
IMPLEMENTASI *TEXT MINING* DALAM ANALISA SENTIMEN MASYARAKAT TERHADAP KETANGGAPAN PEMERINTAH DALAM PENANGGULANGAN BENCANA BANJIR

Apri Affandi¹, Muhammad Sabir Ramadhan²

Universitas Asahan, Asahan

e-mail: ¹apriaffandi777@gmail.com, ²ramasabir@gmail.com.

Abstract: *The development of information technology and social media has made it easier for people to express their opinions on various issues, including the government's responsiveness in handling flood disasters. The large number of comments generated on social media makes manual opinion analysis ineffective because it requires a long time and involves a large volume of data. Therefore, an automated method is needed to process and classify public sentiment. This study aims to implement text mining to analyze public sentiment toward the government's responsiveness in flood disaster management using the K-Nearest Neighbour (KNN) algorithm. The research data were collected from social media platforms such as Instagram, TikTok, and X, consisting of public comments related to flood disasters in Aceh and North Sumatra. The research stages include data collection, text preprocessing, feature extraction, data splitting into training and testing sets, sentiment classification using the KNN method, and evaluation of classification results. The system was developed as a web-based application using PHP programming language and MySQL database. The results of this study are expected to provide an overview of public opinion tendencies in the form of positive, negative, and neutral sentiments and serve as evaluation material for the government to improve its response in flood disaster management.*

Keywords: *Text Mining, Sentiment Analysis, K-Nearest Neighbour (KNN), Social Media, Flood Disaster, Government Responsiveness.*

Abstrak: Perkembangan teknologi informasi dan media sosial menyebabkan masyarakat semakin mudah menyampaikan pendapat terhadap berbagai isu, termasuk ketanggapan pemerintah dalam penanggulangan bencana banjir. Banyaknya komentar yang muncul di media sosial menjadikan analisis opini secara manual tidak efektif karena membutuhkan waktu yang lama dan jumlah data yang besar. Oleh karena itu, diperlukan metode otomatis untuk mengolah dan mengklasifikasikan sentimen masyarakat. Penelitian ini bertujuan untuk mengimplementasikan text mining dalam menganalisis sentimen masyarakat terhadap ketanggapan pemerintah dalam penanggulangan bencana banjir menggunakan algoritma K-Nearest Neighbour (KNN). Data penelitian diperoleh dari platform media sosial Instagram, TikTok, dan X yang berisi komentar masyarakat terkait bencana banjir di wilayah Aceh dan Sumatera Utara. Tahapan penelitian meliputi pengumpulan data, preprocessing teks, ekstraksi fitur, pembagian data training dan testing, proses klasifikasi menggunakan metode KNN, serta evaluasi hasil klasifikasi. Sistem dibangun berbasis web menggunakan bahasa pemrograman PHP dan database MySQL. Hasil penelitian diharapkan dapat memberikan gambaran kecenderungan opini masyarakat dalam bentuk sentimen positif, negatif, dan netral serta membantu sebagai bahan evaluasi bagi pemerintah dalam meningkatkan kualitas respon terhadap bencana banjir.

Kata kunci: *Text Mining, Analisis Sentimen, K-Nearest Neighbour (KNN), Media Sosial, Bencana Banjir, Ketanggapan Pemerintah.*

PENDAHULUAN

Bencana banjir merupakan salah satu bencana alam yang paling sering terjadi di Indonesia dan menimbulkan dampak besar terhadap kehidupan masyarakat, baik dari aspek sosial, ekonomi, maupun lingkungan. Tingginya curah hujan, buruknya sistem drainase, serta rendahnya kesadaran lingkungan menjadi faktor utama terjadinya banjir di berbagai wilayah (Rahmadini Puja et al, 2026). Perkembangan teknologi informasi saat ini mendorong masyarakat untuk menyampaikan pendapat, kritik, serta penilaian terhadap kinerja pemerintah melalui media sosial. Media sosial berkembang menjadi sarana efektif untuk mengekspresikan opini secara terbuka, termasuk terkait respons pemerintah dalam penanggulangan bencana banjir. Opini masyarakat tersebut membentuk sentimen publik yang mencerminkan kondisi emosional masyarakat berupa komentar positif, negatif, maupun netral terhadap kebijakan dan tindakan pemerintah.

