujuan untuk memberikan gambaran secara jelas mengenai prediksi penjualan baju seragam sekolah pada CV UNI M.

Adapun teknik analisis data dalam penelitian kualitatif secara umum dimulai dari :

1. Analisis Data

Selama pengumpulan data peneliti menetapkan fokus penelitian, penyusunan temuan sementara berdasarkan data yang terkumpul, pembuatan rencana pengumpulan data berikutnya, penetapan sasaran pengumpulan data.

2. Reduksi Data

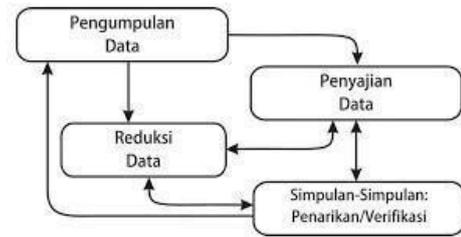
Peneliti dapat melakukan pemilihan data yang hendak dikode mana yang dibuang mana yang diambil yang merupakan ringkasan, cerita, apayang sedang berkembang.

3. Penyajian Data

Penyajian data yaitu menyajikan sekumpulan informasi yang tersusun dan memberikan kemungkinan adanya penarikan kesimpulan dan pengambilan tindakan. Penyajian data yang telah diperoleh dari lapangan terkait dengan seluruh permasalahan penelitian kemudian dipilih sesuai dengan yang dibutuhkan, lalu dikelompokkan kemudian diberikan batasan masalah. Dari penyajian data tersebut, maka diharapkan dapat memberikan kejelasan data yang detail dan *substantive* dengan data pendukung.

4. Penarikan Kesimpulan

Kesimpulan dapat diverifikasi selama kegiatan berlangsung yang merupakan tinjauan ulang pada catatan lapangan yang ada.

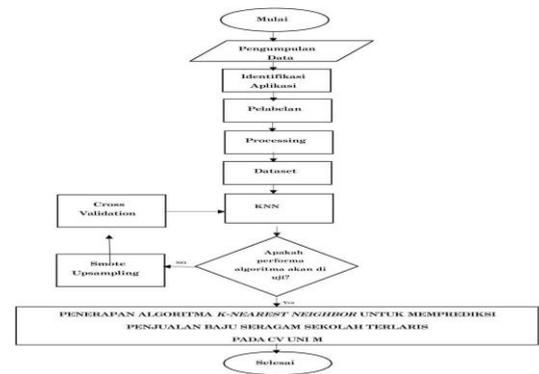


Gambar 2 Metode Analisa Kualitatif
Sumber : Penelitian mandiri 2024

3. HASIL DAN PEMBAHASAN

3.1 Hasil

3.1.1 Pemodelan Data



Gambar 3 Permodelan Data
Sumber : Penelitian mandiri 2024

Berikut penjelasan berdasarkan pemodelan data yang peneliti buat:

1. Pengumpulan Data
2. Identifikasi Aplikasi
3. *Text Preprocessing*
4. *Smote Upsampling*
5. *Cross Validation*

Setelah melakukan analisa terhadap perancangan dengan tahapan data mining untuk menghasilkan prediksi penjualan produk seragam terlaris pada CV UNI M Algoritma *K-Nearest Neighbor*, analisa ini berakhir dengan melakukan proses *data mining* yang sesungguhnya, maka hasil yang dicapai oleh peneliti adalah untuk mengetahui prediksi penjualan seragam terlaris berdasarkan data penjualan produk seragam dari data penjualan di tahun 2023 bulan Juli, Agustus, September, Oktober, November, Desember yang ada di CVUNI M.

Peneliti menggunakan aplikasi *RapidMiner* untuk mempermudah proses data mining, yang menghasilkan informasi prediksi penjualan seragam terlaris. Dan didapatkan hasil prediksi penjualan produk seragam terlaris sebanyak 1.521 jenis produk yang terjual.

