



## ALGORITMA K-NEAREST NEIGHBOR PADA ANALISA SENTIMEN REVIEW PRODUK ROUTER

Rizqi Agung Permana<sup>1</sup>, Sucitra Sahara<sup>2\*</sup>

<sup>1</sup>rizqiagung@gmail.com, <sup>2</sup>sucitra.scr@bsi.ac.id

<sup>1</sup>Teknologi Komputer, STMIK Antar Bangsa, <sup>2</sup>Teknologi Informasi, Universitas Bina Sarana Informatika

### Abstrak

Review komentar yang telah dikumpulkan datanya dan akan diproses yaitu berupa *text positif* dan *text negatif* yang dilakukan pada pengklasifikasian data menggunakan k-Nearest Neighbors (k-NN) method. k-NN method merupakan salah satu algoritma yang cukup tepat untuk pengenalan pola data pada klasifikasi text. Pertumbuhan penggunaan router yang lekat berkaitan dengan kegiatan harian bagi pengguna laptop, *smartphone* dan tablet agar jaringan internet dapat dijalankan secara optimal, karena fungsi router yaitu dapat mengelola lalu lintas antar jaringan dengan meneruskan paket data ke alamat IP yang dituju. Maka dari itu peneliti mencoba memberikan kemudahan dalam pemilihan perangkat router yang tepat bagi pengguna mulai dari kualitas dan performa yang tinggi hingga kualitas yang sering dipertanyakan, sehingga peneliti mengadakan *screening* perangkat untuk opini atau komentar mengenai produk router oleh pengguna yang sudah menggunakan perangkat tersebut dan dituliskan ke dalam media online atau publik komentar. Dan dari pengolahan data sampel komentar yang telah dilakukan yaitu dengan mencoba beberapa metode algoritma seperti Naive Bayes, Neural Network, k-NN, Decision Tree dan Machine Support Vector Machine dengan berbagai tahap pengujian, barulah akan mendapatkan nilai akurasi dan AUC dari masing-masing algoritma tersebut sehingga didapatkan hasil pengujian tertinggi yaitu dengan menggunakan algoritma k-NN.

**Kata kunci:** k-NN, Sentimen Analisis, Review Produk

### Abstract

*Review comments that have collected the data and will be processed in the form of positive text and negative text which is carried out on classifying data using the k-Nearest Neighbors (k-NN) method, the k-NN method is an algorithm that is quite precise for recognizing data patterns in classification text. The growth in the use of routers that are closely connected to users who are connected to daily activities for laptops, smartphones and tablets so that the internet network can run optimally, because the router's function is to manage traffic between networks by forwarding data packets to the intended IP address. Therefore researchers try to provide convenience in selecting the right router device for users ranging from high quality and performance to quality that is often questioned, so researchers hold device screening for opinions or comments about router products by users who have used the device and written it to in online media or public comments. And from processing the comment sample data that has been done, namely by trying several algorithm methods such as Naive Bayes, Neural Network, k-NN, Decision Tree and Machine Support Vector Machine with various stages of testing, then you will get the accuracy and AUC values of each algorithm so that the highest test results are obtained by using the k-NN algorithm.*

**Keywords:** k-NN, Sentimen Analisis, Review Produk

### 1. Pendahuluan

Penelitian sebelumnya yang telah dilakukan pada pengklasifikasian sentimen pada review suatu produk berupa *text* yang diambil secara online layaknya analisa sentimen produk handphone dengan membandingkan metode SVM beserta metode NB, SVM adalah jenis *classifier* yang kemudian diberi satu *dataset* sebagai contoh pengujian, yang dimana setiap kategori sebagai kepemilikan satu dari dua kategorian yaitu algoritma SVM. Diambil dari jurnal mengenai analisis performa metode K-Nearest Neighbor untuk identifikasi jenis kaca, dengan menghasilkan perolehan nilai dari performa yang terbaik yaitu pada jumlah  $k=3$ , dengan akurasi sebesar 64%, dan besaran presisi mencapai 63%, kemudian recall angkanya sebesar 71% f-measure dengan hasil sebesar 67% [5]. Dengan menggunakan kategori

