



Klasifikasi Kebutuhan Sparepart Dengan Algoritma K-Nearest Neighbor Untuk Meningkatkan Penjualan Sparepart

Virza Putra, Gatot Tri Pranoto, Fibi Eko Putra

Fakultas Teknik, Program Studi Teknik Informatika, Universitas Pelita Bangsa, Bekasi, Indonesia

Email: ¹ virzaputra20@gmail.com ²gatot.pranoto@pelitabangsa.ac.id ³fibi@pelitabangsa.ac.id

Email Penulis Korespondensi: virzaputra20@gmail.com

Pemenuhan suku cadang yang cukup akan menjadi faktor penunjang kepercayaan konsumen terhadap perusahaan. Pendekatan metode klasifikasi dapat diterapkan dalam menganalisis data untuk menerapkan data mining dengan metode klasifikasi terhadap kebutuhan sparepart yang dihasilkan dengan memanfaatkan data testing yang terdiri dari 100 *record* dataset dengan pembagian rasio 90% data latih (*data training*) dan 10% data uji (*data testing*). Implementasi model algoritma *K-Nearest Neighbour* pada data uji (*data testing*) sebanyak 100 objek data, mendapatkan hasil yang menunjukkan sebuah wawasan baru berupa klasifikasi tingkat kebutuhan ringan dan tinggi berdasarkan 2 kategori. No merupakan kategori kebutuhan ringan, terdiri dari 89 objek data, kategori Yes merupakan kategori kebutuhan tinggi. Evaluasi dan pengujian performa menggunakan aplikasi RapidMiner Sstudio mampu memberikan hasil yang optimal dengan skenario yang dimodelkan. Model algoritma ini memiliki nilai *Accuracy* sebesar *accuracy*: 93.00% +/- 6.40% (micro average: 93.00%).

Kata Kunci : Data Mining, *K-NN*, *Spare part*, *Bahan Baku*, *Klasifikasi*.

Abstract– Adequate supply of spare parts will be a supporting factor for consumer confidence in the company. The classification method approach can be applied in analyzing data to apply data mining with the classification method for spare parts needs generated by utilizing data testing consisting of 100 record datasets with a ratio of 90% training data (training data) and 10% test data (data testing). . Implementation of the K-Nearest Neighbor algorithm model on test data (data testing) of 100 data objects, obtaining results that show a new insight in the form of classification of low and high level needs based on 2 categories. No is a category of light needs, consisting of 89 data objects, the category Yes is a category of high needs. Performance evaluation and testing using the RapidMiner Sstudio application is able to provide optimal results with the scenarios that are modeled. This algorithm model has an Accuracy value of *accuracy*: 93.00% +/- 6.40% (micro average: 93.00%).

Keywords: Data Mining, *K-NN*, Spare parts, Raw Materials, Classification.

1. PENDAHULUAN

Penjualan merupakan suatu kegiatan primer ekonomi yang relatif penting dalam rangka pencapaian tujuan perusahaan. Secara umum penjualan dilakukan dengan mempunyai tujuan eksklusif yaitu mendatangkan suatu keuntungan kepada seseorang yang memasarkan produk-produk atau jasa-jasa tertentu [1]. Di dalam kegiatan operasionalnya perusahaan terbagi menjadi beberapa divisi dan setiap divisi memiliki kontribusi masing-masing. Salah satu divisi adalah penjualan suku cadang kendaraan, dimana divisi ini memberikan pelayanan terhadap kebutuhan pelanggan terkait komponen kendaraan yang dibutuhkan. Serta sebagai pendukung kebutuhan divisi perawatan kendaraan. Menghasilkan perkiraan penjualan suku cadang, merupakan salah satu faktor penting dalam bisnis ini, karena melakukan kontrol dan perencanaan persediaan, akan memberikan peran penting dalam daya saing perusahaan manapun yang menyediakan barang untuk pelanggan [2].

