

Implementasi Algoritma K-Nearest Neighbor Untuk Rekomendasi Smartphone

Garvin Ardian¹; Teady Matius Surya Mulyana^{1*)}

1. Program Studi Informatika, Universitas Bunda Mulia, Banten 15143, Indonesia

*) Email: tmulyana@bundamulia.ac.id

Received: 21 Juli 2022 | Accepted: 23 April 2023 | Published: 25 April 2023

ABSTRACT

Since 2019, COVID-19 virus pandemic has caused almost all activities to take place at home. The COVID-19 pandemic has made people who stay at home bored and in need of entertainment. One of the entertainments that can be enjoyed at home is a smartphone. Smartphones have many types and models from various brands that make people confused in choosing the one that suits their individual needs. The purpose of this study is to help provide smartphone recommendations based on the needs of the community. The smartphone recommendation program will use the K-Nearest Neighbor algorithm with a value of $K = 3$ to determine the class of smartphone which is divided into 3 classes flagship using 4 parameters. Data will then be classified by calculating the Euclidean distance between the test data and the training data. The test will be performed using a confusion matrix to measure the level of accuracy in class classification from smartphones. From the results of the study, it can be concluded that the program can provide smartphone recommendations with an accuracy of 83%. Data needs to be updated every few years because there will be changes to the specifications for each class within a certain time.

Keywords: Smartphone, classification, K-nearest neighbor, Euclidean distance, confusion matrix

ABSTRAK

Sejak tahun 2019, pandemi virus COVID-19 membuat hampir semua aktivitas dilakukan di rumah. Pandemi COVID-19 membuat orang-orang yang tinggal di rumah bosan dan membutuhkan hiburan. Salah satu hiburan yang bisa dinikmati di rumah adalah smartphone. Smartphone memiliki banyak jenis dan model dari berbagai merek yang membingungkan orang dalam memilih yang sesuai dengan kebutuhan masing-masing. Tujuan dari penelitian ini adalah untuk membantu memberikan rekomendasi smartphone berdasarkan kebutuhan masyarakat. Program rekomendasi smartphone akan menggunakan algoritma K-Nearest Neighbor dengan nilai $K = 3$ untuk menentukan kelas smartphone yang terbagi menjadi 3 kelas flagship dengan menggunakan 4 parameter. Data tersebut kemudian akan diklasifikasikan dengan menghitung jarak Euclidean antara data uji dan data latih. Pengujian akan dilakukan dengan menggunakan confusion matrix untuk mengukur tingkat akurasi pada klasifikasi kelas smartphone. Dari hasil penelitian dapat disimpulkan bahwa program dapat memberikan rekomendasi smartphone dengan akurasi 83%. Data perlu diperbarui setiap beberapa tahun karena akan ada perubahan spesifikasi untuk setiap kelas dalam waktu tertentu.

Kata kunci: Smartphone, klasifikasi, K-Nearest Neighbor, Euclidean Distance, confusion matrix

1. PENDAHULUAN

Pada tahun 2019, pandemi virus Corona atau yang sering disebut COVID-19 melanda dunia yang menimbulkan penerapan *social distancing* dan membuat hampir seluruh aktivitas dilakukan dalam rumah. Di Indonesia, pemerintah mengeluarkan kebijakan Pemberlakuan Pembatasan Kegiatan Masyarakat (PPKM), dan *Work From Home* (WFH) sebagai tindakan untuk mencegah penyebaran COVID-19 yang semakin meluas[1].

Masyarakat yang melakukan *Work From Home* (WFH) merasa jenuh dan membutuhkan hiburan akibat pandemi COVID-19. Salah satu hiburan kesenangan mereka dengan bermain sosial media, bermain game, dan menonton film. Hal ini membuat penggunaan *smartphone* di Indonesia mencapai 167 juta orang atau 89% dari total penduduk di Indonesia[2].