pengelompokan teks, baru dilakukan pembentukan dokumen yang dikelompokkan menurut kategori masing-masing, penyeleksian terhadap kategori spam pada email, memproses pendalaman opini mining dan sentimen analis. Algoritma yang digunakan pada kategori teks saat ini berkembang cepat, diantaranya yaitu metode SVM, pohon pengambilan keputusan. Algoritma k-NN merupakan model penentu jenis klasifikasi berdasar dengan sampel awal data yang dikategorikan tidak dapat membentuk representasi nilai dari sebuah deklaratif suatu kategori data teks, tapi tergantung dengan label dikategori yang menempel pada sebuah dokumen training yang hampir sama dengan dokumen pengujian. Kategori teks merupakan klasifikasi yang tepat pada pengelolaan dataset dimana perkembangan informasi pada saat ini sudah sangat cepat. Kategori teks membuat beberapa pengelolaan informasi atau data tersebut menjadi lebih efektif dan lebih efisien.

Algoritma PSO Teknik merupakan pencarian berbasis populasi berdasarkan gambarannya layaknya kelompok hewan terbang, contoh pada algoritma di komputasi evolusi yang lainnya model PSO dapat menciptakan solusi pada populasi yang baru didalam masalah iterasi nya. Algoritma PSO digunakan untuk memecahkan suatu kasus pada optimasi serta penyeleksian fitur [2]. Definisi data mining dapat diartikan adalah salah satu kategori ilmu yang mengkombinasikan teknik *machine learning*, pengenalan pola, statistik, database, visualisasi [8].

Salah satu mekanisme komputasi evolusi yang mampu menciptakan solusi dengan merata secara optimal dalam pencarian ruang dengan melewati interaksi dan berhubungan langsung secara individu disuatu kumpulan partikel [1]. Makna confusion matrix yakni merupakan algoritma atau model yang memakai pengukuran suatu kinerja pada pengklasifikasian [10].

Beberapa pengguna produk yang berminat membeli produk perangkat router, kadang gambling dalam mengambil produk yang terbaik dan berkualitas dari sekian jenis produk router yang tersedia di toko online, didukung begitu banyak produk perangkat yang kurang tepat dalam menghantarkan internet ke perangkat masing-masing seperti PC dan smartpone. Ada pula yang memalsukan produk perangkat ini yang dapat merugikan dari segi penjual maupun pembeli karena peminat perangkat router yang semakin menambah sehingga pemilik merk merasa keteter akan kegiatan pemalsuan produk [3].

Keuntungan yang dipunyai di algoritma K-NN antara lain dapat melakukan training data dengan cepat dan tepat, sederhana dan bisa dipelajari secara efektif jika data training tepat. Sedangkan kekurangan metode K-NN antara lain dinilai k masih bias, komputasi yang lengkap, dan mudah terkecoh dengan atribut-atribut yang belum relevan[9]. Pada penelitian, algoritma pengklasifikasian k-NN dengan optimasi PSO yang merupakan metode dalam pemilihan suatu fitur yang akan peneliti terapkan pada klasifikasi *text* sebuah opini publik review produk perangkat router yang posting pada situs penjualan produk online perangkat router dalam bentuk opini komentar pembeli dan pengguna, data sampel kata-kata bentuk komentar nantinya peneliti olah dan klasifikasikan berdasar pada objek yang ditetapkan, pengolahan dataset tersebut nantinya akan menjadikan 2 kategori yakni berupa komen positif dan komen negatif, hasil akhirnya akan dituangkan pada aplikasi web berupa review produk perangkat router yang akan diketahui produk tersebut termasuk kategori aman dan bagus atau tidak.

