
ANALISA DATA MINING UNTUK PREDIKSI PENJUALAN PRODUK TERLARIS MENGGUNAKAN METODE K-NEAREST NEIGHBOR (STUDI KASUS: TOKO SOLIDGEAR)

Shandi Noris¹, Yudha Ardiyansyah²

^{1,2}Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Pamulang
E-mail : ¹dosen00354@unpam.ac.id, ²yudha.ardiyansyah07@gmail.com

ABSTRAK

Toko Solidgear merupakan sebuah perusahaan pakaian yang menyediakan dan memproduksi berbagai macam produk pakaian pria seperti T-SHIRT, Jaket Hoodie, Crewneck, Tas dan Topi. Dilihat dari banyaknya permintaan kostumer akan produk, maka diperlukanya prediksi untuk penjualan yang paling diminati oleh konsumen dan agar membantu pihak perusahaan dapat mengenal minat dari konsumen berdasarkan penjualan produk terlaris. Penelitian ini mengimplementasikan teknik data mining untuk memprediksikan produk terlaris menggunakan algoritma *K-Nearest Neighbor* dan *Rapidminer*. *K-nearest neighbor* merupakan metode yang digunakan untuk mengklarifikasi atau mencari data terhadap suatu objek yang berasal dari pembelajaran yang jarak antar data yang paling dekat dengan objek yang di cari. Dari hasil prediksi yang dilakukan mendapatkan hasil 9 produk T-SHIRT dinyatakan laris, 5 produk JAKET HOODIE dinyatakan laris, 5 produk TAS dinyatakan laris, 6 produk CREWNECK dinyatakan laris, dan 6 produk Topi dinyatakan laris. Dari hasil analisa penelitian ini dapat disimpulkan bahwa produk yang di nyatakan tidak laris disebabkan karena desain pada produk sudah terlalu lama yang menyebabkan kurang diminati oleh pelanggan. Oleh karena itu agar produk dapat meningkatkan minat dari konsumen dan terjual dengan maksimal perusahaan dapat memperbarui atau mengubah desain dari produk yang kurang diminati.

Keywords: Prediksi; Data Mining; *K-Nearest Neighbor*.

1. PENDAHULUAN

Penjualan adalah sebuah faktor terpenting bagi perusahaan. Karena dengan adanya penjualan, perusahaan akan mendapat keuntungan yang lebih supaya bisa melanjutkan usaha tersebut. Namun Persaingan bisnis di era perdagangan bebas seperti sekarang ini sangat ketat, setiap perusahaan dituntut untuk mempersiapkan dirinya secara profesional dan fleksibel sehingga perusahaan tidak hanya mampu bertahan, tetapi juga mampu tumbuh dan berkembang. Oleh karena itu diperlukanya proses data mining untuk memprediksi hal tersebut. Prediksi pada penjualan terlaris dapat bermanfaat untuk memperkirakan apakah produk dapat dinyatakan laris atau tidak, yang dapat juga membantu menentukan pontensi penjualan dari produk tesebut di masa depan. Data mining merupakan sebuah analisa terhadap data yang sudah ada untuk menemukan hubungan yang jelas serta menyimpulkannya data yang belum diketahui sebelumnya dengan cara terkini dipahami dan berguna bagi pemilik data tersebut. Dalam data mining, data disimpan secara elektronik dan diproses secara otomatis oleh komputer menggunakan teknik dan perhitungan tertentu (Baharuddin, 2019).

Prediksi dapat di lakukan dengan menggabungkan data yang telah ada untuk memprediksikan hasil hal baru yang akan terjadi selanjutnya, singkatnya prediksi merupakan memperkirakan hal baru yang akan terjadi selanjutnya. Algoritma *K-Nearest Neighbor* merupakan salah satu dari metode klasifikasi yang terdapat pada data mining, metode ini digunakan untuk mengklarifikasi atau mencari data terhadap suatu objek yang berasal dari pembelajaran yang jarak antar data yang paling dekat dengan objek yang ingin di cari. Keuntungan dari metode *K-Nearest Neighbor* ini untuk

digunakan dalam meprediksi data-data yang efektif apabila data trainingnya besar (Pratama, 2019).

