

# Klasifikasi Vendor Menggunakan K-Nearest Neighbor (KNN) dan Weighted Product (WP)

Nia Desfiani S\*, Tacbir Hendro Pudjiantoro, Puspita Nurul Sabrina

Jurusan Informatika, Fakultas Sains & Informatika

Universitas Jenderal Achmad Yani

Jl. Terusan Sudirman, Cimahi

[niadesfianis17@if.unjani.ac.id\\*](mailto:niadesfianis17@if.unjani.ac.id)

**Abstrak**—Banyaknya *vendor* yang bekerja sama sehingga perusahaan memerlukan informasi mengenai solusi bisnis sebagai pendukung pengambilan keputusan. Maka penelitian ini melakukan proses data mining pada data pembelian dengan menggunakan metode K-NN dan *Weighted Product* dikarenakan lebih efektif dalam pencarian data dan mampu menyeleksi dengan baik. Proses klasifikasi pada penelitian ini memiliki tujuan memberitahu pengetahuan dari algoritma *K-Nearest Neighbor* dengan *Weighted Product* juga mampu memberikan informasi mengenai *vendor* yang sesuai dengan jarak, harga, *quantity*, dan total. Sehingga dapat disimpulkan dengan menggunakan metode K-NN dan *Weighted Product* ini efektif dan akurat dengan nilai akurasi yang didapatkan sebesar 75.46%. Hasil pengujian akurasi dengan *Confusion Matrix* dihasilkan *Accuracy* 73,48%, *Precision* 74%, *Recall* 77%, *F1-Score* 75%. Namun persentase yang dihasilkan cukup rendah tidak tercapai hingga persentase 100%, dikarenakan beberapa faktor yang mempengaruhi diantaranya jumlah data yang kurang banyak dan varian serta atribut yang digunakan kurang banyak.

**Kata kunci:** *Data Mining; K-Nearest Neighbor; Weighted Product; vendor;*

## I. PENDAHULUAN

*Data mining* merupakan suatu proses pencarian secara otomatis untuk menemukan pola atau model dari suatu database yang besar [1]. Umumnya, dalam *data mining* terdapatnya berbagai teknik-teknik, salah satunya ialah teknik klasifikasi yang dimana adalah proses untuk menemukan pola atau fungsi yang menjelaskan serta membedakan konsep maupun kelas data, dengan tujuan untuk dapat memperkirakan kelas dari suatu objek yang labelnya tidak diketahui [2]. Dibutuhkan sebuah metode pendukung dalam pembuatan suatu sistem. Metode yang digunakan ialah metode K-Nearest Neighbor dan metode Weighted Product atau biasa dikenal dengan singkatan WP. Dimana metode K-Nearest Neighbor ialah termasuk pada golongan *supervised*, dimana pada hasil *query instance* yang terbaru dilakukan klasifikasi sengan berdasarkan banyaknya kedekatan jarak dari kelompok atau fitur yang sudah ada[3]. Metode K-Nearest Neighbor dapat memilih kumpulan data yang dapat digolongkan atau di kelas kan menjadi rekomendasi dan tidak rekomendasi. Dengan banyaknya penggunaan metode K-NN dalam pengklasifikasian, metode K-NN juga mendapat suatu kritikan dalam penggunaannya karena apabila data uji dan data latih sedikit atau kurang banyak maka akan dianggap kurang baik penggunaannya. Maka dari itu harus ditambah dengan tambahan

metode lain. Dengan berbagai pertimbangan yang telah dipertimbangan dengan baik maka didekatkan dengan metode Weighted Product (WP). Dimana pada metode Weighted Product dapat digunakan dalam penyesuaian dan melengkapi serta menentukan kumpulan hasil penyeleksian klasifikasi oleh metode K-NN dengan melakukan perankingan untuk mendapatkan hasil yang terbaik [4].

Menurut penelitian Wang, dkk tahun 2013 dalam penelitian tersebut peneliti menyatakan metode K-Nearest neighbor ini dapat membantu penyaringan data yang ada pada database yang buruk. Dari kesimpulan penelitian ini pun menunjukkan bahwa pada metode K-Nearest Neighbor lebih *efisien* dan sangat baik dalam proses pencarian data [5] Menurut penelitian Behnam, Pourghassem (2015:67) dalam penelitian tersebut menunjukkan metode K-Nearest Neighbor mampu digunakan dalam optimasi pencarian data [6].