Namun tingginya jumlah data teks yang dihasilkan dari media sosial menyebabkan analisis opini secara manual menjadi tidak efektif dan membutuhkan waktu yang lama. Oleh karena itu diperlukan pendekatan berbasis teknologi informasi untuk mengolah dan menganalisis data teks dalam jumlah besar. Salah satu pendekatan yang dapat digunakan adalah *text mining*, yaitu proses mengekstraksi informasi penting dari data teks yang tidak terstruktur. *Text mining* bertujuan mengubah teks dalam jumlah besar menjadi informasi yang bernilai dengan teknik komputasi untuk menemukan pola, tren, serta informasi tersembunyi. Dengan *text mining*, dapat diketahui apa yang masyarakat katakan, pikirkan, dan butuhkan terhadap suatu topik, sehingga membantu pengambilan keputusan yang lebih tepat dan meningkatkan kualitas layanan.

Text mining merupakan metode penting dalam proses pencarian

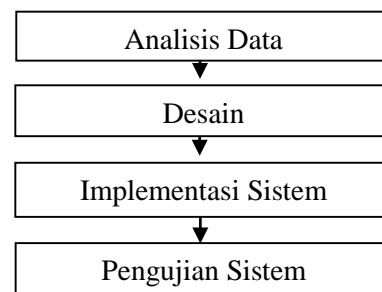
pengetahuan dari data berbentuk teks. *Text mining* digunakan ketika jumlah teks yang dimiliki sangat banyak, sehingga jika dibaca dan dianalisis satu per satu oleh manusia akan memakan waktu lama dan tidak efisien. Pada dasarnya, *text mining* membantu komputer untuk membaca dan memahami tulisan seperti yang dilakukan manusia, meskipun tentu dengan cara yang berbeda (Algoritma et al., 2026).

Analisis sentimen merupakan bagian dari *text mining* yang bertujuan mengidentifikasi opini masyarakat ke dalam kategori positif, negatif, atau netral.

METODE

Algoritma

Penelitian ini menggunakan algoritma *K-Nearest Neighbour* (KNN), yaitu metode *supervised learning* yang mengklasifikasikan data berdasarkan kedekatan jarak dengan data latih terdekat. Struktur kerja sistem *clustering* yang dibangun dapat dilihat pada gambar berikut:



Gambar 1 Kerangka Kerja

Tahapan Analisis

Penelitian diawali dengan pengumpulan data dengan metode (kualitatif), dilanjutkan dengan tahap *preprocessing*. Data yang tervalidasi kemudian diproses menggunakan komputasi algoritma *K-NEAREST NEIGHBORS* (K-NN).

Implementasi KNN

Data dikelompokkan ke dalam 3 kluster, yaitu Positif, Netral, dan Negatif. Melalui algoritma KNN, setiap data

terhadap pusat kluster dihitung secara iteratif. Keputusan akhir kluster ditentukan oleh nilai keanggotaan terdekat, dan komputasi akan berhenti (konvergen) ketika selisih fungsi objektif telah mencapai batas toleransi *error* yang ditentukan. Rumus dari algoritma KNN yaitu:

$$d(x,y)=\sum_{i=1}^n(x_i-y_i)^2$$

Keterangan:

d = jarak

x_i = nilai atribut data uji

y_i = nilai atribut data latih

n = jumlah atribut

Alat dan Evaluasi

Komputasi algoritma KNN diimplementasikan ke dalam sistem berbasis web PHP dan MySQL guna menjamin efisiensi dan presisi perhitungan. Hasil clustering yang diperoleh dapat digunakan oleh Pemerintah sebagai bahan analisis dalam mengelompokkan tingkat akurasi

sentimen masyarakat berdasarkan variabel respon, penanganan, transparansi, bantuan dan infrastruktur sehingga mendukung pengambilan keputusan dalam menganalisis sentimen atau komentar masyarakat terhadap ketanggapan pemerintah dalam penanggulangan bencana banjir.