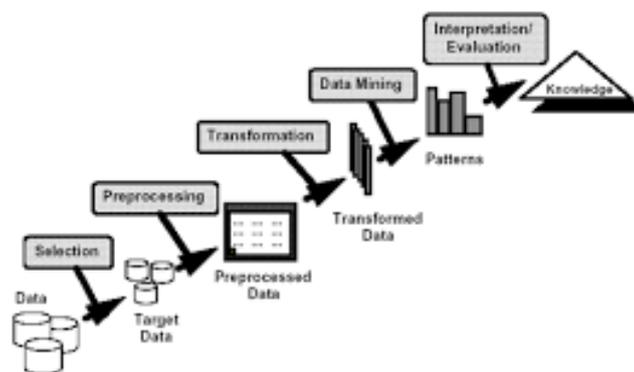
Toko Solid Gear merupakan sebuah perusahaan yang bergerak di bidang busana yang menyediakan berbagai macam produk busana. Toko Solid Gear saat ini sudah mengembangkan penggunaan sistem informasi penjualan untuk mendukung kinerja dari pihak perusahaan maupun layanan informasi penjualan. Perusahaan juga telah menerapkan penjualan sistem penjualan secara online dalam memasarkan produknya. Produk yang dijual antara lain seperti: berbagai macam produk T-SHIRT, Jaket Hoodie, Crewneck, Tas dan Topi. Dilihat dari banyaknya permintaan kostumer akan produk busana, maka diperlukanya prediksi untuk penjualan yang paling diminati oleh konsumen dan agar pihak perusahaan dapat mengenal minat dari konsumen berdasarkan penjualan produk terlaris.

2. METODE PENELITIAN

Metode yang digunakan pada penelitian ini adalah dengan menggunakan metode *Knowledge Discovery in Database* dan Metode *K-Nearest Neighbor* sebagai berikut:

a) Metode *Knowledge Discovery in Database*

Knowledge Discovery in Database (KDD) adalah proses menentukan informasi yang berguna serta pola-pola yang ada dalam data. Informasi ini terkandung dalam basis data yang berukuran besar yang sebelumnya tidak diketahui dan potensial bermanfaat (Baharuddin, 2019). Berikut tahapan proses KDD:



Gambar 1 Tahapan pada KDD

1) *Data Selection*.

Pada proses ini dilakukan pemilihan himpunan data, menciptakan himpunan data target, atau memfokuskan pada subset variable (sampel data) dimana penemuan (*discovery*) akan dilakukan. Hasil seleksi disimpan dalam suatu berkas yang terpisah dari basis data operasional.

2) *Pre-Processing* dan *Cleaning Data*.

Pre-Processing dan *Cleaning Data* dilakukan membuang data yang tidak konsisten dan *noise*, duplikasi data, memperbaiki kesalahan data, dan bisa diperkaya dengan data eksternal yang relevan.

3) *Transformation*.

Proses ini mentransformasikan atau menggabungkan data ke dalam yang lebih tepat untuk melakukan proses mining dengan cara melakukan peringkasan (agregasi).

4) *Data Mining*

Data Mining proses mencari pola atau informasi menarik dalam data terpilih dengan menggunakan teknik, metode atau algoritma tertentu sesuai dengan tujuan dari proses KDD secara keseluruhan.

5) *Interpretation / Evaluasi*

Interpretation / Evaluasi merupakan proses untuk menerjemahkan pola-pola yang dihasilkan dari *Data Mining*. Mengevaluasi (menguji) apakah pola atau informasi yang ditemukan bersesuaian atau bertentangan dengan fakta atau hipotesa sebelumnya. Pengetahuan yang diperoleh dari pola-pola yang terbentuk dipresentasikan dalam bentuk visualisasi.

b) Prosedur Metode *K-Nearest Neighbor*

Langkah-langkah untuk menghitung algoritma K-NN (Rani, 2019):

- 1) Menentukan nilai k.
- 2) Menghitung kuadrat jarak *euclid* (*query instance*) masing-masing objek terhadap training data yang diberikan. Persamaan (1) menunjukkan perhitungan kuadrat jarak *euclid* (*query instance*) masing-masing objek terhadap training data yang diberikan.
- 3) Kemudian mengurutkan objek-objek tersebut ke dalam kelompok yang mempunyai jarak *euclid* terkecil.
- 4) Menghitung rata-rata dari nilai objek pada jangkauan K dengan menggunakan kategori *K-Nearest Neighbor* yang terdekat (jangkauan K), maka dapat diprediksi nilai *query instance* yang telah dihitung.