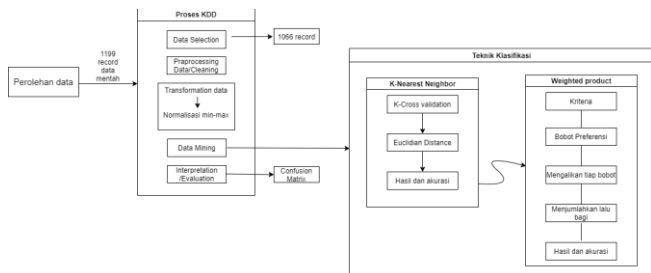
Selanjutnya menurut penelitian yang dilakukan oleh Nihru Nafi Dzikulloh (2017) penelitian tersebut menggunakan metode K-Nearest Neighbor dan Weighted Product dikarenakan mampu dan memiliki keunggulan dapat mengklasifikasikan data calon pegawai yang belum diketahui dengan menggunakan data uji dan data latih menghasilkan nilai akurasi sebesar 94 %, nilai *precision* sebesar 100% & nilai *recall* 80%. Sehingga dengan menggunakan pendekatan metode ini akan lebih efisien dikarenakan waktu yang dibutuhkan lebih singkat dalam melakukan perhitungan, namun masih terdapatnya kekurangan pada nilai akurasi dan nilai *recall* yang belum mencapai 100% [2].

Berdasarkan beberapa penelitian tersebut menyatakan bahwa metode K-Nearest Neighbor dan Weighted Product kinerjanya memiliki akurasi yang baik dan lebih efektif pada proses pencarian data, sehingga dalam dunia bisnis terdapat istilah *vendor* [7]. Dimana definisi dari *vendor* ini merupakan perusahaan yang berkontribusi dalam barang-barang atau jasa [8] Maka, seiring banyaknya *vendor* yang bekerja sama sehingga perusahaan memerlukan informasi mengenai solusi bisnis sebagai pendukung pengambilan keputusan. Penelitian ini melakukan proses data mining pada data pembelian dengan menggunakan metode K-Nearest Neighbor dan Weighted Product dikarenakan lebih efektif dalam pencarian data dan mampu menyeleksi dengan baik. Proses klasifikasi pada penelitian ini memiliki tujuan memberitahu pengetahuan dari algoritma K-Nearest Neighbor dengan Weighted Product juga mampu memberikan informasi mengenai *vendor* yang sesuai dengan jarak, harga, *quantity*, dan total.

Selain memberikan pengetahuan pada suatu algoritma mengenai ke efektifan dan ke akuratanya, juga mampu memberikan informasi bagi perusahaan mengenai rekomendasi *vendor* yang akan diajak bekerja sama kembali.

## II. METODE

Metode yang digunakan dalam penelitian ini dapat dilihat pada Gambar 1. Diagram Alir Metode.



Gambar 1. Diagram alir metode penelitian

### A. Perolehan Data

Dalam penelitian ini tahap pertama yang dilakukan ialah studi pustaka untuk mengumpulkan berbagai informasi dan sebagai referensi yang berkaitan dengan kata kunci penelitian selanjutnya pengumpulan data untuk perolehan data, yang dimana pada tahap ini dengan melakukan survey ke tempat atau observasi setelah melalui tahap survey selanjutnya melakukan wawancara dengan pihak perusahaan. Data yang digunakan pada penelitian ini yaitu Data Pembelian dari bulan Juni hingga Agustus 2019.

### B. Proses KDD

Proses KDD atau Knowledge Discovery In Databases ialah merupakan keseluruhan proses dalam data untuk mencari & mengidentifikasi sebuah pola [9]. Proses KDD ini dapat dilihat sebagai berikut.

#### 1) Data Selection

Pada tahap data selection ialah proses dalam pemilihan data dari banyaknya data-data, sebelum penggalian informasi dimulai maka tahap ini perlu dilakukan [10]. Data *vendor* akan melalui proses seleksi terlebih dahulu lalu data hasil seleksi ini digunakan untuk suatu proses *data mining*, selanjutnya disimpan dalam suatu file yang terpisah dari basis data operasional.

#### 2) Praprocessing Data

Selanjutnya pada tahap ini ialah proses yang dilakukan sebelum masuk dalam proses *data mining* menggunakan teknik klasifikasi, dimana sebagian besar data ialah data mentah yang kotor maksudnya data tidak lengkap, adanya noise atau tidak konsisten [11]. Maka, perlunya dilakukan *preprocessing* data karena apabila datanya tidak berkualitas akan menghasilkan juga kualitas mining yang tidak baik sehingga dilakukannya data *cleaning* untuk menghasilkan data mining yang berkualitas., pada proses ini dilakukannya pembersihan data yang kosong atau hilang, serta data yang mengandung kesalahan.