HASIL DAN PEMBAHASAN

Analisa Data

Menentukan Data Kriteria

Tahap awal komputasi adalah menetapkan kriteria berdasarkan data yang di peroleh dari media sosial. Data uji merupakan faktor pendukung untuk melakukan proses pengujian, data uji yang digunakan dalam penelitian ini sebanyak 5 (Lima) variabel utama sebagai tolak ukur penilaian, yang dapat dilihat pada Tabel 1 berikut:

Tabel 1 Data Kriteria

No	Kode Variabel	Nama Variabel
1	00001	Respon (Kecepatan Respon Pemerintah)
2	00002	Penanganan (Kualitas Penanganan)
3	00003	Transparansi (Transparansi Informasi)
4	00004	Bantuan (Ketersediaan Bantuan)
5	00005	Infrastruktur (Infrastruktur dan Penanggulangan)

Penentuan Jumlah Kluster

Sebelum proses komputasi iteratif *K-NEAREST NEIGHBOR* (KNN) dilakukan, terlebih dahulu ditetapkan target pengelompokan sebanyak 3 (tiga) kluster. Penentuan jumlah kluster ini bertujuan untuk mengelompokkan tingkat akurasi berdasarkan karakteristik data yang digunakan. Adapun pembagian kluster tersebut terdiri dari kategori produksi positif, netral, dan negatif, yang dapat dilihat pada Tabel 2 berikut:

Tabel 2 Kategori Kluster Sentimen

Masyarakat

No	Kelas	Keterangan
1	Klaster 1	Positif
2	Klaster 2	Netral
3	Klaster 3	Negatif

Dataset

Dataset diperoleh dari sosial media yaitu sentimen masyarakat terhadap pemerintah dalam penanggulangan bencana banjir. Dalam kasus ini perolehan dataset dapat dilihat pada Tabel 3 berikut ini:

Tabel 3 Dataset

No	Kode	Respon	Penanganan	Transparansi	Bantuan	Infrastruktur
----	------	--------	------------	--------------	---------	---------------

1	00010	0	0	0	0	0
2	00012	0	1	0	0	0
3	00014	0	0	1	1	0
4	00018	0	0	0	1	1
5	00019	0	0	0	1	1
6	00027	0	0	0	0	0
7	00036	0	0	0	0	0
8	00037	0	0	0	1	1
9	00043	0	1	0	0	1
10	00044	0	0	1	1	0
....						
60	00060	1	0	1	1	1

Nilai Cluster

Keputusan akhir clustering didasarkan pada pembagian data ke dalam 5 kluster utama. Ketiga kluster ini

merepresentasikan tingkat akurasi algoritma KNN dalam menganalisa sentimen masyarakat, yang meliputi:

Tabel 4 Clustering Nilai

NO	Nama	Respon	Penanganan	Transparansi	Bantuan	Infrastruktur
1	Samir	1	1	1	0	0
2	Suri	0	0	1	1	1
3	Irfan	1	1	1	1	0
4	Radi	0	0	0	1	1
5	Sumin	1	1	1	0	1

Pengujian

Setelah *dataset* disiapkan, langkah selanjutnya adalah menetapkan parameter dasar komputasi untuk algoritma *K-NEAREST NEIGHBOR* (KNN). Parameter ini berfungsi sebagai batasan agar perulangan (iterasi) dapat berjalan dan berhenti pada titik konvergen.

ditetapkan pada pengujian ini adalah:

- Jumlah Cluster : 3
- Pembobotan : 3

Adapun parameter yang

PERHITUNGAN UNTUK DATA UJI: SAMIR.

Perhitungan Jarak ke Data Training Terdekat (Kode 00002)

$$d = \sqrt{(1 - 1)^2 + (0 - 0)^2 + (1 - 1)^2 + (1 - 1)^2 + (1 - 1)^2}$$

$$d = \sqrt{0^2 + 0^2 + 0^2 + 0^2 + 0^2} = \sqrt{0} = 0$$

B. Perhitungan Jarak ke Data Training Terdekat Kedua (Kode 00060)

$$d = \sqrt{(1 - 1)^2 + (0 - 0)^2 + (1 - 1)^2 + (1 - 1)^2 + (1 - 1)^2}$$

$$d = \sqrt{0^2 + 0^2 + 0^2 + 0^2 + 0^2} = \sqrt{0} = 0$$

C. Perhitungan Jarak ke Data Training Terdekat Ketiga (Kode 00046)

$$d = \sqrt{(1 - 0)^2 + (0 - 0)^2 + (1 - 1)^2 + (1 - 1)^2 + (1 - 1)^2}$$

$$d = \sqrt{1^2 + 0^2 + 0^2 + 0^2 + 0^2} = \sqrt{1} = 1$$

Keputusan Akhir Samir (K=3):

Jarak 0 = Positif (Kode 00002)