Perhitungan metode *K-Nearest Neighbor* dapat dirumuskan dengan perhitungan sebagai berikut:

$$d_e = \sqrt{\sum_{k=5}^m (fd_{i,k} - k_j)^2}$$

Keterangan:

- d_e = Jarak euclidean
 fd_i = Data training
k = Data testing
m = Jumlah data penelitian

3. HASIL DAN PEMBAHASAN

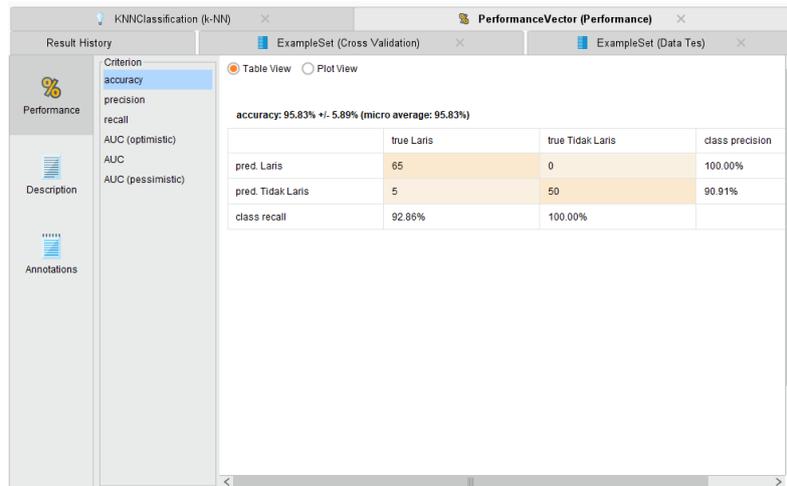
a) Hasil Penelitian

Setelah melakukan analisa dan penelitian pada tahap data mining untuk menghasilkan prediksi penjualan produk dan potensi penjualan produk terlaris pada Toko Solidgear dengan menggunakan K-nearest Neighbour, analisa dan penelitian ini berakhir dengan proses penelitian data mining dengan menggunakan aplikasi rapidminer, Penelitian ini menggunakan aplikasi Rapidminer yang digunakan untuk mempermudah dalam proses data mining yang dilakukan, yang menghasilkan data prediksi penjualan produk dan potensi penjualan produk terlaris. Dari hasil prediksi 5 produk yang dilakukan mendapatkan hasil yaitu: 9 produk T-SHIRT dinyatakan laris, 5 produk JAKET HOODIE dinyatakan laris, 5 produk TAS dinyatakan laris, 6 produk CREWNECK dinyatakan laris, dan 6 produk Topi dinyatakan laris dengan tingkat akurasi data yang di hasilkan dari cross validatin sebesar 95.83%

b) Pembahasan

Rapid miner merupakan salah satu perangkat lunak (*software*) pengolahan data mining yang mengelolah data set untuk mencari pola data yang sesuai dengan tujuan pengolahan data tersebut, tidak semua algoritma dapat sesuai dengan atau dapat mengelolah data set yang ada,