Banyaknya transaksi penjualan suku cadang yang terjadi berdampak pada sulitnya melakukan perkiraan untuk penyediaan serta distribusi suku cadang. Sebagai akibatnya menyebabkan ketidakstabilan persediaan suku cadang untuk memenuhi permintaan konsumen. Serta mengganggu proses perbaikan dari divisi perawatan kendaraan. Ketidakstabilan ini akan menghambat proses bisnis sehingga berkurang kemampuan pada pemenuhan permintaan sekaligus mengurangi kepercayaan konsumen. Dalam pemasaran suatu produk ada beberapa faktor akan sangat berpengaruh terhadap penjualan, salah satunya adalah strategi dalam penjualan. Jika suatu produsen memiliki stok barang yang terlalu banyak sedangkan penjualan tidak sebanding, maka yang di hasilkan adalah kerugian. [3].

Pemenuhan suku cadang yang cukup akan menjadi faktor penunjang kepercayaan konsumen terhadap perusahaan. Dalam memperkirakan penjualan suku cadang terdapat beberapa faktor yang harus diperhatikan terkait jenis pelayanan perawatan kendaraan yang dipilih, tipe kendaraan dari setiap konsumen, jumlah dari pembelian suku cadang berdasarkan cakupan waktu, harga dari setiap suku cadang yang dibeli serta jenis suku cadang yang dipilih konsumen. Dengan faktor tersebut perusahaan dapat meramalkan jumlah penjualan di bulan atau tahun berikutnya. Pesatnya teknologi memberikan banyak metode dalam penyelesaian suatu masalah terkait prediksi dan peramalan. Salah satu teknik yang dapat diaplikasikan ialah *machine learning*. Pada dasarnya *machine learning* mempunyai tujuan untuk mengkaji sebuah prosedur pemecahan dengan melakukan pembelajaran secara otomatis, dengan kontribusi yang sangat sedikit dilakukan oleh manusia pada umumnya. *Machine learning* dapat didefinisikan secara luas sebagai metode komputasi yang menggunakan pengalaman untuk meningkatkan kinerja atau membuat prediksi yang akurat [4].

Terdapat penelitian sebelumnya mengenai *regresi linier*, yang pernah dilakukan oleh Arief Kurniadi, Jasmir, Yudi Novianto. Dalam penelitian tersebut menggunakan data penjualan. Peneliti mengambil data transaksi penjualan pada

bulan Januari sampai dengan November sebanyak 150 data transaksi. Dalam melakukan analisisnya peneliti menggunakan tools *Weka*. Metode yang digunakan dalam penelitian ini adalah Regresi Linier Berganda. Setelah dilakukan analisis prediksi regresi linier pada tools *Weka* menggunakan *10-Fold Cross Validation*, maka di dapatlah akurasi tertinggi yaitu pesanan terbanyak adalah 7038 dengan rata-rata 625,367 nama obat *finpro box 30 tablet*, penjualan terbanyak adalah 6599 dengan rata-rata 628,753 nama obat *finpro box 30 tablet*, saldo terbanyak adalah 988 dengan rata-rata 106,967 nama *obat ottopan syrup 60 ml botol* [6].

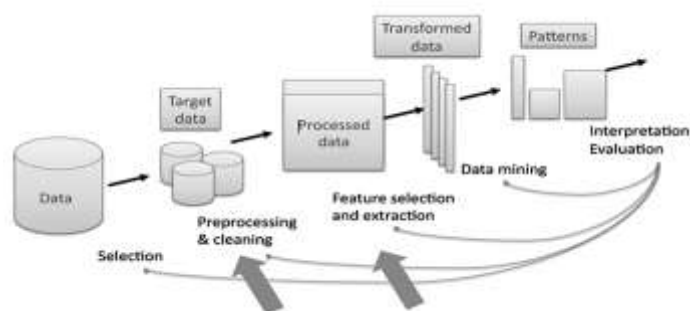
Dalam mengukur keakuratan dari hasil ramalan menggunakan *regerensi linier*, terdapat penelitian lain yang mendukungnya. Penelitian tersebut dilakukan oleh Firahmi Rizky, Yohanni Syahra, Ita Mariami, Yusnawati. Untuk dapat menyelesaikan permasalahan yang ada maka salah satu cara yang dapat dilakukan untuk memprediksi pemakaian stok barang adalah dengan menggunakan teknik data mining. Adapun teknik yang digunakan dalam hal ini adalah Algoritma Regresi Linier berganda. Regresi Linier berganda adalah Sebagai kajian terhadap ketergantungan satu variabel dengan variabel lain dengan tujuan membuat estimasi rata-rata nilai variabel yang sudah diketahui. Hasil dari penelitian ini adalah, Penjualan barang yang memiliki keterkaitan / hubungan yang diolah dengan teknik data mining menggunakan algoritma regresi linier berganda dapat membantu pihak Manajemen perusahaan dalam menentukan stok barang dibulan-bulan berikutnya [7].