Analisa sentimen adalah metode atau model yang mengekstrak data berupa *text* untuk mendapatkan informasi sentimen yang bernilai jenis positif atau negatif [7]. Analisis ini menunjukkan pengaruh pada pengguna internet pada perangkat komputer maupun smartpone dan tablet sebagai kosumsi sehari hari, sehingga pembuatan aplikasi dengan program bahasa PHP untuk analisis review pada produk perangkat router yang peneliti rancang sangat perlu untuk prediksi dengan baik dan benar sebelum calon pelanggan membeli produk yang diinginkan, dan sebagai acuan pertimbangan pengguna dalam memilih produk tersebut.

Analisis pada sentimen bisa dikatakan jenis pengolahan kata dalam bentuk bahasa ataupun kalimat untuk mengetahui pendapat atau perasaan terhadap masyarakat yang keterkaitan dengan suatu produk. Analisis review sentimen yang juga disebut sumber pendapat opini online, keterlibatan kedalam pembangunan sistem untuk mengelompokkan dan mendalami pendapat opini mengenai produk yang di jual pada daftar barang di toko online, blog, ulasan produk, postingan tweet, dll.

## 2. Metode

Pada pengolahan review *text* yang akan diproses kedalam *text mining* yang pada dasarnya ada beberapa diantara karakteristik yang memiliki dimensi agak tinggi, dimana masih terdapat sebuah *noise* atau *miss* pada data, dan terdapat adanya struktur *text* yang tidak baik performanya. Data akan diolah

peneliti di beberapa studi kasus klasifikasi sentimen analisis yang di ambil dari sebuah situs e-commerce layaknya website [www.amazon.com](http://www.amazon.com) (review produk), [www.reviewcentre.com](http://www.reviewcentre.com), dll yang telah menjadi situs online ulasan opini produk oleh banyak pengguna, situs tersebut adalah sebuah situs review yang sudah profesional seperti [www.dpreview.com](http://www.dpreview.com), [www.zdnet.com](http://www.zdnet.com) dan situs pendapat opini pengguna produk mengenai *scope* yang luas.

Pengaturan data awal opini atau komentar sebuah review online ke dalam data lebih terstruktur yang diubah berupa bobot pada nilai numerik maka dari itu data yang dapat diolah sebagai sumber data awal yang dapat diproses. Proses pengubahan bentuk menjadi data yang lebih terstruktur yang sesuai dengan kebutuhan model yang dipakai untuk proses pengolahan data text mining, biasanya nanti akan diolah menjadi sebuah nilai numerik, proses tersebut populer disebut *text processing*.

Banyak proses yang akan dilakukan berupa *preprocessing* yang pertama *tokenization* yaitu peneliti akan melakukan penerapan menggunakan *tokenize* pada pemisahan suatu kata/huruf dari semua tanda pada bacaan dan data simbol. Yang kedua *stopward removal* yakni pemilahan kata yang dikategorikan tidak diperlukan dalam proses pengolahan data pada sentimen contohnya berupa kata sambung *if, or, the, of, etc.c.* yan ketiga *stemming* yang merupakan proses perubah bentuk sebuah kata biasa menjadi jenis kata dasar pada metode perubahan dari bentuk sebuah *text* dasar menjadi kata dasar ini harus menyesuaikan dari struktur suatu bahasa dalam tahap yaitu proses *stemming* [4].

Langkah penerapan algoritma k-NN:

1. Terlebih dahulu menentukan parameter k nya untuk jumlah tangga yang terdekat.
2. Perhitungan *Euclidean Distance* diantara data percobaan dan data yang akan diuji. Rumus dari *Euclidean Distance* bisa dilihat pada Persamaan  $d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$  (1) dimana ketentuan  $d(x,y)$  merupakan jarak dari *euclidean* diantara data coba pada titik x dan didata uji y,x data coba, y data uji, dan i variabel suatu data, serta ketentuan n dimensi dari atribut.
3. Menyusun jarak pada hasil perhitungan dilangkah ke 2, yakni mulai nilai terendah menuju ke nilai yang tertinggi.
4. Koleksi kategori y yang sesuai dengan nilai k.
5. Memilah jumlah pada kelas dari nilai tetangga yang mendekati jarak paling dekat, dan akan menentukan kelas yang akan menjadi kelas pada data dan nantinya diprediksi dahulu sesuaidengan hasil dikelas yang dominan [6].