Tujuan dari penelitian ini adalah algoritma K-NN mampu untuk mengklasifikasikan jenis-jenis *smartphone* berdasarkan harga, RAM, prosesor, kamera dan ruang penyimpanan dan aplikasi yang dibuat mampu untuk memberikan rekomendasi pembelian *smartphone* kepada *user*. Manfaat dari penelitian baik dari segi praktis ataupun teoritis adalah memberikan kemudahan bagi masyarakat untuk dapat memilih *smartphone* yang tepat dan memberikan referensi *smartphone* yang sesuai bagi masyarakat berdasarkan kebutuhan penggunaan.

Algoritma K-NN sudah banyak digunakan untuk klasifikasi di banyak bidang. Contohnya untuk klasifikasi pada lagu[3], untuk klasifikasi pada follower twitter yang menggunakan Bahasa tertentu[4], untuk klasifikasi pemilihan mobil[5], untuk klasifikasi tulang manusia[6], dan untuk pemilihan *smartphone* masih banyak yang menggunakan metode sistem pendukung keputusan[7][8][9] dan juga menggunakan algoritma lain seperti Fuzzy[10][11]. Untuk itu, peneliti akan melakukan klasifikasi pada *smartphone* dengan algoritma *K-Nearest Neighbor* guna membantu masyarakat.

2. METODE/RANCANGAN PENELITIAN

2.1. Identifikasi Masalah

Banyaknya penggunaan pada *smartphone* membuat banyak pula kelas dan jenis-jenis *smartphone* yang tersedia untuk memenuhi pasaran. Kelas *smartphone* yang sedang beredar dimulai dari kelas entry hingga flagship dengan harga dan spesifikasi yang telah ditentukan[12]. Banyaknya kelas ini membuat masyarakat bingung dalam menentukan pilihan mereka. Untuk itu, diperlukan rekomendasi *smartphone* yang dapat digunakan membantu masyarakat memilih *smartphone* yang sesuai dengan kelas dan spesifikasi yang diinginkan. Penelitian yang dilakukan untuk mengukur tingkat akurasi dari algoritma *K-Nearest Neighbor* dalam menentukan kelas *smartphone* dan juga dapat memberikan rekomendasi *smartphone*.

2.2. K-Nearest Neighbor

Algoritma K-Nearest Neighbor (KNN) adalah algoritma yang digunakan untuk melakukan klasifikasi terhadap suatu objek, berdasarkan k buah data latih yang jaraknya paling dekat dengan objek tersebut. Syarat nilai k adalah tidak boleh lebih besar dari jumlah data latih, dan nilai k harus ganjil dan lebih dari satu[4]. Berikut Langkah-langkah dari cara kerja algoritma *K-Nearest Neighbor*[13]:

1. Tentukan parameter K.
2. Hitung jarak antara data yang akan dievaluasi dengan semua data latih.
3. Urutkan jarak data.
4. Tentukan jarak terdekat sampai urutan K.
5. Pasangkan kelas yang bersesuaian.

6. Cari jumlah kelas dari tetangga yang terdekat dan tetapkan kelas tersebut sebagai kelas data yang dievaluasi.

2.3. Euclidean Distance

Euclidean Distance adalah perhitungan jarak untuk mengukur jarak antar dua titik dalam *euclidean space* yang mempelajari hubungan antara sudut dan jarak (Derisma, Firdaus, & Yusya, 2016). Rumus persamaan dari *Euclidean Distance* di tunjukkan pada persamaan 1[14]:

$$D = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \quad (1)$$

Dengan keterangan :

D = Jarak *Euclidean*

X₁ = Sampel data X

X₂ = Data uji / testing X

Y₁ = Sampel data Y

Y₂ = Data uji / testing Y

2.4. Confusion Matrix

Confusion Matrix merupakan metode yang digunakan dalam perhitungan akurasi pada data berbentuk mining (data mining). Contoh tabel *Confusion Matrix* yang dapat dilihat pada Tabel 1.