Klasifikasi proses pengidentifikasian obyek ke dalam sebuah kelas atau kelompok berdasarkan atribut data yang akan digunakan yang bertujuan untuk menempatkan objek dan variabel data. Algoritma K-Nearest Neighbor salah satu algoritma yang dapat digunakan dalam *classifier* suatu dataset yang menjadi suatu pengetahuan dan informasi, teknik dari *classifier* bagian dari sebuah algoritma yang dapat diuji atau bagian dari data mining yang dapat diprediksi dari sebuah data diabetes (kencing manis), berdasarkan sekumpulan data dari atribut-atribut. Penelitian guna untuk mencari hasil prediksi yang nantinya akan dijadikan acuan yang mengenai pencegahan diabetes dan dapat dimanfaatkan untuk menentukan hasil yang baik dalam pengujian data. Hasil dari pengolahan data tersebut menjadi sebuah informasi dan pengetahuan yang diharapkan, sehingga dapat digali suatu potensi atau pengetahuan yang lebih baik atau akurat dalam pembacaan datanya, tepat dan cepat dari data tersebut sehingga dapat menganalisa pencegahan diabetes dan menemukan peluang-peluang yang baru serta menemukan rencana strategis dan untuk mencegah diabetes, selain itu bisa digunakan sebagai sarana untuk mengambil keputusan. Data mining merupakan bentuk penggalian data yang digunakan untuk menggali pengetahuan dari jumlah data yang besar. Berdasarkan uraian masalah diatas maka penelitian ini menggunakan salah satu metode dalam data mining yaitu dengan Algoritma K-Nearest Neighbor untuk menentukan klasifikasi data.

2. METODOLOGI PENELITIAN

2.1 Tahapan Penelitian

Dalam melakukan penelitian seorang peneliti memiliki pedoman yang secara bertahap yang Dalam melakukan analisa dan mencari pola data untuk dijadikan sebuah dataset dalam memudahkan penelitian dan dapat berjalan dengan sistematis dan memenuhi tujuan yang diinginkan maka dibuat langkah – langkah dalam tahapan penelitian yang akan dilakukan berikut :



Gambar 1 Tahapan Penelitian

Dalam pengujian data menggunakan teknik klasifikasi dan tahapan - tahapan pada data mining dengan algoritma k-means, pengolahan data dan yang akan dijadikan dataset dalam penelitian ini Dari data penjualan komputer tersebut akan menjadi dataset.

2.2 Pengujian Metode

K Nearest Neighbor (KNN) adalah metode untuk melakukan klasifikasi terhadap objek berdasarkan data pembelajaran yang jaraknya paling dekat dengan objek tersebut. Untuk pemilihan atribut terdiri dari n neighbors (biasa disebut k). parameter k pada testing ditentukan berdasarkan nilai k optimum pada saat training. Nilai k optimum diperoleh dengan mencoba-coba. Menghitung kuadrat jarak euclid (euclidean distance) masing-masing obyek terhadap data sampel yang diberikan.

$$D(a, b) = \sqrt{\sum_{k=1}^d (a_k - b_k)^2}$$

Berdasarkan rumus 4 dimana matriks D(a,b) adalah jarak skalar dari kedua vektor a (data latih) dan b (data uji) darimatriks denganukurand dimensi. Mengurutkanhasilobjek-objek tersebut ke dalam kelompok yang mempunyai jarak euclidian terkecil. Mengumpulkan kategori y (klasifikasi nearest neighbor berdasarkan nilai k) Dengan menggunakan kategori nearest neighbor yang paling mayoritas, maka dapat diprediksikan kategori objek tersebut.