3.2 Pembahasan

Data yang digunakan dalam penelitian ini yaitu data penjualan produk seragam berdasarkan penjualan tahun 2023 dari bulan Juli, Agustus, September, Oktober, November. Kemudian data tersebut diseleksi dan akan digunakan untuk diolah dalam memprediksi penjualan seragam terlaris. Adapun atribut yang digunakan dalam penentuan prediksi penjualan seragam terlaris adalah atribut nama barang, jumlah barang, dan bulan.

Tabel1 Data *Selection* CV UNI M Tahun 2023

No	Nama Barang	Jumlah	Bulan
1	Celana UNIM TK Orange 0	1	Juli
2	Rok Unim Putih SP17	1	Juli
3	Rok Unim Putih SP18	2	Juli

No	Nama Barang	Jumlah	Bulan
4	PNS Pdk W BAN L-60	2	September
5	PNS Pdk W BAN XL-60	1	September

Sumber : Penelitian mandiri 2024

Tahap *preprocessing* dilakukan untuk pengelompokan jenis penjualan produk seragam berdasarkan jumlah penjualan 6 bulan pada tahun 2023 untuk mempermudah dalam proses perhitungan prediksi. Setelah data terkelompokan

lalu semua dijumlahkan sehingga menjadi data penjualan untuk semua seragam.

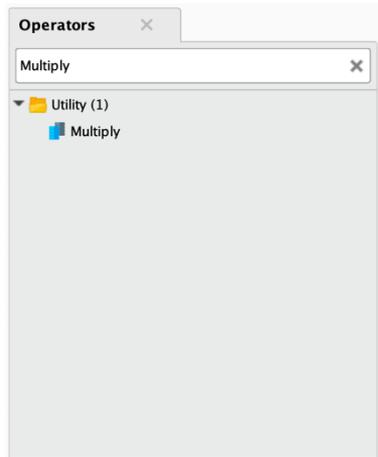
Tabel 2 Pengelompokan Data
Sumber : Penelitian mandiri 2024

Nama Barang	Jul	Agst	Sept	Okt	Nov	Des
celana unim larici kempol34	0	0	0	0	0	1
Celana UNIM SD 22	0	0	0	0	0	1
Hem UNIM SMP 9	0	0	0	0	0	1
Hem Unim SMA Tisu 10	0	0	0	0	0	3
Dress Karen Mode Shimer	0	0	0	0	0	2
Hem Korpri Uni-M W M	0	0	0	0	0	4
Dress Mofei F	0	0	0	0	0	1
Dress Mofei L	0	0	0	0	0	9
Lacos Hijau Tosca PolosXL	0	0	0	0	0	1
Topi Pramuka Pet Bludru	0	0	0	0	0	1

Pada tahap *transformation* ini hasil dari pengelompokan data *preprocessing* kemudian digunakan untuk data training. Proses pembentukan data *training* berdasarkan data yang ada, data harus di seleksi terlebih dahulu untuk menentukan atribut mana yang dapat mempengaruhi penjualan seragam terlaris yang disebut data target, dimana data target merupakan data yang berisikan atribut yang akan menjadi atribut yang relevan dan mendukung dalam proses *data mining*. Adapun data *training* yang digunakan adalah seperti yang ditunjukkan.

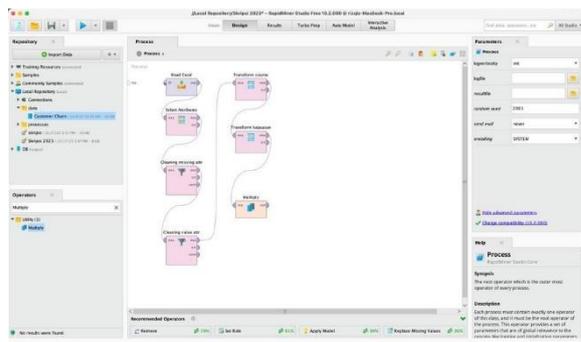
Tahapan *modelling* adalah langkah kunci dalam data mining yang sangat penting. Tahap ini menggunakan beberapa algoritma yaitu, *Naïve Bayes*, *K-Nearest Neighbors*, dan *Decision Tree*. Berdasarkan hasil dari proses *modelling* ini dapat dilakukan proses evaluasi dan memilih algoritma yang memiliki tingkat akurasi tertinggi di antara keempat algoritma tersebut. Pilihan ini kemudian dapat digunakan untuk melakukan prediksi *customer churn*.