Tabel 1. *Confusion Matrix*

Aktual	Prediksi	
	TRUE	FALSE
TRUE	TP	FP
FALSE	FN	TN

Dengan Keterangan pada Tabel 1 sebagai berikut:

1. True Positive (TP): Jumlah data yang bernilai positif dan diprediksi benar positif.
2. True Negative (TN): Jumlah data yang bernilai negatif dan diprediksi benar negatif.
3. False Positive (FP): Jumlah data yang bernilai negatif tetapi diprediksi sebagai positif.
4. False Negative (FN): Jumlah data yang bernilai positif tetapi diprediksi sebagai negatif.

Perhitungan tingkat akurasi dengan menggunakan metode *confusion matrix* di tunjukkan pada persamaan 2 sebagai berikut[15] :

$$Accuracy = \frac{TP+TN}{TP+FP+FN+TN} \quad (2)$$

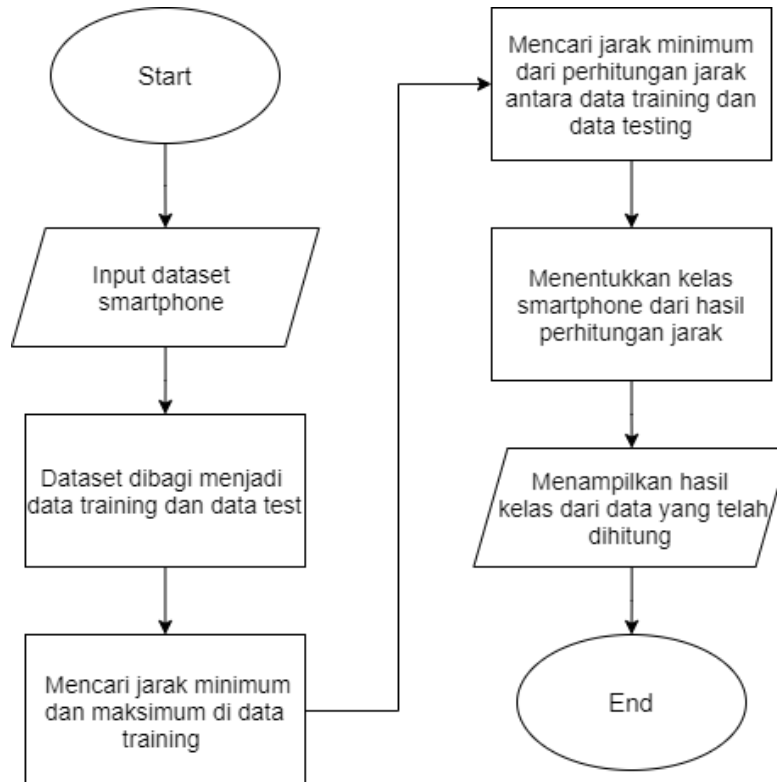
2.5. Impelentasi Algoritma *K-Nearest Neighbor*

Berdasarkan **Gambar 1** Algoritma *K-Nearest Neighbor* akan menggunakan data dari dataset *smartphone* yang meliputi besar kamera *smartphone*, besar ukuran ruang penyimpanan, besar ukuran RAM, dan harga *smartphone*.

Kemudian akan dibagi oleh sistem menjadi data *training* dan data *testing* dimana data testing yang diambil sejumlah 30% dari total data yang dimasukkan. Selanjtunya sistem akan menentukan jarak minimum dan jarak maksimum dari data *training* untuk tiap kelas yang telah di tentukan.

Setelah mendapatkan jarak maksimum dan minimum pada data *training* untuk setiap kelas selanjutnya data *testing* akan mencari jarak minimum yang paling sesuai dari setiap data dengan menggunakan parameter K adalah 3. Setelah mendapatkan 3 jarak minimum yang telah ditentukan, selanjutnya algoritma K-NN akan menentukan jenis klasifikasi dari hasil voting pada parameter K. Selanjutnya sistem akan menampilkan hasil klasifikasi dari data tersebut.

Alur proses *flowchart* dari sistem algoritma *K-Nearest Neighbor* yang di buat dapat dilihat pada Gambar 1.