3. HASIL DAN PEMBAHASAN

3.1 Evaluasi Algoritma K-Nearest Neighbour

Perhitungan algoritma K-NN dari penulis menggunakan data uji (*data testing*) sebanyak 12 *record* data yang didapat pada dataset sebelumnya.

Tabel 1 Data uji (*data testing*) yang digunakan dalam simulasi perhitungan

No.	Date	Type_Motor	Harga	Jenis_Service
1	Sampling 1	1	45000	0
2	Sampling 2	3	65000	0
3	Sampling 3	4	125000	0
4	Sampling 4	3	75000	0
5	Sampling 5	2	50000	1
6	Sampling 6	3	55000	0
7	Sampling 7	1	57000	0
8	Sampling 8	5	75000	1
9	Sampling 9	2	55000	0
10	Sampling 10	2	20000	1
11	Sampling 11	7	35000	1
12	Sampling 12	2	100000	0

Selanjutnya dalam menerapkan algoritma K-NN pada data uji (*data testing*) tersebut ditentukan dengan langkah-langkah seperti dibawah berikut:

1. Menentukan jumlah tetangga yang akan diperhitungkan, dinotasikan dalam (**k**), dalam hal ini penulis menentukan 3 tetangga terdekat atau dengan kata lain (**k = 3**)
2. Menghitung jarak dengan perhitungan *Euclidean Distance* dari masing-masing setiap tetangga terhadap objek target data uji (*data testing*), kemudian diurutkan dari yang terkecil sampai terbesar.
3. Menentukan objek target data uji (*data testing*) masuk dalam suatu kelompok dengan melihat nilai mayoritas sebagai hasil prediksi dari data uji tersebut.

Dengan langkah yang ditetapkan tersebut, maka berikut adalah perhitungan dari masing – masing objek data uji yang terdapat 12 *record* data pada Tabel 4.1 menggunakan persamaan rumus perhitungan jarak *Euclidean Distance* berikut

$$D = \sqrt{(x1 - x2)^2 + (y1 - y2)^2 + (z1 - z2)^2}$$

1. **Dataset samplin 13, dengan nilai x1 = 2, y1 = 195000, dan z1 = 0.**

Didapatkan hasil dari perhitungan jarak *Euclidean Distance* antara objek dataset ke 511 dengan 5 tetangga (k = 5) terdekat dapat dilihat dari Tabel 2 berikut.

Tabel 2 Euclidean Distance terhadap objek dataset sampling 13

Perhitungan Manual				Tinggi	Jarak Terhadap K	K=3
Date	Type_Motor	Harga	Jenis_Service	Klasifikasi		
Sampling 13	2	195000	0	No		
Sampling 1	1	45000	0	No	150000	
Sampling 2	3	65000	0	No	130000	
Sampling 3	4	125000	0	No	70000	1
Sampling 4	3	75000	0	No	120000	3
Sampling 5	2	50000	1	Yes	145000	
Sampling 6	3	55000	0	No	140000	
Sampling 7	1	57000	0	No	138000	
Sampling 8	5	75000	1	Yes	120000	
Sampling 9	2	55000	0	No	140000	
Sampling 10	2	20000	1	Yes	175000	
Sampling 11	7	35000	1	Yes	160000	
Sampling 12	2	100000	0	No	95000	2

Berdasarkan dari 3 tetangga terdekat yang diurutkan dari yang terkecil sampai dengan yang terbesar, keseluruhan tetangga memiliki label No, sehingga secara mayoritas dapat disimpulkan bahwa dataset spray ke 13 termasuk kedalam Katagori No.

2. **Dataset sampling 26, dengan nilai $x_1 = 6$, $y_1 = 25000$, dan $z_1 = 0$.**

Didapatkan hasil dari perhitungan jarak *Euclidean Distance* antara objek dataset ke 26 dengan 3 tetangga ($k = 3$) terdekat dapat dilihat dari Tabel 3 berikut.