#### 3) Transformation Data

Pada tahap ini dilakukannya proses perubahan data menjadi bentuk data yang sesuai [10]. Dalam melakukan transformasi data pada penelitian ini juga dengan menggunakan pendekatan transformasi data ialah proses normalisasi dimana sebuah atribut numerik diskalakan dalam range yang lebih kecil misalnya -1.0 sampai 1.0 atau 0.0 sampai 1.0. Pada Penelitian ini transformasi data ini menggunakan Normalisasi min-max dengan skala range 0-1.

### 4) Data Mining

Data Mining disini merupakan proses ekstraksi data menjadi informasi baru, dengan menggunakan teknik yang tepat proses *data mining* akan mendapatkan hasil yang optimal, dapat juga diartikan sebagai proses penemuan pola pada data dalam database besar. Pengaplikasian *data mining* ini digunakan dalam bidang Marketing dan Bisnis yaitu Recommender System dalam teknik klasifikasi, sehingga hasil luaran dari teknik ini berupa rekomendasi dari data yang telah diolah dengan sesuai jarak, quantity, harga, dan total.

- Teknik Klasifikasi

Penelitian ini dengan menggunakan teknik klasifikasi dimana dilakukannya klasifikasi pada data *vendor* dengan mencoba menemukan hubungan antara atribut masukan dan atribut target. Semakin besar suatu data yang diproses atau diolah maka semakin besar juga waktu pemrosesannya. Maka, dalam klasifikasi ini dilakukan pemilihan rekomendasi *vendor* yang tepat performansinya [12]. Sehingga dalam penelitian ini pengklasifikasian *vendor* dengan menggunakan metode K-Nearest Neighbor & Weighted Product.

- Metode K-Nearest Neighbor

Pada penelitian ini klasifikasi yang digunakan ialah K-Nearest Neighbor. Algoritma ini pengklasifikasiannya berdasarkan data latih atau data uji, untuk melihatnya dari jarak yang sangat dekat dengan objek tersebut berdasarkan nilai k [13]. Sebelum dilakukannya perhitungan pada metode ini maka langkah pertama harus dapat menentukan data latih serta data uji. Selanjutnya menggunakan rumus *euclidean distance* sebagai dilakukannya proses perhitungan untuk mencari jarak. Dan sebelum dilakukannya proses mencari jarak data pada tetangga dengan menentukan nilai k tetangga, dan mendefinisikan suatu jarak menggunakan rumus *euclidean distance*, dimana antara dua titik yaitu titik data latih dan titik data uji [14]. Kelebihan dari metode ini menghasilkan data yang efektif dan jelas [15]. Berikut merupakan langkah-langkah dalam K-NN [16].

- Menentukan data latih dan data uji
- Menentukan jumlah pada tetangga k.
- Menghitung jarak objek dengan masing-masing data kelompok. Perhitungan jarak menggunakan rumus *euclidian distance* (Sreemathy dan Balamurungan, 2012). yang ditunjukkan pada persamaan 1:

$$d(x, y) = \sqrt{\sum_{i=1}^m (X_i - Y_i)^2} \quad (1)$$

Dimana D adalah jarak, x dan y adalah data latih dan data uji.

- Urutkan data yang mempunyai jarak terkecil ke terbesar.
- Lalu didapatkan hasil pengklasifikasian.

Pada saat dilakukannya proses perhitungan metode K-Nearest Neighbor ini memiliki cara untuk menentukan jarak dengan klasifikasi kelas rekomendasi dan tidak rekomendasi.