Jarak 0 = Positif (Kode 00060)

Jarak 1 = Negatif (Kode 00046)

Hasil Voting: Positif (2) > Negatif (1) = POSITIF

PERHITUNGAN UNTUK DATA UJI: SURI

Perhitungan Jarak ke Data Training Terdekat (Kode 00014)

$$d = \sqrt{(0-0)^2 + (0-0)^2 + (1-1)^2 + (1-1)^2 + (0-0)^2}$$

$$d = \sqrt{0^2 + 0^2 + 0^2 + 0^2 + 0^2} = \sqrt{0} = 0$$

B. Perhitungan Jarak ke Data Training Terdekat Kedua (Kode 00044)

$$d = \sqrt{(0-0)^2 + (0-0)^2 + (1-1)^2 + (1-1)^2 + (0-0)^2}$$

$$d = \sqrt{0^2 + 0^2 + 0^2 + 0^2 + 0^2} = \sqrt{0} = 0$$

C. Perhitungan Jarak ke Data Training Terdekat Ketiga (Kode 00046)

$$d = \sqrt{(0-0)^2 + (0-0)^2 + (1-1)^2 + (1-1)^2 + (0-1)^2}$$

$$d = \sqrt{0^2 + 0^2 + 0^2 + 0^2 + (-1)^2} = \sqrt{1} = 1$$

Keputusan Akhir Suri (K=3):

Jarak 0 → Negatif (Kode 00014)

Jarak 0 → Negatif (Kode 00044)

Jarak 1 → Negatif (Kode 00046)

Hasil Voting: Negatif (3) > Positif (0) = NEGATIF

PERHITUNGAN UNTUK DATA UJI: IRFAN

Perhitungan Jarak ke Data Training Terdekat (Kode 00001)

$$d = \sqrt{(1-1)^2 + (1-1)^2 + (1-1)^2 + (1-1)^2 + (1-1)^2}$$

$$d = \sqrt{0^2 + 0^2 + 0^2 + 0^2 + (0)^2} = \sqrt{0} = 0$$

B. Perhitungan Jarak ke Data Training Terdekat Kedua (Kode 00003)

$$d = \sqrt{(1-1)^2 + (1-1)^2 + (1-1)^2 + (1-1)^2 + (1-1)^2}$$

$$d = \sqrt{0^2 + 0^2 + 0^2 + 0^2 + (0)^2} = \sqrt{0} = 0$$

C. Perhitungan Jarak ke Data Training Terdekat Ketiga (Kode 00005)

$$d = \sqrt{(1-1)^2 + (1-1)^2 + (1-1)^2 + (1-1)^2 + (1-1)^2}$$

$$d = \sqrt{0^2 + 0^2 + 0^2 + 0^2 + (0)^2} = \sqrt{0} = 0$$

Keputusan Akhir Irfan (K=3):

Jarak 0 → Positif (Kode 00001)

Jarak 0 → Positif (Kode 00003)

Jarak 0 → Positif (Kode 00005)

Hasil Voting: Positif (3) > Negatif (0) = POSITIF

PERHITUNGAN UNTUK DATA UJI: RADI

Perhitungan Jarak ke Data Training Terdekat (Kode 00018)

$$d = \sqrt{(0-0)^2 + (0-0)^2 + (0-0)^2 + (1-1)^2 + (1-1)^2}$$

$$d = \sqrt{0^2 + 0^2 + 0^2 + 0^2 + (0)^2} = \sqrt{0} = 0$$

B. Perhitungan Jarak ke Data Training Terdekat Kedua (Kode 00019)

$$d = \sqrt{(0-0)^2 + (0-0)^2 + (0-0)^2 + (1-1)^2 + (1-1)^2}$$

$$d = \sqrt{0^2 + 0^2 + 0^2 + 0^2 + (0)^2} = \sqrt{0} = 0$$

C. Perhitungan Jarak ke Data Training Terdekat Ketiga (Kode 00037)

$$d = \sqrt{(0-0)^2 + (0-0)^2 + (0-0)^2 + (1-1)^2 + (1-1)^2}$$

$$d = \sqrt{0^2 + 0^2 + 0^2 + 0^2 + (0)^2} = \sqrt{0} = 0$$

Keputusan Akhir Radi (K=3):

Jarak 0 → Negatif (Kode 00018)

Jarak 0 → Negatif (Kode 00019)