Gambar 1. Alur Flowchart K-NN

2.6. Pengumpulan Data

Pengumpulan data untuk dataset smartphone di ambil dari salah satu website *e-commerce* “Tokopedia” dengan menggunakan web scrapping yang dimasukkan ke dalam format data csv dimana masing-masing kategori yang diambil sebagai berikut :

1. Storage : Menggunakan data dari ukuran ruang penyimpanan tanpa ruang penyimpanan eksternal yang tercantum di tokopedia.
2. Kamera : Menggunakan data dari kamera utama sebagai acuan untuk pengambilan data di tokopedia.
3. RAM : Menggunakan data dari RAM utama yang tercantum di tokopedia.
4. Harga : Menggunakan data dari harga asli tanpa diskon yang tercantum di tokopedia.

Hasil dari pengumpulan data yang telah di simpan dalam csv dapat dilihat pada tabel berikut ini :

1. Tabel Data *Training*

Tabel 2. Sampel Data Training

Nama Smartphone	Storage	Kamera	RAM	Harga	Kelas
Samsung Galaxy M62	256	64	8	5999000	Flagship
Vivo V21 5G	128	64	8	5799000	Flagship
Samsung Galaxy A52s 5G 8/128GB	128	64	8	5499000	Flagship
Realme GT Neo 3T	128	64	8	5499000	Flagship
Samsung Galaxy A52 8/256GB	256	64	8	5399000	Flagship
Samsung Galaxy A22 5G	128	48	6	3299000	High Level
Realme Narzo 50 5G	128	48	6	3299000	High Level
Vivo Y21T	128	50	6	3099000	High Level
OPPO A54	128	13	6	3099000	High Level
OPPO A53 6/128GB	128	13	6	3099000	High Level
Redmi 10A 3/64GB	64	13	3	1599000	Mid Level
Redmi 10A 3/32GB	64	13	3	1499000	Mid Level
Vivo Y01	32	13	2	1499000	Mid Level
Redmi 10A 3/32GB	32	13	3	1499000	Mid Level
Redmi 9A 2/32GB	32	13	2	1399000	Mid Level

2. Tabel Data Uji

Tabel 3. Sampel Data Uji

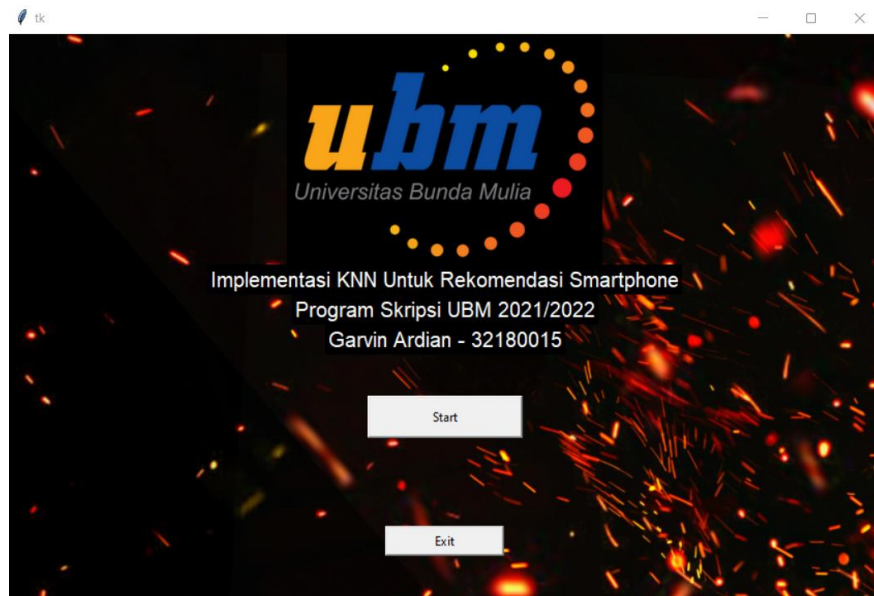
Nama Smartphone	Storage	Kamera	Ram	Harga
OPPO Reno 7 LTE	256	64	8	5199000
Realme 9 Pro +	256	50	8	5199000
Samsung Galaxy M32 8/128GB	128	64	8	3299000
Redmi Note 10 5G 8/128GB	128	48	8	3099000
Redmi 9A 3/32GB	32	13	3	1499000
Realme C11 (2021)	32	8	2	1349000