Berikut ini persamaan yang ada pada model *Confution Matrix*:

- a. Nilai *accuracy* merupakan bentukan proporsi jumlah dari prediksi benar. Perhitungannya akurasi dengan persamaan berikut:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

- b. *Sensitivity* yang digunakan dalam proporsi TP pada tupel yang positif, dihitung dengan persamaan berikut:

$$Sensitivity = \frac{TP}{TP + FN}$$

- c. *Specificity* dipakai pada perbandingan suatu proporsi TN terhadap tupel berupa negatif, dengan perhitungan menggunakan persamaan sebagai berikut:

$$Specificity = \frac{TN}{TN + FP}$$

- d. PPV (*positive predictive value*) yakni bentuk proporsi pada kasus diagnosa yang bentuknya positif, perhitungan penggunaan persamaannya sebagai berikut:

$$PPV = \frac{TP}{TP + FP}$$

- e. NPV (*negative predictive value*) yakni proporsi pada kasus dengan nilai diagnosa yang negatif, dengan perhitungan menggunakan persamaan berikut:

$$NPV = \frac{TN}{TN + FN}$$

## 2.1. Proses Pengolahan Dataset

Berikut proses pada klasifikasi *text* analisa sentimen review opini:

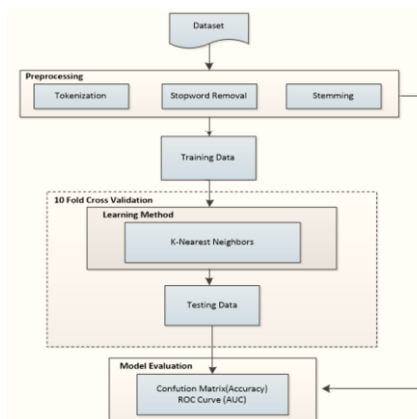
1. *Definisi dataset*  
Pengumpulan sebuah dataset yang telah melingkupi suatu situs berdomain, misal dataset review komentar film, dataset review komentar produk, dll.
2. *Pre-processing*  
Proses awal yang dilakukan berupa proses *tokenization*, *stopwords removal*, dan proses *stemming*.
3. *Transformation*  
Proses representasi pada angka akan dihitung data yang berupa *text*. Representasi *binary* yang diolah dan akan menghitung keberadaan atau tidak ada pada sebuah kata *text* di dalam dokumen-dokumen opininya. Dan dilihat berapa kali kata yang muncul didalam dokumen opini tersebut akan digunakan pada skema pembobotan nilai dari data yang *text*. Proses yang telah ditentukan yakni metode TF-IDF, *Frequency transformation*, *Binary transformation*.
4. *Feature Selection*  
Fitur ini menghasilkan klasifikasi *text* yang efisien efektif dengan menghilangkan jumlah data untuk bisa dianalisa saat identifikasi fitur real yang selanjutnya di olah kembali. Metode atau model penentuan fitur yang sering digunakan *information gain*, *knowledge*, *minimum frequency*, dll.
5. *Classification*  
Klasifikasi yang sering dipakai yaitu pengklasifikasi seperti Algoritma *Naive Bayes*, *SVM*, dll.

## 2.2. Tahapan Pengolahan Penelitian

Tahap pengolahan penelitian kali ini mengedepankan penerapan model atau algoritma pada kegiatan penelitian review opini pada produk perangkat router yaitu model k-NN yang diproses beberapa tahap. Algoritma k-NN nantinya menghasilkan metode dimana peneliti terlebih dahulu menentukan besaran nilai k untuk mengetahui tingkatan dan nilai akurasi yang keluar menjadi tinggi pada saat tahap pengujian dataset. Besaran nilai pada akurasi yang tertinggi nantinya digunakan peneliti untuk menentukan nilai bobot tersebut optimal atau tidaknya, dan setelah diolah dan diproses ternyata nilai bobot tersebut sudah baik artinya.