Tabel 3 Euclidean Distance terhadap objek dataset sampling 26

Perhitungan Manual				Tinggi	Jarak Terhadap K	K=3
Date	Type_Motor	Harga	Jenis_Service	Klasifikasi		
Sampling 26	6	25000	0	No		
Sampling 14	2	600000	0	No	575000	
Sampling 15	2	65000	0	No	40000	
Sampling 16	2	65000	0	No	40000	
Sampling 17	2	60000	1	Yes	35000	
Sampling 18	4	25000	0	No	2	1
Sampling 19	2	50000	0	No	25000	
Sampling 20	2	45000	0	No	20000	
Sampling 21	2	82000	0	No	57000	
Sampling 22	2	172000	0	No	147000	
Sampling 23	2	15000	0	No	10000	3
Sampling 24	2	176000	0	No	151000	
Sampling 25	1	20000	0	No	5000	2

Berdasarkan dari 3 tetangga terdekat yang diurutkan dari yang terkecil sampai dengan yang terbesar, terdapat 3 tetangga memiliki label no, sehingga secara mayoritas dapat disimpulkan bahwa dataset sampling termasuk kedalam katagori no.

3. **Dataset sampling 36, dengan nilai $x_1 = 2$, $y_1 = 65000$, dan $z_1 = 0$.**

Didapatkan hasil dari perhitungan jarak *Euclidean Distance* antara objek dataset ke 36 dengan 3 tetangga ($k = 3$) terdekat dapat dilihat dari Tabel 4 berikut.

Tabel 4 Euclidean Distance terhadap objek dataset sampling 36

Perhitungan Manual				Tinggi	Jarak Terhadap K	K=3
Date	Type_Motor	Harga	Jenis_Service	Klasifikasi		
Sampling 39	2	65000	0	No		
Sampling 27	1	15000	0	No	50000	
Sampling 28	2	90000	0	No	25000	
Sampling 29	5	15000	0	No	50000	
Sampling 30	2	5000	1	Yes	60000	
Sampling 31	2	245000	0	No	180000	
Sampling 32	2	210000	0	No	145000	
Sampling 33	2	70000	0	No	5000	1
Sampling 34	1	50000	0	No	15000	
Sampling 35	4	55000	0	No	10000	2
Sampling 36	1	45000	0	No	20000	3
Sampling 37	7	35000	0	No	30000	
Sampling 38	1	195000	0	No	130000	

Berdasarkan dari 3 tetangga terdekat yang diurutkan dari yang terkecil sampai dengan yang terbesar, keseluruhan tetangga memiliki label no, sehingga secara mayoritas dapat disimpulkan bahwa dataset sampling termasuk kedalam katagori no.

Pengujian performa model klasifikasi dengan algoritma KNN pada aplikasi RapidMiner Studio didapatkan hasil seperti pada Gambar 1 di bawah ini.



Gambar 1 Hasil evaluasi model KNN dengan Cross Validation

Dan berikut merupakan gambar AUC (*pessimistic*) yang dihasilkan dalam evaluasi model klasifikasi serta algoritma KNN pada aplikasi RapidMiner Studio



Gambar 2 AUC (*pessimistic*) pada aplikasi RapidMiner Studio

3.2 Pembahasan Hasil dan Pengujian

Dengan menggunakan data latih (*data training*) dan data uji (*data testing*) dari dataset yang dimiliki, maka tahapan model pada penerapan algoritma *K-Nearest Neighbour* menghasilkan prediksi terhadap objek – objek untuk mengelompokkan ke dalam label No, label Yes. Total data yang digunakan ialah sebanyak 850 record data, sehingga dengan komposisi tersebut maka terdapat 510 *record* yang dijadikan sebagai data latih (*data training*) dan 340 *record* dataset sebagai data uji (*data testing*).