3. HASIL DAN PEMBAHASAN

3.1. Hasil

Hasil dari aplikasi yang telah dibuat untuk rekomendasi *smartphone* dengan algoritma *K-Nearest Neighbor* adalah sebagai berikut :

- Halaman Menu utama



Gambar 2. Halaman Menu Utama

Pada Gambar 2 menampilkan halaman utama yang akan pertama kali dilihat oleh *user*. Pada halaman ini terdapat gambar Logo UBM di bagian atas dan judul dari program skripsi yang dibuat yaitu “IMPLEMENTASI KNN UNTUK REKOMENDASI SMARTPHONE”. Selanjutnya terdapat nama peneliti bertuliskan “Garvin Ardian” dan NIM dari peneliti bertuliskan “32180015”. Setelah itu terdapat dua tombol di halaman utama aplikasi. Tombol *start* yang ketika di pilih akan berpindah ke halaman *input user*. Selain itu, ada tombol *exit* yang ketika dipilih akan mengakhiri program aplikasi rekomendasi *smartphone*.

- Halaman Menu Spesifikasi



Gambar 3. Halaman Menu Spesifikasi

Gambar 3 menampilkan halaman dari *user input* aplikasi yang akan diinput oleh *user* dengan empat pilihan spesifikasi *smartphone* yang nantinya akan digunakan untuk menentukan klasifikasi kelas dari *smartphone* tersebut. Setelah itu, terdapat tombol *search* yang akan menampilkan hasil

klasifikasi dari spesifikasi-spesifikasi yang telah dimasukkan dengan menggunakan algoritma K-Nearest Neighbor. Pada bagian bawah halaman terdapat tombol exit untuk mengakhiri program rekomendasi smartphone.

- Halaman Output Spesifikasi

Gambar 4. Halaman Output Spesifikasi

Gambar 4 menampilkan halaman output aplikasi yang merupakan sambungan dari halaman input. Halaman output ini akan muncul setelah user selesai memasukkan spesifikasi-spesifikasi yang diperlukan dan telah menekan tombol *search*. Pada halaman ini akan muncul sebuah *textbox* yang akan menampilkan hasil dari klasifikasi dari data *input* dengan menggunakan algoritma *K-Nearest Neighbor*. Selanjutnya terdapat tombol *list smartphone* yang digunakan untuk menampilkan rekomendasi smartphone dari kelas yang telah di klasifikasikan saat ditekan.

3.2. Pembahasan

Hasil dari pengujian klasifikasi yang dilakukan menggunakan python mendapatkan hasil yang ditunjukkan pada Tabel 4.

Tabel 4. Hasil Pengujian Data Uji

Hasil Pengujian Data Uji			
Nama Smartphone	Hasil Prediksi	Data Aktual	Kesesuaian
Samsung Galaxy S22 Ultra 5G 12/256GB	Flagship	Flagship	Sesuai
Samsung Galaxy S22+ 5G	Flagship	Flagship	Sesuai
Samsung Galaxy S22 5G	Flagship	Flagship	Sesuai
Vivo X70 Pro 5G	Flagship	Flagship	Sesuai
Vivo X60 Pro 5G	Flagship	Flagship	Sesuai
Vivo X60 5G	Flagship	Flagship	Sesuai
Samsung Galaxy A53 5G	Flagship	Flagship	Sesuai
Poco F3 8/256GB	Flagship	Flagship	Sesuai
OPPO Reno 6	High Level	Flagship	Tidak Sesuai