Langkah pertama dalam melakukan *modelling* ini kita perlu melakukan penambahan *operator Multiply*. *Operator Multiply* digunakan untuk menampung atau memproses *modelling* dengan beberapa algoritma sekaligus. Cari *operator Multiply* di menu *operator* seperti gambar 4.



Gambar 4 Operator Multiply di menu Operator
Sumber : Penelitian mandiri 2024

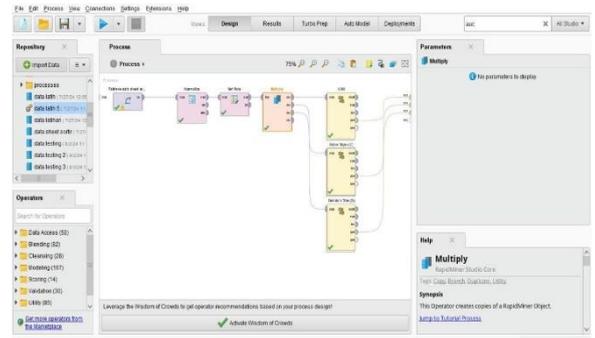
Langkah selanjutnya adalah *drag operator Multiply* bagian menu Process. Selanjutnya, hubungkan operator tersebut dengan operator lain pada tahapan sebelumnya seperti pada gambar 5



Gambar 5 Tampilan operator Multiply di menu Process
Sumber : Penelitian mandiri 2024

Langkah selanjutnya adalah penambahan *operator Cross Validation* di menu proses. Kemudian drag *Cross Validation* ke bagian menu process sebanyak empat kali. Selanjutnya, lakukan *rename* pada operator keempat *Cross Validation* menjadi masing-masing algoritma yang telah ditentukan.

Langkah terakhir pada pemodelan ini adalah menghubungkan *output* keempat *Cross Validation* ke *response* di RapidMiner.



Gambar 6 Tampilan akhir dari modeling
Sumber : Penelitian mandiri 2024

Operator	Accuracy	Precision	Recall	F1 Score
Model	98.98%	95.50%	98.98%	97.23%

Gambar 7 Hasil dari algoritma K-Nearest Neighbors
Sumber : Penelitian mandiri 2024

4. KESIMPULAN

Dalam penelitian ini, dilakukan perbandingan akurasi empat algoritma *K-Nearest Neighbors*, *Naïve Bayes*, dan *Decision Tree* dalam konteks prediksi penjualan terlaris di CV UNI M. Hasil analisis data menunjukkan bahwa algoritma *K-Nearest Neighbors* adalah yang terbaik dengan tingkat *accuracy* 98.98% dan nilai *precision* 95.50%. Kesimpulan ini memberikan informasi krusial untuk pemilihan algoritma terbaik dalam menangani permasalahan prediksi penjualan Terlaris

DAFTAR PUSTAKA

A. G. Pramesti, Q. J. Adrian, and Y. Fernando, "Perancangan Ui/Ux Pada Aplikasi Pemesanan Buket Menggunakan Metode User Centered Design (Studi Kasus: Bouquet Lampung)," J. Inform. dan Rekayasa Perangkat Lunak, vol. 3, no. 2, pp. 179–184, 2022

Prabowo and D. Darwis, "Implementasi Algoritma Levensthein Distance Guna," vol. 3, no.1, pp. 57–65, 2022.

Han, J., & Kamber, M. (2023). Data Mining: Concepts and Techniques. 4th Edition. Morgan Kaufmann Publishers.

M. Akbar and Y. Rahmanto, “Desain Data Warehouse Penjualan Menggunakan Nine Step Methodology Untuk Business Intelegency Pada Pt Bangun Mitra Makmur,” *J. Inform. Dan Rekayasa Perangkat Lunak*, vol. 1, no. 2, pp. 137– 146, 2020, doi: 10.33365/jatika.v1i2.331.