Berikutnya dalam rasio yang digunakan pada skenario pemodelan, sebanyak 100 *record* data dijadikan data uji (*data testing*) dan diproses melalui aplikasi RapidMiner Studio. Hasil dari proses tersebut menunjukkan terdapat 89 objek yang masuk ke dalam klasifikasi kategori No atau status kebutuhan ringan, berikutnya ada sebanyak 11 objek yang masuk ke dalam klasifikasi kategori kebutuhan tinggi. Berikut merupakan grafik hasil dari klasifikasi pada aplikasi RapidMiner Studio.



Gambar 3. Grafik sebaran klasifikasi pada aplikasi RapidMiner Studio

Pembahasan dilakukan guna untuk mendapatkan nilai *accuracy*, *precision*, *recall* untuk algoritma K-NN. Algoritma K-NN, akurasi dilakukan dengan cara jumlah TP + TN dibagi jumlah total data testing yang diuji.

Tabel 4 Accuracy

accuracy: 93.00% +/- 6.40% (micro average: 93.00%)

	true No	true Yes	class precision
pred. No	88	6	93.62%
pred. Yes	1	5	83.33%
class recall	98.88%	45.45%	

Nilai precision dihitung dengan cara membagi jumlah data benar yang bernilai positif (True Positive) dibagi dengan jumlah data benar yang bernilai positif (True Positive) dan data salah yang bernilai positif (False Positive).

Tabel 5 Precision

precision: 83.33% (positive class: Yes)

	true No	true Yes	class precision
pred. No	88	6	93.62%
pred. Yes	1	5	83.33%
class recall	98.88%	45.45%	

Nilai recall dihitung dengan cara membagi data benar yang bernilai positive (True Positive) dengan hasil penjumlahan dari data benar yang bernilai positif (True Positive) dan data salah yang bernilai negatif (False Negative).

Tabel 6 Recall

recall: 45.00% +/- 47.17% (micro average: 45.45%) (positive class: Yes)

	true No	true Yes	class precision
pred. No	88	6	93.62%
pred. Yes	1	5	83.33%
class recall	98.88%	45.45%	

Hasil dari pengujian data testing yang diutarakan yaitu sebanyak 100 data , kemudian hasil dari data sparepart menyatakan tingkat ke *accuracy*, *recall* dan *persicion*, tingkat *accuarcy* 93.00 % , *recall* 98,88 % dan *percision* 93.69 % putus dalam klasifikasi. hasil dari penelitian ini merupakan rangkuman hasil percobaan dari algoritma tersebut dalam menghasilkan hasil optimal dalam klasifikasi spare part, mengetahui hasil perhitungan Analisa data dengan cara mengklasifikasikan data dalam bentuk label No dan Yes menggunakan algoritma K-NN menghasilkan hasil kebutuhan ringan 89 dan kebutuhan tinggi 11.

4. KESIMPULAN

Pendekatan metode klasifikasi dapat diterapkan dalam menganalisis data untuk menerapkan data mining dengan metode klasifikasi terhadap kebutuhan sparepart yang dihasilkan dengan memanfaatkan data testing yang terdiri dari 100 *record* dataset dengan pembagian rasio 90% data latih (*data training*) dan 10% data uji (*data testing*). Implementasi model algoritma *K-Nearest Neighbour* pada data uji (*data testing*) sebanyak 100 objek data, mendapatkan hasil yang menunjukkan sebuah wawasan baru berupa klasifikasi tingkat kebutuhan ringan dan tinggi berdasarkan 2 kategori. No merupakan kategori kebutuhan ringan, terdiri dari 89 objek data, katagori Yes merupakan kategori kebutuhan tinggi dan evaluasi dan pengujian performa menggunakan aplikasi RapidMiner Sstudio mampu memberikan hasil yang optimal dengan skenario yang dimodelkan. Model algoritma ini memiliki nilai *Accuracy* sebesar accuracy: 93.00% +/- 6.40% (micro average: 93.00%).