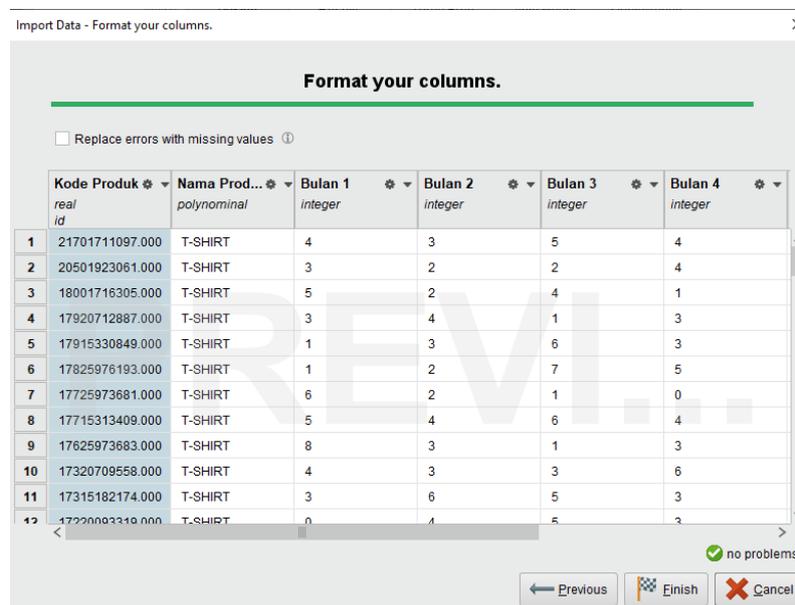
harus di lakukan peyesuaian data pola data dan sesuai dengan tujuan dari pengelolaan data tersebut.



	true Laris	true Tidak Laris	class precision
pred. Laris	85	0	100.00%
pred. Tidak Laris	5	50	90.91%
class recall	92.86%	100.00%	

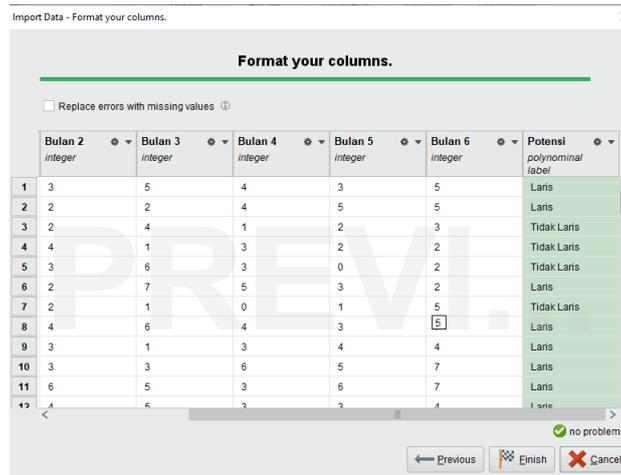
Gambar 2 Akurasi data prediksi potensi penjualan

Berdasarkan dari gambar 2 hasil dari pemodelan yang telah di proses di dalam *data mining* memiliki keakuratan data sebesar 95.83%. Susunan operator pada *data mining* adalah *drag and drop operator read excel* lalu lakukan importing data tabel microsoft excel kedalam proses membaca data training dalam format excel, Kemudian ganti atribut Kode produk menjadi ID dan atribut Potensi yang akan menjadi data yang akan dicari dengan menjadi label lalu setelah selesai klik finish. Seperti gambar berikut:



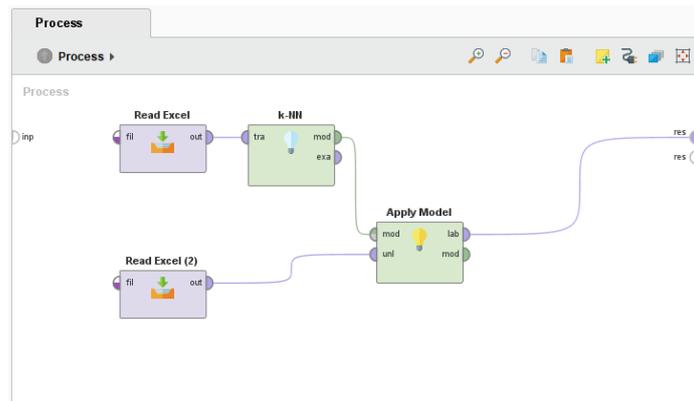
	Kode Produk	Nama Prod...	Bulan 1	Bulan 2	Bulan 3	Bulan 4
1	21701711097.000	T-SHIRT	4	3	5	4
2	20501923061.000	T-SHIRT	3	2	2	4
3	18001716305.000	T-SHIRT	5	2	4	1
4	17920712887.000	T-SHIRT	3	4	1	3
5	17915330849.000	T-SHIRT	1	3	6	3
6	17825976193.000	T-SHIRT	1	2	7	5
7	17725973681.000	T-SHIRT	6	2	1	0
8	17715313409.000	T-SHIRT	5	4	6	4
9	17625973683.000	T-SHIRT	8	3	1	3
10	17320709558.000	T-SHIRT	4	3	3	6
11	17315182174.000	T-SHIRT	3	6	5	3
12	17220093319.000	T-SHIRT	0	4	5	3