Jarak 0 → Negatif (Kode 00037)

Hasil Voting: Negatif (3) > Positif (0) = NEGATIF

PERHITUNGAN UNTUK DATA UJI: SUMIN

Perhitungan Jarak ke Data Training Terdekat (Kode 00013)

$$d = \sqrt{(1 - 1)^2 + (1 - 1)^2 + (1 - 1)^2 + (0 - 0)^2 + (0 - 0)^2}$$

$$d = \sqrt{0^2 + 0^2 + 0^2 + 0^2 + (0)^2} = \sqrt{0} = 0$$

B. Perhitungan Jarak ke Data Training Terdekat Kedua (Kode 00025)

$$d = \sqrt{(1 - 1)^2 + (1 - 1)^2 + (1 - 1)^2 + (0 - 0)^2 + (0 - 0)^2}$$

$$d = \sqrt{0^2 + 0^2 + 0^2 + 0^2 + (0)^2} = \sqrt{0} = 0$$

C. Perhitungan Jarak ke Data Training Terdekat Ketiga (Kode 00031)

$$d = \sqrt{(1 - 1)^2 + (1 - 1)^2 + (1 - 1)^2 + (0 - 0)^2 + (0 - 0)^2}$$

$$d = \sqrt{0^2 + 0^2 + 0^2 + 0^2 + (0)^2} = \sqrt{0} = 0$$

Keputusan Akhir Sumin (K=3):

Jarak 0 → Positif (Kode 00013)

Jarak 0 → Positif (Kode 00025)

Jarak 0 → Positif (Kode 00031)

Hasil Voting: Positif (3) > Negatif (0) = POSITIF

Tabel 5 REKAPITULASI HASIL UJI

Nama	Vektor Atribut [x1,x2,x3,x4,x5]	3 Tetangga Terdekat (Kode)	Pola Sentimen Terpilih	Keputusan Klasifikasi Akhir
Samir	[1, 0, 1, 1, 1]	00002, 00060, 00046	2 positif, 1 negatif	POSITIF
Suri	[0, 0, 1, 1, 0]	00014, 00044, 00046	0 positif, 3 negatif	NEGATIF
Irfan	[1, 1, 1, 1, 1]	00001, 00003, 00005	3 positif, 0 negatif	POSITIF
Radi	[0, 0, 0, 1, 1]	00018, 00019, 00037	0 positif, 3 negatif	NEGATIF
Sumin	[1, 1, 1, 0, 0]	00013, 00025, 00031	3 positif, 0 negatif	POSITIF

Pembahasan

1. Menu login merupakan menu yang akan muncul saat admin menjalankan aplikasi, selanjutnya input username dan password, berikut tampilan menu login saat menu login dipilih.



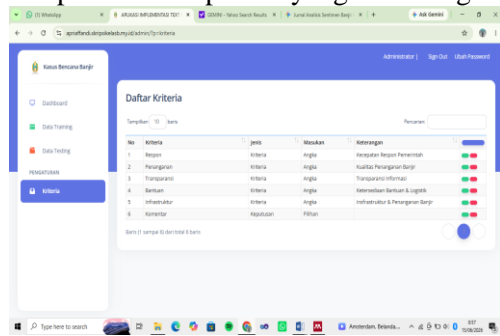
Gambar 1 Halaman Login

2. Menu utama berisikan menu kriteria, alternatif, perhitungan dan password, menu-menu tersebut berfungsi untuk melakukan pengolahan data yang ada di sistem. Berikut tampilan dari menu utama dari aplikasi yang dirancang.



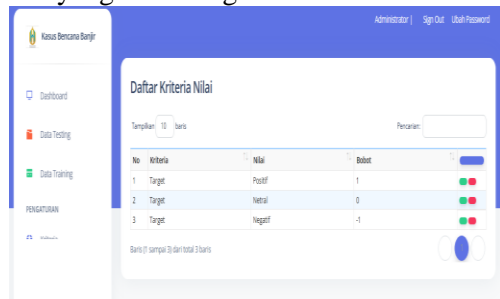
Gambar 2 Halaman Menu Utama

3. Menu kriteria digunakan untuk menambahkan data kriteria atau variable. Berikut tampilan dari menu priode dari aplikasi yang dirancang.



Gambar 3 Halaman Menu Kriteria

4. Menu kriteria nilai digunakan untuk menambahkan, edit dan hapus cluster. Berikut tampilan dari menu alternatif yang dirancang.