OPPO Reno 4	High Level	Flagship	Tidak Sesuai
OPPO Reno 7 LTE	High Level	Flagship	Tidak Sesuai
Realme 9 Pro +	High Level	Flagship	Tidak Sesuai
Samsung Galaxy A52 8/128GB	High Level	High Level	Sesuai
Samsung Galaxy M33 5G	High Level	High Level	Sesuai
Samsung Galaxy A32 5G 8/128GB	High Level	High Level	Sesuai
Redmi Note 11 Pro LTE 8/128GB	High Level	High Level	Sesuai
Redmi Note 10 Pro 8/128GB	High Level	High Level	Sesuai
Redmi Note 10 Pro 6/128GB	High Level	High Level	Sesuai
Samsung Galaxy A32 5G 6/128GB	High Level	High Level	Sesuai
Redmi Note 11 Pro LTE 6/128GB	High Level	High Level	Sesuai
Poco M4 Pro 8/256GB	High Level	High Level	Sesuai
Samsung Galaxy A23	High Level	High Level	Sesuai
Samsung Galaxy M32 8/128GB	High Level	High Level	Sesuai
Redmi Note 10 5G 8/128GB	High Level	High Level	Sesuai
OPPO A53 4/128GB	Mid Level	Mid Level	Sesuai
Redmi 10 6/128GB	Mid Level	Mid Level	Sesuai
Samsung Galaxy A13 LTE 4/128GB	Mid Level	Mid Level	Sesuai
Vivo Y21	Mid Level	Mid Level	Sesuai
OPPO A16K	Mid Level	Mid Level	Sesuai
Vivo Y15s 3/64GB	Mid Level	Mid Level	Sesuai
Realme C31 4/64GB	Mid Level	Mid Level	Sesuai
Samsung Galaxy A03s	Mid Level	Mid Level	Sesuai
Samsung Galaxy M12	Mid Level	Mid Level	Sesuai
Realme C31 3/32GB	Mid Level	Mid Level	Sesuai
Redmi 9A 3/32GB	Mid Level	Mid Level	Sesuai
Realme C11 (2021)	Mid Level	Mid Level	Sesuai

Dari hasil pengujian Tabel 4, maka dapat dibuat tabel dari confusion matrix untuk menghitung nilai akurasi dari klasifikasi yang telah dilakukan yang ditunjukkan pada Tabel 5.

Tabel 5. Tabel Hasil Confusion Matrix

True Class	Predicted Class			
		Flagship	High Level	Mid Level
	Flagship	8	4	0
	High Level	0	12	0
	Mid Level	0	0	12

Berdasarkan Tabel 5 confusion matrix diatas dapat dihitung tingkat akurasi dari K-NN yang sudah dibuat sebagai berikut :

$$Accuracy = TP / \text{Jumlah data} = \frac{8+12+12}{36} = 0,89 \times 100\% = 89\%.$$

Dengan demikian, akurasi dari algoritma *K-Nearest Neighbor* untuk klasifikasi *smartphone* dengan parameter $K = 3$ sebesar 89%

4. KESIMPULAN DAN SARAN

Berdasarkan hasil yang sudah didapatkan dari penelitian, dapat ditarik beberapa kesimpulan bahwa Algoritma *K-Nearest Neighbor* dapat berjalan dengan baik untuk program sederhana rekomendasi *smartphone*. Dan hasil akurasi dari 102 data *training* dan 36 data uji adalah 83% dengan menggunakan $K_{neighbor} = 3$ dan juga hasil dari klasifikasi sangat bertumpu pada data harga *smartphone*. Sehingga saran untuk penelitian selanjutnya untuk menambahkan data parameter agar tidak bergantung pada parameter harga saja dan data harus diperbaharui beberapa tahun karena perubahan jaman dapat merubah spesifikasi *smartphone* dikelas yang sama.