**REFERENCES**

- [1] V. Sihombing, "Sistem Informasi Penjualan Mobil Suzuki Di Dealer Bagan Batu," *Sistemasi*, vol. 7, no. 2, pp. 113–119, 2018.
- [2] E. Žunić, K. Korjenić, K. Hodžić, and D. Đonko, "Application of Facebook's Prophet Algorithm for Successful Sales Forecasting Based on Real-world Data," *Int. J. Comput. Sci. Inf. Technol.*, vol. 12, no. 2, pp. 23–36, 2020, doi: 10.5121/ijcsit.2020.12203.
- [3] Siska Delima, "Prediksi Penyediaan Sparepartkendaraan Roda Dua Dengan Metode Naive Bayes(Studi Kasus : Toko Dewi Motor)," *Jupersatek*, vol. 3, no. 2, pp. 720–726, 2020.
- [4] M. Mohri, A. Rostamizadeh, and A. Talwalkar, *Foundations of machine learning*. MIT press, 2018.
- [5] M. Rianto and R. Yunis, "Analisis Runtun Waktu Untuk Memprediksi Jumlah Mahasiswa Baru Dengan Model Random Forest," *Paradig. - J. Komput. dan Inform.*, vol. 23, no. 1, pp. 99–105, 2021, doi: 10.31294/p.v23i1.9781.
- [6] I. Yenidogan, A. Cayir, O. Kozan, T. Dag, and C. Arslan, "Bitcoin Forecasting Using ARIMA and PROPHET," *UBMK 2018 - 3rd Int. Conf. Comput. Sci. Eng.*, no. February 2019, pp. 621–624, 2018, doi: 10.1109/UBMK.2018.8566476.
- [7] R. H. Kiswanto, "Spesifikasi Komputer Rakitan Berdasarkan Kebutuhan dan Anggaran Menggunakan Algoritma Backtracking," *J. Eksplora Inform.*, vol. 10, no. 1, pp. 1–12, 2020, doi: 10.30864/eksplora.v10i1.358.
- [8] E. E. Lavindi, W. Wijanarto, and A. Rohmani, "Aplikasi Hybrid Filtering Dan Naive Bayes Untuk Sistem Rekomendasi Pembelian Laptop," *JOINS (Journal Inf. Syst.*, vol. 4, no. 1, pp. 54–64, 2019, doi: 10.33633/joins.v4i1.2518.
- [9] J. Y. Marpaung, G. L. Ginting, and N. Silalahi, "Penerapan Metode Fuzzy Tsukamoto Dalam Penentuan Harga Laptop Bekas," *Build. Informatics, Technol. Sci.*, vol. 2, no. 2, pp. 115–126, 2020, doi: 10.47065/bits.v2i2.310.
- [10] Suyanto, *Data Mining*. Yogyakarta: Informatika, 2017.
- [11] Retno Tri vulandari, *Data Mining*. Yogyakarta: Gava Media, 2017.
- [12] G. widi N. Dicky Nofriansyah, *No Title*. Yogyakarta: Cv Budi Utama, 2017.
- [13] J. S. Informasi and F. Teknik, "Optimasi teknik klasifikasi modified k nearest neighbor menggunakan algoritma genetika," no. September, pp. 130–134, 2014.
- [14] M. Bagas, A. Darmawan, F. Dewanta, and S. Astuti, "Analisis Perbandingan Algoritma Decision Tree , Random Forest , dan Naive Bayes untuk Prediksi Banjir di Desa Dayeuhkolot Comparative Analysis of Decision Tree , Random Forest , and Naive Bayes Algorithm for Flood Prediction at Dayeuhkolot Village," vol. 9, no. 1, pp. 52–61.
- [15] P. E. Wiraswendro and H. Soetanto, "Application of Random Forest Classifier Algorithm in Indonesian Sign Language System (Sibi) Detection System," *Bit (Fakultas Teknol. Inf. Univ. Budi Luhur)*, vol. 19, no. 2, p. 75, 2022, doi: 10.36080/bit.v19i2.2043.
- [16] Y. Silalahi, Kristiani Desri., Murfi, Hendri., Satria, "Studi Perbandingan Pemilihan Fitur untuk Support Vector Machine pada Klasifikasi Penilaian Risiko Kredit," vol. 1, no. 2, pp. 119–136, 2017.