- Metode *Weighted Product*

Metode *Weighted Product* ialah metode pengambilan keputusan dengan cara perkalian untuk menghubungkan suatu rating atribut, yang dimana pada tiap nilai atribut terlebih dahulu harus dipangkatkan dengan bobot atribut yang berkaitan [17]. *Weighted Product* merupakan salah satu analisis keputusan multi-kriteria atau MCDA. Pada MCDA yang disediakan ialah kesimpulan keputusan alternatif terbaik yang dijelaskan mengenai beberapa kriteria keputusan [18]. Untuk setiap alternatif keputusan dibandingkan dengan alternatif keputusan yang lain dengan mengalikan jumlah rasio, satu untuk setiap kriteria keputusan. Setiap rasio dilakukan penaikan pangkat agar setara dengan bobot relatif dari kriteria yang telah sesuai. Metode *Weighted Product* juga memerlukan proses normalisasi di karenakan memperbanyak hasil penilaian dari setiap atribut. Pada hasil perkalian belum memiliki makna apabila belum dibandingkan dengan nilai standar. Bobot untuk atribut hasil kerja berfungsi sebagai ranking positif diproses perkalian, sedangkan pada bobot efisiensi biaya berfungsi sebagai ranking negatif. Berikut merupakan langkah-langkah *Weighted Product* [19].

- Langkah pertama yaitu menentukan alternatif pilihan.
- Tentukan kriteria yang digunakan sebagai acuan dalam pemilihan *vendor*.
- Tentukan rating kecocokan tiap alternatif setiap kriteria yang sebelumnya telah ditentukan.
- Selanjutnya tentukan bobot preferensi tiap-tiap kriteria.
- Setelah menentukan alternatif pilihan, kriteria serta bobot preferensinya maka langkah selanjutnya yaitu mengalikan semua bobot atribut/alternatif pilihan lalu jumlahkan hasil perkalian
- Bagi nilai V untuk setiap alternatif dengan nilai untuk tiap alternatif
- Menghasilkan urutan terbaik

### 5) Interpretation / Evaluation

Pada tahap ini ialah proses penerjemahan pola-pola pada hasil proses data mining, sehingga pola-pola ini perlu ditampilkan atau diperlihatkan secara mudah agar dapat dimengerti oleh pihak yang berkepentingan[20].

## III. HASIL DAN DISKUSI

### A. Data Selection

Dalam tahap data selection ini dari jumlah total dataset 1066 record dan 9 kolom atau 9 atribut yang terdapat dalam dataset seperti nama vendor, kota, alamat, barang, jarak, quantity, harga, total, dan label rekomendasi perlu dilakukan data selection pada

atribut-atribut atau kolom yang ada pada dataset menjadi atribut Jarak, Quantity, Harga, dan Total. Atribut-atribut yang digunakan harus memiliki tipe numeric dengan tujuan agar data dapat dihitung berdasarkan formula K-Nearest Neighbor. Berikut merupakan atribut yang digunakan untuk proses klasifikasi vendor dapat dilihat pada Tabel 1. Atribut Data.

TABEL 1. ATRIBUT DATA YANG DIGUNAKAN

NO	Atribut	Keterangan
1	Jarak Lokasi	Menjelaskan tentang jarak lokasi perusahaan ke lokasi para <i>vendor</i> . Dengan atribut jarak ini menjadi pertimbangan dalam waktu pembelian.
2	Quantity	Menjelaskan tentang banyak atau sedikitnya ketersediaan barang. Ini pun menjadi pertimbangan dikarenakan akan berpengaruh pada harga barang tersebut.
3	Harga ( <i>price</i> )	Menjelaskan tentang terjangkau atau tidak terjangkaunya harga bahan baku per satuan dari para <i>vendor</i> . Pada atribut harga ini merupakan pertimbangan yang penting.
4	Total	Menjelaskan tentang keseluruhan harga beli.

### B. Praprocessing Data

Pada proses *praprocessing* data ini maka akan dilakukannya proses pembersihan data yang kosong atau hilang, serta data yang mengandung kesalahan. Pada penelitian ini menggunakan data pembelian dari bulan Juni 2019 hingga bulan Agustus 2019, jumlah total data mentah ini sebanyak 1199 *record*, namun karena adanya data kosong dan data yang mengandung kesalahan maka setelah dilakukannya data *cleaning* jumlah total yang digunakan yaitu menjadi sebanyak 1066 *record*. Data mentah ini diubah menjadi dataset. Dataset terbagi menjadi 2 yaitu data latih dan data uji. Jumlah total dataset yang dipakai ialah sebanyak 1066 *record* dengan 9 kolom. Pada data latih dengan jumlah data sebanyak 836 *record*, dan data uji sebanyak 230 *record*. Maka, dataset ini digunakan untuk memperoleh informasi mengenai *vendor-vendor* yang direkomendasikan atau tidak direkomendasikan. Berikut merupakan contoh data pembelian sebelum dilakukannya proses data *cleaning* dan sesudah dilakukannya data *cleaning* dapat dilihat pada Gambar 2. Data *Vendor*.