Gambar 3 Importing data training pada Operator Read Excel



Gambar 4 Label data training pada Operator Read Excel

Selanjutnya *drag and drop* operator k-NN dan operator apply model, kemudian memasukan nilai $K=5$ dan hubungkan *operator read excel* dengan operator k-NN. Setelah itu hubungkan k-NN dengan *operator apply model*, lalu ulangi lagi *operator read excel* (2) lalu lakukan *importing data testing*, sambungkan dengan *operator apply model* setelah itu hubungkan dengan *result*, Seperti gambar berikut:



Gambar 5 Susunan operator algoritma K-NN

Tahap ini adalah tahap terakhir dari proses data mining dengan menggunakan *rapid miner* dimana setelah semua operator terhubung kemudian klik *icon play* tombol F11, maka hasilnya akan muncul di dalam tab result, yang berisikan sebuah hasil prediksi dari seluruh data yang memenuhi nilai $k=5$, didapatkan hasil prediksi penjualan seperti gambar berikut:

Row No.	Kode Produk	Potensi	prediction(Potensi)	confidence(Laris)	confidence(Tidak Laris)	Nama Produk	Bulan 1
1	2151171997	Laris	Laris	1	0	T-SHIRT	2
2	20501823091	Tidak Laris	Tidak Laris	0.183	0.807	T-SHIRT	3
3	18001716305	Tidak Laris	Tidak Laris	0.399	0.601	T-SHIRT	4
4	17920712887	Tidak Laris	Tidak Laris	0.385	0.615	T-SHIRT	2
5	17915330849	Laris	Laris	0.794	0.206	T-SHIRT	5
6	17825976193	Laris	Laris	1	0	T-SHIRT	6
7	17722972081	Laris	Laris	1	0	T-SHIRT	3
8	17715313409	Laris	Laris	1	0	T-SHIRT	8
9	17629873883	Laris	Laris	0.608	0.392	T-SHIRT	8
10	17320706058	Laris	Laris	1	0	T-SHIRT	3
11	17315182174	Laris	Laris	1	0	T-SHIRT	2
12	17220803319	Laris	Laris	1	0	T-SHIRT	3

Gambar 6 Hasil prediksi penjualan T-SHIRT

Gambar 6 merupakan hasil prediksi potensi penjualan T-SHIRT di mana prediksi penjualan ini menghasilkan 9 produk laris dan 3 produk tidak laris.

Row No.	Kode Produk	Potensi	prediction(Potensi)	confidence(Laris)	confidence(Tidak Laris)	Nama Produk	Bulan 1
1	20501891701	?	Tidak Laris	0	1	JAKET HOOD...	2
2	20402073138	?	Tidak Laris	0.194	0.806	JAKET HOOD...	3
3	17920709220	?	Laris	0.588	0.414	JAKET HOOD...	1
4	17810998108	?	Tidak Laris	0	1	JAKET HOOD...	2
5	17025976290	?	Tidak Laris	0.201	0.799	JAKET HOOD...	0
6	17525973924	?	Tidak Laris	0.401	0.599	JAKET HOOD...	2
7	17320470091	?	Laris	1	0	JAKET HOOD...	1
8	17225974145	?	Tidak Laris	0	1	JAKET HOOD...	3
9	17141320967	?	Tidak Laris	0	1	JAKET HOOD...	0
10	17125974388	?	Laris	0.607	0.393	JAKET HOOD...	1
11	17025970938	?	Laris	1	0	JAKET HOOD...	2
12	17025973987	?	Laris	0.808	0.192	JAKET HOOD...	3

Gambar 7 Hasil prediksi penjualan JAKET HOODIE

Gambar 7 merupakan hasil prediksi potensi penjualan JAKET HOODIE di mana prediksi penjualan ini menghasilkan 5 produk laris dan 7 produk tidak laris.