Disimpulkan ternyata bahwa struktur pada algoritma telah dirancang untuk mencapai ideal akurasi bobot dalam suatu pemecahan masalah salah satunya sentimen analisis. Data training atau sampel yang digunakan pada klasifikasi *text* yakni 100 review opini positif dan 100 review opini negatif pada review produk perangkat router. Data review komentar yang akan diolah masih berupa kelompok *text* yang terpisah dalam bentuk dokumen *text*. Sebelum data sampai tahap klasifikasi, dataset harus melewati proses dan tahapan supaya data tersebut dapat diolah dan diproses dengan performa yang baik.

Berikut gambar usulan peneliti yang telah dirancang dalam review produk perangkat router:



Gambar 1. Metode Usulan

Pada evaluasi kali ini penulis mengusulkan penggunaan model dalam kegiatan review perangkat router yaitu algoritma k-NN, yang telah dilakukan ada dua tahapan penerapan. Algoritma k-NN akan menghasilkan model dan nantinya peneliti akan mendapatkan nilai k yang tertinggi untuk pencarian tingkat akurasi yang paling tinggi pada pengujian di dataset sampel. Nilai bobot pada akurasi yang paling tinggi akan diambil peneliti dalam menentukan apakah nilai tersebut optimal atau tidak, dan hasilnya ternyata nilai tersebut menghasilkan akurasi yang sudah baik. Disimpulkan bahwa struktur algoritma yang dirancang mencapai ideal dalam pemecahan masalah penelitian.

### 3. Hasil dan Pembahasan

#### 3.1. Hasil Klasifikasi Text Algoritma k-NN

Data *training* digunakan dalam pengolahan pengklasifikasian *text* diambil sampel 100 review positif dan 100 review negatif pada perangkat router. Data review yang diolah masih berupa sekumpulan dokumen txt yang masih terpisah. Sebelum dikelompokkan, data tersebut harus melewati tahapan-tahapan agar data bisa diolah dengan baik dan tepat. Pada penelitian ini, pengolahan data akan mendapatkan hasil dari pengujian dari algoritma dan optimasi yang akan didapat berupa akurasi dan AUCnya.

Data *training* yang diolah dalam pengklasifikasian *text* pada review produk perangkat router, akan diproses lagi yang dimana data awalnya masih berupa sekumpulan *text* terpisah dalam bentuk dokumen *text*. Sebelum diklasifikasikan, data harus melewati proses tahapan lagi supaya datanya dapat diproses dengan baik dan tepat. Berikut ini pengelompokan dokumen berdasarkan pengklasifikasiannya:

Tabel 1. Vector Boolean dengan Kelas Label Hasil Klasifikasi

Dokumen	Great	Working	Bad	Nomax	Class
Positif4.txt	1	0	0	0	Poss
Positif2.txt	1	1	0	0	Poss
Positif83.txt	0	1	0	0	Poss
Positif10.txt	1	1	0	0	Poss
Positif36.txt	1	0	0	0	Poss
Negatif95.txt	0	0	1	1	Negg
Negatif3.txt	0	0	1	0	Negg
Negatif44.txt	0	0	1	0	Negg
Negatif12.txt	0	0	0	1	Negg
Negatif52.txt	0	0	1	0	Negg

#### 3.2. Penerapan Algoritma k-NN

Eksperimen pada pengujian dataset terhadap algoritma k-NN, nilai hasil *query* pada instansi dipenelitian ini ditentukan dengan melakukan pengujian dan percobaan dengan memasukkan nilai k (jumlah nilai tetangga terdekat).

Tabel 2. Eksperimen Metode k-NN

Nilai K	Accuracy	AUC
1	65.50%	0.542
2	65.71%	0.547
3	70.00%	0.541
4	72.42%	0.643
5	74.53%	0.646
6	74.79%	0.738
7	74.81%	0.746
8	75.57%	0.749
9	78.31%	0.818
10	83.20%	0.879

Tabel 2 menunjukkan penerapan algoritma k-NN dengan nilai k=10 yang nilai bobot akurasinya sampai 83.20% dan nilai AUC 0.879 mendapatkan hasil yang paling tinggi diantara penunjukan nilai k yang lainnya.