Gambar 2. Data *Vendor* sebelum data *cleaning*

Dapat dilihat pada Gambar 2. Data *Vendor* terdapat banyaknya data kosong dan data yang mengandung kesalahan, maka dari itu dilakukannya pembersihan data dapat dilihat pada Gambar 3.

Gambar 3. Data vendor setelah data cleaning

### C. Transformation Data

Dataset diproses dengan cara membagi total record dataset menjadi 8:2, jumlah total dataset yaitu sebanyak 1066 record dimana 80% jumlah dataset akan menjadi data latih sebanyak 836 record, dan 20% jumlah dataset akan menjadi data uji sebanyak 230 record. Pada data vendor ini bukan merupakan data bertipe data *numeric* sehingga perlu dilakukannya transformasi data. Maka, pada tahap Transformasi Data ini dilakukan untuk mengubah data dengan tipe data *Character* menjadi *Numeric*. Pada Transformasi data ini dengan menggunakan Normalisasi Data Min-Max.

- Normalisasi Min-Max

Pada tahap ini agar rentang nilai seimbang pada setiap atribut-atributnya yang digunakan maka harus dilakukannya normalisasi data. Untuk mendapatkan hasil nilai dengan rentang 0-1, maka *praprosesing* data pada atribut-atribut yang digunakan yaitu dengan menggunakan normalisasi data min-max. Pada normalisasi min-max melakukan transformasi *linier* terhadap data asli dengan persamaan (2) berikut.

$$N_x = \frac{MinRange + (X - MinValue)(MaxRange - MinRange)}{MaxValue - MinValue} \quad (2)$$

Diketahui :  $MinRange = 0$  ,  $MaxRange = 1$

Berikut merupakan dilakukannya transformasi data dengan menggunakan normalisasi data min-max pada Tabel. 2 Data Mentah.

TABEL 2. DATA MENTAH

No	Nama Vendor	Jarak	Qty	Harga	Total	Keterangan
1.	DUNIA GARMENT ACCESSORIE S, CV	39km	24 PCS	19.800,00	475.200	Tidak Rekomendasi
2.	SRI INDAH LABETAMA, PT	26km	190 PCS	16.250,00	3.087.500	Tidak Rekomendasi
3.	FORTUNINDO	54km	100 PCS	26.100,00	2.610.000	Tidak Rekomendasi
4.	YES, TOKO	5km	210 PCS	13.400,00	2.814.000	Rekomendasi
5.	CIPTA CEKAS	20km	245 PCS	8.100,00	1.984.500	Rekomendasi
6.	TERUNA PERKASA, PO	17 km	217 PCS	4.900,00	1.063.300	Rekomendasi
7.	MAJU JAYA	61km	49 PCS	12.500,00	612.500	Tidak Rekomendasi

No	Nama Vendor	Jarak	Qty	Harga	Total	Keterangan
8.	LOTUS PERMAL, PT	39km	108 PCS	17.150,00	1.852.200	Tidak Rekomendasi
9.	HIDUP BARU, CV	65km	97 PCS	23.700,00	2.298.900	Tidak Rekomendasi
10.	VICTORIA LABEL, PT	14 km	319 PCS	9.200,00	2.934.800	Rekomendasi

Maka didapatkan tabel nilai hasil normalisasi data min-max, dapat dilihat pada Tabel 3. Normalisasi

TABEL 3. NORMALISASI DATA MIN-MAX

No	Nama Vendor	Jarak	Qty	Harga	Total	Keterangan
1.	DUNIA GARMENT ACCESSORIE S, CV	0,58	0,003	0,70	0,00000003	Tidak Rekomendasi
2.	SRI INDAH LABETAMA, PT	0,36	0,56	0,53	1	Tidak Rekomendasi
3.	FORTUNINDO	0,83	0,26	1	0,82	Tidak Rekomendasi
4.	YES, TOKO	0,01	0,63	0,40	0,90	Rekomendasi
5.	CIPTA CEKAS	0,26	0,75	0,15	0,58	Rekomendasi
6.	TERUNA PERKASA, PO	0,21	0,65	0,000004	0,22	Rekomendasi
7.	MAJU JAYA	0,95	0,08	0,36	0,05	Tidak Rekomendasi
8.	LOTUS PERMAL, PT	0,58	0,28	0,58	0,53	Tidak Rekomendasi
9.	HIDUP BARU, CV	1	0,25	0,89	0,70	Tidak Rekomendasi
10.	VICTORIA LABEL, PT	0,16	1	0,20	0,94	Rekomendasi

### D. Data Mining

Pengaplikasian data mining ini digunakan dalam bidang Marketing dan Bisnis yaitu *Recommender System* dengan teknik klasifikasi, sehingga hasil luaran dari teknik ini berupa rekomendasi dari data yang telah diolah dengan atribut jarak, quantity, harga, dan total. Berikut merupakan metode K-Nearest Neighbor & Weighted Product.