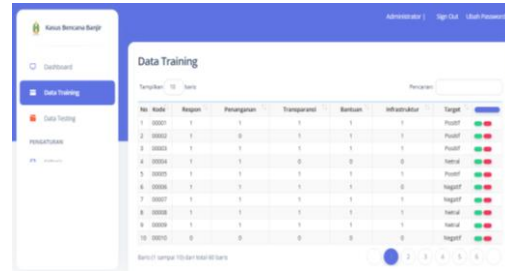
Gambar 4 Halaman Menu Kriteria Nilai

5. Menu data testing digunakan untuk membuat dan memasukkan data baru, terdapat menu tambah, edit dan hapus. Berikut tampilan dari menu perhitungan dari aplikasi yang dirancang.



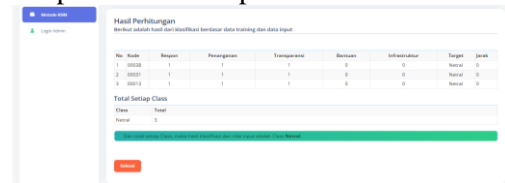
Gambar 5 Halaman Menu Data Testing

6. Menu data training digunakan sebagai data mutlak yang sudah di input, dan bisa di edit, dan juga hapus. Berikut tampilan dari menu password dari aplikasi yang dirancang.



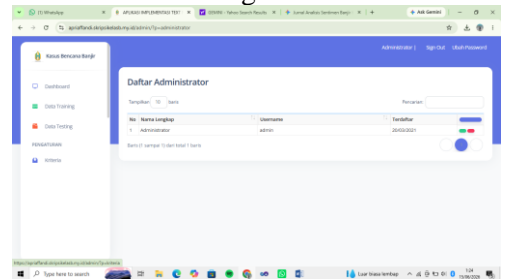
Gambar 6 Halaman Data Training

7. Menu tampilan hasil adalah hasil dari perhitungan data baru setelah di uji dengan data training. Berikut adalah tampilan menu tampilan hasil.



Gambar 7 Tampilan Hasil

8. Menu Administrator digunakan untuk merubah akun dan password. Berikut adalah tampilan menu administrator. g



Gambar 8 Menu Administrator

SIMPULAN

Hasil analisis menunjukkan bahwa metode *K-NEAREST NEIGHBOR* mampu menganalisis data berdasarkan tingkat kedekatan nilai ke dalam beberapa *cluster*, sehingga menghasilkan pengelompokan yang lebih fleksibel dan akurat. Implementasi algoritma *K-NEAREST NEIGHBOR* ke dalam aplikasi yang dirancang berhasil dilakukan. Aplikasi yang dirancang dilengkapi dengan fitur penginputan data, proses *clustering*, serta laporan hasil pengelompokan. Dengan adanya aplikasi ini, Pemerintah dapat mengatasi apa saja

hal yang perlu dilakukan ketika banyak komentar masyarakat yang masuk, baik komentar positif, netral ataupun negatif, serta mendukung proses pengambilan keputusan berdasarkan hasil *clustering* yang diperoleh.

DAFTAR PUSTAKA

Algoritma, P., Forest, R., & Naïve, D. A. N. (2026). *Perbandingan algoritma random forest, svm, dan naïve bayes dalam analisis sentimen ulasan spotify di play store berbasis smote*. 10(1), 17–27.

Dewi, R., & Saleh, M. (2026). *Banjir Sumatera Sebagai Pemicu Kesadaran Publik tentang*

Kerusakan Lingkungan. 2(1), 50–56.

Saputra, A., Nurdiani, I., Nurhidayah, U. S., Maesaroh, S., Informatika, T., Teknik, F., Bakti, U. M., & Barat, T. J. (2026). *Analisis Sentimen Masyarakat Terhadap Penanganan Banjir Bandang di Pulau Sumatra Berdasarkan Komentar Youtube Menggunakan Metode Support Vector Machine (SVM)*. 2(1), 2225–2239.

Safhira, N. (2025). *Penanggulangan Bencana Banjir dan Dampak Psikologis Bagi Masyarakat*. 1(2), 36–40.

Silalahi, Fujiama Diapoldo, S.Kom, M. K. (2022). *Manajemen Databse MySQL, Yayasan Prima Agus Teknik : Semarang*.