Row No.	Kode Produk	Potensi	prediction(Potensi)	confidence(Laris)	confidence(Tidak Laris)	Nama Produk	Bulan 1
1	18302058534	?	Laris	0.793	0.207	TAS	1
2	17925976029	?	Tidak Laris	0	1	TAS	2
3	17848265525	?	Tidak Laris	0.211	0.789	TAS	3
4	17825973789	?	Tidak Laris	0.199	0.801	TAS	4
5	17720098676	?	Tidak Laris	0.186	0.814	TAS	4
6	17615189502	?	Laris	0.805	0.195	TAS	3
7	17189243566	?	Tidak Laris	0.207	0.793	TAS	2
8	17125974011	?	Tidak Laris	0.197	0.803	TAS	2
9	17020704442	?	Tidak Laris	0.204	0.796	TAS	2
10	16925974107	?	Tidak Laris	0.203	0.797	TAS	3
11	16748271354	?	Laris	0.607	0.393	TAS	4
12	16715334999	?	Laris	0.796	0.204	TAS	5

Gambar 8 Hasil prediksi penjualan TAS

Gambar 8 merupakan hasil prediksi potensi penjualan TAS di mana prediksi potensi penjualan ini menghasilkan 4 produk laris dan 8 produk tidak laris.

Row No.	Kode Produk	Potensi	prediction(Potensi)	confidence(Laris)	confidence(Tidak Laris)	Nama Produk	Bulan 1
1	18302058534	?	Laris	1.000	0	TAS	1
2	17925976029	?	Tidak Laris	0	1	TAS	2
3	17848265525	?	Tidak Laris	0.211	0.789	TAS	3
4	17825973789	?	Tidak Laris	0.199	0.801	TAS	4
5	17720098676	?	Tidak Laris	0.186	0.814	TAS	4
6	17615189502	?	Laris	0.805	0.195	TAS	3
7	17189243566	?	Tidak Laris	0.207	0.793	TAS	2
8	17125974011	?	Tidak Laris	0.197	0.803	TAS	2
9	17020704442	?	Tidak Laris	0	1	TAS	2
10	16925974107	?	Tidak Laris	0.203	0.797	TAS	3
11	16748271354	?	Laris	0.607	0.393	TAS	4
12	16715334999	?	Laris	0.597	0.403	TAS	5

Gambar 9 Hasil prediksi penjualan CREWNECK

Gambar 9 merupakan hasil prediksi potensi penjualan CREWNECK di mana prediksi penjualan ini menghasilkan 6 produk laris dan 6 produk tidak laris.

Row No.	Kode Produk	Potensi	prediction(Potensi)	confidence(Laris)	confidence(Tidak Laris)	Nama Produk	Bulan 1
1	17615339955	?	Tidak Laris	0.391	0.609	TOPI	3
2	17525876215	?	Laris	1	0	TOPI	2
3	17425975048	?	Tidak Laris	0	1	TOPI	4
4	17320707145	?	Laris	0.810	0.190	TOPI	4
5	17320466446	?	Tidak Laris	0.407	0.593	TOPI	3
6	17225976659	?	Tidak Laris	0.394	0.606	TOPI	3
7	17202082580	?	Tidak Laris	0	1	TOPI	2
8	16915328891	?	Tidak Laris	0	1	TOPI	1
9	1682655134	?	Tidak Laris	0.219	0.782	TOPI	3
10	16795020858	?	Laris	0.800	0.200	TOPI	4
11	16725974430	?	Tidak Laris	0.210	0.790	TOPI	4
12	16720100484	?	Laris	0.804	0.196	TOPI	3

Gambar 10 Hasil prediksi penjualan TOPI

Gambar 10 merupakan hasil prediksi potensi penjualan topi di mana prediksi penjualan ini menghasilkan 4 produk laris dan 8 produk tidak laris.