#### 1) Metode K-Nearest Neighbor

Pada metode K-Nearest Neighbor memiliki cara untuk menentukan jarak terdekat. Data yang telah dilakukan preprocessing akan menjadi dataset untuk data yang akan diuji dengan mengukur kedekatannya dengan data yang ada, perhitungan *distance* yang akan digunakan adalah *euclidian distance*. Perhitungan K-NN ini menggunakan dataset sebanyak 1066 record dimana data latih sebanyak 836 record dan data uji sebanyak 230 record.

- Menentukan nilai K tetangga terdekat dengan K-Cross Validation.

Model atau algoritma dilatih oleh subset pembelajaran dan divalidasi oleh subset validasi. Selanjutnya pemilihan jenis Cross Validation dapat didasarkan pada ukuran dataset. Dataset ini dibagi menjadi 2 data latih sebanyak 80% dan data uji 20%. K-Cross Validation digunakan karena dapat mengurangi waktu komputasi dengan tetap menjaga keakuratan estimasi. Maka dari itu berikut merupakan simulasi dari K-Fold Cross Validation pada Gambar 4.



Gambar 4. K-Cross Validation

Setelah melakukan k-cross validation dimana penentuan nilai K tetangga terdekat secara akurat, maka berikut didapatkan hasil K-cross validation nilai K tetangga terdekat sebanyak 5, dapat dilihat pada Gambar 5. Hasil K-Cross.

no	K	Akurasi
1	3	74.104113300493%
2	5	75.461822660099%
3	7	69.873768472906%
4	9	65.57881773399%

Gambar 5. Hasil K-Cross Validation

Selanjutnya ialah memasukkan data baru dengan menginputkan jarak, quantity, harga, total untuk mendapatkan informasi mengenai vendor. Dengan nilai K yang telah didapat sebanyak 5, untuk mencari jarak terdekat yaitu menggunakan rumus *euclidian distance* pada persamaan 1. Berikut merupakan hasil klasifikasi vendor dengan K-NN, dapat dilihat pada Gambar 6 dan Gambar 7.

#	Nama vendor	Barang	Jarak	Qty	Harga	Total	Kelas	Ranking (weight product)
1	FASILAMINDO	JASA LAMINATING KAIN + BUSA 41 HL 4 MM + TRICOT PUTH (BLESS)	20	3	7400	21460	1	0.26282613773608
2	FASILAMINDO	JASA LAMINATING KAIN + BUSA 41 HL 4 MM + TRICOT PUTH (BLESS)	20	1	7400	10340	1	0.18877057678467
3	OBRAS KILAT, TOKO	VETERBAND UK 1 CM PANAH PUTH	23	481	133	64132	1	0.2196543661513
4	DUNIA BENANG	CENTANG NO. 22 UK 15 MM ANTIQ BRASS	23	1000	100	100000	1	0.24826043898915
5	OBRAS KILAT, TOKO	STRINGPIN 3 HITAM	23	1000	12	12000	1	0.000486409874973

Gambar 6. Hasil klasifikasi K-NN

Sehingga kesimpulan dari klasifikasi dengan algoritma K-Nearest Neighbor ini menghasilkan nilai akurasi 75.46% dengan nilai tetangga terdekat sebanyak K = 5.

Jumlah K	Akurasi
5	75.461822660099%

Gambar 7. Nilai K dan Akurasi Metode

## 2) Metode Weighted Product

Agar hasil klasifikasi yang didapatkan maksimal maka hasil klasifikasi yang didapatkan dengan metode K-Nearest Neighbor ini diranking dengan menggunakan metode Weighted Product. Pada proses perhitungan Weighted Product dengan cara perkalian untuk menghabungkan suatu rating atribut, dimana pada tiap nilai atribut terlebih dahulu harus dipangkatkan dengan bobot atribut yang berkaitan. Berikut merupakan hasil klasifikasi Weighted Product, dapat dilihat pada Gambar 8. Hasil Perangkingan.