DAFTAR PUSTAKA

- [1] H. Yunida, "Pengaruh Ppkm Darurat Dan Protokol Kesehatan 5 M Terhadap Penurunan Covid-19," J. Innov. Knowl., vol. 1, no. 4, pp. 547–556, 2021.
- [2] Z. Hanum, "Kemenkominfo: 89% Penduduk Indonesia Gunakan Smartphone," Media Indonesia, 2021. .
- [3] A. Thomas, R. Candra, and H. Agung, "Perancangan Rekomendasi Lagu Menggunakan Algoritma K-Nearest Neighbor," J. Inf. Eng. Educ. Technol., vol. 4, no. 1, 2020, doi: 10.26740/jieet.v4n1.p41-46.
- [4] M. Rivki and A. M. Bachtiar, "IMPLEMENTASI ALGORITMA K-NEAREST NEIGHBOR DALAM PENGKLASIFIKASIAN FOLLOWER TWITTER YANG MENGGUNAKAN BAHASA INDONESIA," J. Sist. Inf., vol. 13, no. 1, 2017, doi: 10.21609/jsi.v13i1.500.
- [5] N. L. G. P. Suwirmayanti, "Penerapan Metode K-Nearest Neighbor Untuk Sistem Rekomendasi Pemilihan Mobil," Techno.Com, vol. 16, no. 2, 2017, doi: 10.33633/tc.v16i2.1322.
- [6] I. Handayani, "Application of K-Nearest Neighbor Algorithm on Classification of Disk Hernia and Spondylolisthesis in Vertebral Column," Indones. J. Inf. Syst., vol. 2, no. 1, 2019, doi: 10.24002/ijis.v2i1.2352.
- [7] I. Mulyadin and D. S. Winarso, "Sistem Pendukung Keputusan Pemilihan Smartphone Menggunakan Metode Simple Additive Weighting," CAHAYAtch, vol. 7, no. 2, p. 88, 2019, doi: 10.47047/ct.v7i2.13.
- [8] H. Hertyana et al., "Sistem Pendukung Keputusan Untuk Rekomendasi Pembelian Smartphone Dengan Menggunakan Metode Topsis," J. Tek. Inform. UNIKA St. Thomas, vol. 5, no. 1, 2020.
- [9] N. Shodik, N. Neneng, and I. Ahmad, "SISTEM REKOMENDASI PEMILIHAN SMARTPHONE SNAPDRAGON 636 MENGGUNAKAN METODE SIMPLE MULTI ATTRIBUTE RATING TECHNIQUE (SMART)," J. Nas. Pendidik. Tek. Inform., vol. 7, no. 3, 2019, doi: 10.23887/janapati.v7i3.15727.
- [10] C. T. Prasetyo, F. A. Hermawati, and E. Ronando, "SISTEM REKOMENDASI PEMILIHAN SMARTPHONE BERDASARKAN KARAKTERISTIK SOSIO-DEMOGRAFIS PENGGUNA MENGGUNAKAN METODE FUZZY TAHANI," KONVERGENSI, vol. 14, no. 1, 2019, doi: 10.30996/konv.v14i1.2769.

- [11] P. Pujiana, “sistem Pendukung Keputusan Pemilihan Smartphone dengan Pendekatan Fuzzy Multiple Attribute Decision Making,” *J. Sains Mat. dan Stat.*, vol. 7, no. 2, 2021, doi: 10.24014/jsms.v7i2.12921.
- [12] Aqmal Maulana, “Ini Bedanya HP Entry-level, Mainstream, High-end, dan Flagship,” *droidlime*, 2015. <https://www.droidlime.com/fitur/ini-bedanya-hp-entry-level-mainstream-high-end-dan-flagship/> (accessed May 20, 2022).
- [13] M. F. Naufal, “Analisis Perbandingan Algoritma SVM, KNN, dan CNN untuk Klasifikasi Citra Cuaca,” *J. Teknol. Inf. dan Ilmu Komput.*, vol. 8, no. 2, 2021, doi: 10.25126/jtiik.2021824553.
- [14] Y. Miftahuddin, S. Umaroh, and F. R. Karim, “PERBANDINGAN METODE PERHITUNGAN JARAK EUCLIDEAN, HAVERSINE, DAN MANHATTAN DALAM PENENTUAN POSISI KARYAWAN,” *J. Tekno Insentif*, vol. 14, no. 2, 2020, doi: 10.36787/jti.v14i2.270.
- [15] Diana Rahmawati, K. A. Wibisono, and F. D. W, “Analisis Sistem Parkir Cerdas Berbasis Ocr (Optical Character Recognition) Menggunakan Metode Confussion Matrix,” *Simp. Nas. Teknol. Terap.*, 2017.