### 3.3. Hasil Analisis Evaluasi dan Hasil

Hasil pengujian yang dilakukan peneliti diawal pembahasan, pengukuran nilai akurasi penggunaan *confusions matrix* dan kurva model ROC menunjukkan bahwa hasil pengujian algoritma k-NN bernilai cukup menunjukkan hasil performa baik, nilai *accuracy* terhadap model algoritma k-NN 83.20% pada tetangga k=10, yang terdapat di Tabel 3.

Tabel 3. Pengujian k-NN

K-NN Algoritm	Nilai K	Accuracy	AUC
	10	83.20%	0.879

Tabel 4. Confusion Matrix pada Model k-NN

Accuracy : 83.20% +/- 8.50% (mikro: 83.20%)			
	True Positif	True Negatif	Class Precision
Prediksi Positif	85	35	70.25%
Prediksi Negatif	25	63	81.01%
Class Recall	85.00%	63.00%	

Berdasarkan tabel diatas diartikan menunjukkan nilai bobot akurasi yang baik (83.20%), dan dilihat dari hasil yang terbentuk pada hasil tabel pengujian bisa dicari nilai akurasinya, jadi pengujian ini merupakan algoritma k-NN yang dapat menaikkan nilai akurasi yang bisa jadi solusi yang baik dalam studi kasus pada klasifikasi sentimen review opini produk perangkat router.

### 3.4. Pembahasan

Dibuktikan dari penerapan algoritma k-NN pada data olah akan mendapat nilai *accuracy* pada klasifikasi *text review* produk perangkat router dalam pengidentifikasian review berupa komentar positif dan review komentar negatif, dengan model klasifikasi text pada analisa review komentar, pengguna nantinya akan mudah dalam mengidentifikasi mana review positif dan mana yang negatif. Dari data contoh review yang akan diolah, akan dipisahkan dahulu menjadi kata atau text, lalu diberikan nilai untuk masing-masing kata tersebut, barulah dapat dilihat kata atau text yang berhubungan dengan sentimen analisis yang sering ditunjukkan dan mempunyai nilai bobot tinggi. Dengan demikian diketahui kalau review tersebut termasuk review opini yang tergolong positif ataupun review opini tergolong negatif.

Dalam penelitian ini hasil pengujian model akan ditunjukkan melalui *confusion matrix* agar menunjukkan mana model klasifikasi yang paling tepat dan benar dengan tanpa lagi menggunakan metode pemilihan fitur yang lainnya, bahwa algoritma k-NN sudah menghasilkan akurasi yang cukup besar yaitu 83.20% dan nilai bobot AUC 0.879.

Peneliti nantinya akan menyediakan aplikasi berbasis web dengan bahasa pemrograman PHP untuk pengujian model dengan menggunakan dataset yang beda dan belum diklasifikasi sesuai pada kelasnya, yang memudahkan calon pengguna untuk memilah pemisahan antara komentar positif dan negatif untuk penentuan produk router yang baik. Implikasi pada aspek sebuah sistem dari hasil evaluasi

dimana menunjukkan penggunaan Algoritma k-NN yang merupakan algoritma paling tepat dalam mengklasifikasikan text pada review produk perangkat router.

Diharapkan dari penelitian ini dapat mempermudah calon pengguna produk perangkat router dalam menentukan produk yang layak untuk mereka miliki, supaya tidak terjadi kesalahan dalam pemilihan produk yang tersedia diberberapa toko online. Penerapan model algoritma terhadap aspek ke penelitian selanjutnya dapat menggunakan algoritma dalam penentuan fitur atau pun dataset sampel dari produk lainnya, sebagai contoh review komentar tempat wisata, review maskapai, atau banyak lainnya yang sesuai dengan kebutuhan masyarakat saat itu.