#	Nama vendor	Barang	Jarak	Qty	Harga	Total	Kelas	Ranking (weight product)
1	FASILAMINDO	JASA LAMINATING KAIN + BUSA 41 HL 4 MM + TRICOT PUTH (BLESS)	20	3	7400	21460	1	0.26282613773608
4	DUNIA BENANG	CENTANG NO. 22 UK 15 MM ANTIQ BRASS	23	1000	100	100000	1	0.24826043898915
3	OBRAS KILAT, TOKO	VETERBAND UK 1 CM PANAH PUTH	23	481	133	64132	1	0.2196543661513
2	FASILAMINDO	JASA LAMINATING KAIN + BUSA 41 HL 4 MM + TRICOT PUTH (BLESS)	20	1	7400	10340	1	0.18877057678467
5	OBRAS KILAT, TOKO	STRINGPIN 3 HITAM	23	1000	12	12000	1	0.000486409874973

Gambar 8. Hasil Perangkingan Weighted Product

Hasil metode Weighted Product ini meranking dari hasil klasifikasi terbesar hingga terkecil sehingga ini merupakan keluaran dari klasifikasi K-NN dan WP.

## E. Interpretation / Evaluation

Pada interpretation atau Evaluation ini proses untuk mengetahui tingkat keberhasilan klasifikasi algoritma K-Nearest Neighbor & Weighted Product dari sistem yang telah dibuat.

Pengujian akurasi ini menggunakan Confusion Matrix dimana sebuah matrik dari prediksi yang akan dibandingkan dengan kelas yang asli dari data inputan.

	Rekomendasi	Tidak Rekomendasi
Rekomendasi	93	33
Tidak Rekomendasi	28	76

Gambar 9. Pengujian Akurasi dengan Confusion Matrix

Pengujian akurasi ini dilakukan menggunakan data uji sebanyak 230 record, 126 record data rekomendasi dan 104 record data Tidak direkomendasi. Hasil pengujian akurasi metode K-Nearest Neighbor dengan Confusion Matrix dapat dilihat pada Gambar 9. Pengujian akurasi dan Tabel 4.

TABEL 4. PENGUJIAN AKURASI CONFUSION MATRIX

Kelas	Accuracy	Precision	Recall	F1-Score
Rekomendasi	73,48%	$0.74 \times 100\% = 74\%$	$0.77 \times 100\% = 77\%$	$0.75 \times 100\% = 75\%$
Tidak Rekomendasi	73,48%	$0.73 \times 100\% = 73\%$	$0.70 \times 100\% = 70\%$	$0.71 \times 100\% = 71\%$

Dari hasil pengujian akurasi dengan Confusion Matrix ini dapat dikatakan berhasil, namun persentase yang dihasilkan cukup rendah tidak tercapai hingga persentase 100%, dikarenakan beberapa faktor yang mempengaruhi dalam proses data mining menggunakan algoritma K-NN dan WP.

#### IV. KESIMPULAN

Maka berdasarkan hasil penelitian yang dilakukan untuk klasifikasi data vendor sesuai jarak, quantity, harga, total yang dapat menghasilkan suatu keputusan yang akan diambil dengan melihat hasil klasifikasi vendor. Dengan menghitung akurasi metode & nilai K menggunakan K-Cross Validation sebanyak 5 dapat disimpulkan dengan menggunakan metode K-NN dan Weighted Product ini lebih efektif dan akurat dengan nilai akurasi yang didapatkan sebesar 75.46%. Hasil pengujian akurasi dengan Confusion Matrix dihasilkan Accuracy 73,48%, Precision 74%, Recall 77%, F1-Score 75%. Namun persentase yang dihasilkan cukup rendah tidak tercapai hingga persentase 100%, dikarenakan beberapa faktor yang mempengaruhi diantaranya jumlah data yang kurang banyak dan varian serta atribut yang digunakan kurang banyak.

#### DAFTAR PUSTAKA

[1] Neighbor, "Penerapan data mining untuk prediksi penjualan produk elektronik terlaris menggunakan metode k-nearest neighbor," 2018.  
 [2] N. N. Dzirkulloh and B. D. Setiawan, "Penerapan Metode K - Nearest Neighbor (knn) dan Metode Weighted Product (wp) Dalam Penerimaan Calon Guru Dan Karyawan Tata Usaha Baru Berwawasan Teknologi (Studi Kasus : Sekolah Menengah Kejuruan Muhammadiyah 2 Kediri)," *Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 1, no. 5, pp. 378–385, 2017.