4. KESIMPULAN

Dari hasil pembahasan di atas, maka dapat diambil kesimpulan sebagai berikut:

- Dari hasil penelitian yang dilakukan menggunakan pemodelan dari algoritma *K-Nearest Neighbor* metode ini digunakan untuk mengklarifikasi atau mencari data terhadap suatu objek yang berasal dari pembelajaran yang jarak antar data yang paling dekat dengan objek yang ingin di cari menggunakan data Toko SolidGear yang diolah dengan menggunakan tahapan dari KDD (*Knowledge Discovery in Database*) sehingga menghasilkan prediksi produk yaitu 9 produk T-SHIRT dinyatakan laris, 5 produk JAKET HOODIE dinyatakan laris, 4 produk TAS dinyatakan laris, 6 produk CREWNECK dinyatakan laris, dan 4 produk Topi dinyatakan laris.
- Dari hasil analisa pada penelitian ini dapat disimpulkan bahwa produk yang di nyatakan tidak laris disebabkan karena desain pada produk sudah terlalu lama yang menyebabkan kurang diminati oleh pelanggan. Oleh karena itu agar produk dapat meningkatkan minat dari konsumen dan terjual dengan maksimal perusahaan dapat memperbarui atau mengubah desain dari produk-produk yang kurang diminati.

DAFTAR PUSTAKA

- A. A. Putri, "Penerapan Data Mining Untuk Memprediksi Penjualan Buah Dan Sayur Menggunakan Metode K-Nearest Neighbor (Studi Kasus : PT . Central Brastagi Utama)," vol. 1, no. 6, pp. 354–361, 2021.
- B. Hardiyanto and F. Rozi, "Prediksi penjualan sepatu menggunakan metode k- nearest neighbor," vol. 04, pp. 13–18, 2020.
- F. A. Pratama, R. Narasati, and D. R. Amalia, "Pengaruh Kata Cashback Terhadap Peningkatan Penjualan Menggunakan Data Mining," vol. 3, no. 2, pp. 1–5, 2019.
- I. A. Nikmatun, U. Diponegoro, I. Waspada, and U. Diponegoro, "IMPLEMENTASI DATA MINING UNTUK KLASIFIKASI MASA STUDI MAHASISWA MENGGUNAKAN ALGORITMA K-NEAREST NEIGHBOR," vol. 10, no. 2, pp. 421–432, 2019.
- J. Basuki, R. No, and P. Lombok, "MODEL DATA MINING UNTUK KAREKTERISTIK DATA TRAVELLER PADA PERUSAHAAN TOUR AND TRAVEL (Studi Kasus : Lombok Ceria Holiday)," vol. 2, no. 2, pp. 61–68, 2019.
- M. M. Baharuddin, T. Hasanuddin, and H. Azis, "ANALISIS PERFORMA METODE K-NEAREST NEIGHBOR UNTUK," vol. 11, no. 28, pp. 269–274, 2019.

- N. Nosiell, S. Sriyanto, and F. Maylani, "Perbandingan Teknik Data Mining Untuk Prediksi Penjualan Pada UMKM Gerabah," *Pros. Semin. Nas. Darmajaya*, vol. 1, no. 0, pp. 72–86, 2021.
- R. L. Hasanah, M. Hasan, and W. E. Pangesti, "KLASIFIKASI PENERIMA DANA BANTUAN DESA MENGGUNAKAN METODE KNN (K-NEAREST NEIGHBOR)," vol. 16, no. 1, pp. 1–6, 2019.
- S. R. Rani, S. R. Andani, and D. Suhendro, "Penerapan Algoritma K-Nearest Neighbor untuk Prediksi Kelulusan Siswa pada SMK Anak Bangsa," *Pros. Semin. Nas. Ris. Inf. Sci.*, vol. 1, no. September, p. 670, 2019, doi: 10.30645/senaris.v1i0.73.