#### 4. Kesimpulan dan Saran

Kesimpulan dari penelitian yang menjelaskan capaian yang didapat pada hasil menggunakan metode yang diusulkan. Adapun saran menjelaskan tentang usulan-usulan bagi penelitian selanjutnya untuk lebih menyempurnakan hasil penelitian yang telah didapatkan. Ringkasan pada klasifikasi text kali ini dengan dataset awal berupa review opini produk perangkat router, yang sesuai yaitu salah satunya klasifikasinya yang digunakan pada Algoritma k-NN. Hal ini dikarenakan algoritma k-NN yaitu metode tergolong sesuai pada pengolahan data yang berupa text dan sangatlah mudah dipahami. Data review yang sudah diambil sampel, akan dikelompokkan menjadi kata dan akan diberikan nilai dimasing-masing kata, dan kita lihat text mana yang berhubungan dengan sentimen opini review komentar yang sering muncul yang mempunyai bobot nilai tinggi.

Demikian dapat diketahui review produk tersebut tergolong review positif atau review opini negatif. Algoritma k-NN juga populer digunakan pada beberapa penelitian pada klasifikasi jenis text yang mempunyai performa cukup baik yang berupa pengolahan data yang telah diproses. Text review komentar yang peneliti proses sudah dapat diklasifikasikan dengan baik kedalam bentuk text jenis positif dan text jenis negatif. Hasilnya pada data accuracy dengan penggunaan algoritma k-NN pada data review produk perangkat router mencapai 83.20% dan nilai AUC sebesar 0.879.

#### Daftar Pustaka

- [1] P. J. García Nieto, E. García-Gonzalo, F. Sánchez Lasheras, and F. J. de Cos Juez, "Hybrid PSO–SVM-based method for forecasting of the remaining useful life for aircraft engines and evaluation of its reliability," *Reliab. Eng. Syst. Saf.*, vol. 138, pp. 219–231, 2015.
- [2] F. Zhao, Y. Liu, C. Zhang, and J. Wang, "A self-adaptive harmony PSO search algorithm and its performance analysis," *Expert Syst. Appl.*, vol. 42, no. 21, pp. 7436–7455, 2015.
- [3] P. Kasih, "Perlindungan hukum bagi masyarakat terhadap peredaran barang palsu," *J. Univ. Atma Jaya Yogyakarta*, vol. 5, no. 6, pp. 12–21, 2016.
- [4] M. Sholehudin, M. Fauzi Ali, and S. Adinugroho, "Implementasi Metode Text Mining dan K-Means Clustering untuk Pengelompokan Dokumen Skripsi (Studi Kasus: Universitas Brawijaya)," vol. 2, no. 11, pp. 5518–5524, 2018.
- [5] M. M. Baharuddin, H. Azis, and T. Hasanuddin, "Analisis Performa Metode K-Nearest Neighbor Untuk Identifikasi Jenis Kaca," *Ilk. J. Ilm.*, vol. 11, no. 3, pp. 269–274, Dec. 2019
- [6] K. Malik, Y. Pratama, and K. Nisa, "Diagnosa Pharyngitis Menggunakan Metode K-Nearest Neighbor (KNN) di Puskesmas Leces Probolinggo," *TRILOGI J. Ilmu Teknol. Kesehat. dan Hum.*, vol. 2, no. 3, pp. 199–206, Dec. 2021
- [7] Karsito and S. Susanti, "Klasifikasi Kelayakan Peserta Pengajuan Kredit Rumah dengan Algoritma Naïve Bayes Di Perumahan Azzura Residencia," *J. Teknol. Pelita Bangsa*, vol. 9, pp. 43–48, 2019
- [8] D. R. W. Z. M. Jollyta, *Konsep Data Mining dan Penerapan*. Yogyakarta: Deepublish, 2020
- [9] G. A. Rosso, "Milton," *William Blake Context*, no. September, pp. 184–191, 2019, doi: 10.1017/9781316534946.021
- [10] F. V. Sari and A. Wibowo, "Analisis Sentimen Pelanggan Toko Online Jd.Id Menggunakan Metode Naïve Bayes Classifier Berbasis Konversi Ikon Emosi," *J. SIMETRIS*, vol. 10, no. 2, pp. 681–686, Nov. 2019