[3] D. eka Riska, Marji, "Klasifikasi Suara Berdasarkan Gender Dengan Metode Knn," *Cybrarians J.*, no. 37, pp. 1–31, 2015, doi: 10.12816/0013114.  
 [4] A. Fahrnasrudin, "Sistem Rekomendasi Pembelian Mobil Menggunakan Metode WWeighted Product (WP)," 2014.  
 [5] X. Wang, X. Ding, A. K. H. Tung, and Z. Zhang, "Efficient and effective knn sequence search with approximate ngrams," *Proc. VLDB Endow.*, vol. 7, no. 1, pp. 1–12, 2013, doi: 10.14778/2732219.2732220.  
 [6] M. Behnam and H. Pourghassem, "Singular Lorenz Measures Method for seizure detection using Knn-Scatter Search optimization algorithm," *2015 Signal Process. Intell. Syst. Conf. Sp. 2015*, pp. 67–72, 2016, doi: 10.1109/SPIS.2015.7422314.  
 [7] E. N. S. Liani, M. Kholil, and S. Safitri, "Pemilihan Alternatif Supplier Menggunakan Pendekatan Vendor Performance Indicator (vpi) Dan Metode Analytical Hierarchy Process (ahp) Di PT. Sumber Berkat Anugerah Indonesia," *Issn 2338-7122*, pp. 13–14, 2017.  
 [8] E. Rahmawati, S. Hadianti, M. F. Akbar, and W. Gata, "Penerapan Algoritma Cuntuk Memprediksi Performa Vendor Online," *Semin. Nas. Teknol. Inf. Univ. Ibn Khaldun Bogor 2018*, pp. 224–231, 2018.  
 [9] Kurniawan, "Data Mining," *J. Chem. Inf. Model.*, vol. 53, no. 9, pp. 8–24, 2014.  
 [10] F. Marisa, "Educational Data Mining (Konsep dan Penerapan)," *Teknol. Inf.*, vol. 4, no. 2, pp. 90–97, 2014, [Online]. Available: <http://ejournal.stimata.ac.id/index.php/TI/article/view/108/148>.  
 [11] K. D. Anda, "data preprocessing," pp. 1–18.  
 [12] S. Bice *et al.*, "Landasan Data Mining," *Resour. Policy*, vol. 7, no. 1, pp. 1–10, 2017.  
 [13] L. A. R. Hakim, A. A. Rizal, and D. Ratnasari, "Aplikasi Prediksi Kelulusan Mahasiswa Berbasis K-Nearest Neighbor (K-NN)," *JTIM J. Teknol. Inf. dan Multimed.*, vol. 1, no. 1, pp. 30–36, 2019, doi: 10.35746/jtim.v1i1.11.  
 [14] A. Nugraha, pratama dwi., Said al faraby, "Klasifikasi Dokumen Menggunakan Metode Knn Dengan Information Gain," *eProceedings Eng.*, vol. 5, no. 1, pp. 1541–1550, 2018.  
 [15] J. T. Elektro, F. Teknik, and U. N. Semarang, "Implementasi metode K-nearest neighbor," 2019.  
 [16] Kpm, R. A. Ramadhani, and e D. LiceFrense, "Buku K-Nears Neighbours," *Buku knn*, vol. 84, no. December, pp. 487–492, 2013.  
 [17] S. Lestari, "Penerapan Metode Weighted Product Model Untuk Seleksi Calon Karyawan," *J. Sist. Inf.*, vol. 5, no. 1, pp. 540–545, 2013.  
 [18] R. Roni, S. Sumijan, and J. Santony, "Metode Weighted Product dalam Pemilihan Penerima Beasiswa Bagi Peserta Didik," *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 3, no. 1, pp. 87–93, 2019, doi: 10.29207/resti.v3i1.834.  
 [19] J. B. Informatika, "Sistem Pendukung Keputusan Untuk Seleksi Calon Ketua Badan Eksekutif Mahasiswa (BEM) STMIK Banjarbaru Dengan Metode Weighted Product (WP) Muhammad Arsyad," vol. 4, no. 1, pp. 51–59, 2016.  
 [20] L. Gitleman, "Penerapan Knowledge Management System Menggunakan Teknik Knowledge Data Discovery," *Pap. Knowl. Towar. a Media Hist. Doc.*, vol. 8, no. 2, pp. 1038–1055, 2016.