

**ANALISIS SENTIMEN OPINI PUBLIK TERHADAP
KESEHATAN MENTAL DI MEDIA SOSIAL MENGGUNAKAN
PERBANDINGAN *METODE LOGISTIC REGRESSION* DAN
*SUPPORT VECTOR MACHINE (SVM)***

TUGAS AKHIR



**ANALISIS SENTIMEN OPINI PUBLIK TERHADAP
KESEHATAN MENTAL DI MEDIA SOSIAL MENGGUNAKAN
PERBANDINGAN *METODE LOGISTIC REGRESSION* DAN
*SUPPORT VECTOR MACHINE (SVM)***

Tugas Akhir

Skema Skripsi

Tugas Akhir diajukan sebagai salah satu persyaratan memperoleh gelar Sarjana

Teknik pada Program Studi Teknik Informatika



**PROGRAM STUDI TEKNIK INFORMATIKA
JURUSAN TEKNIK ELEKTRO DAN INFORMATIKA
FAKULTAS TEKNIK DAN TEKNOLOGI KEMARITIMAN
UNIVERSITAS MARITIM RAJA ALI HAJI
TANJUNGPINANG
2026**

PERNYATAAN MENGENAI TUGAS AKHIR DAN SUMBER INFORMASI SERTA PELIMPAHAN HAK CIPTA

Dengan ini saya menyatakan bahwa tugas akhir berjudul “Analisis Sentimen Opini Publik terhadap Kesehatan Mental di Media Sosial Menggunakan Perbandingan Metode Logistic Regression dan Support Vector Machine (SVM)” adalah benar karya saya dengan arahan dari komisi pembimbing dan belum diajukan dalam bentuk apa pun kepada perguruan tinggi mana pun. Sumber informasi yang berasal atau dikutip dari karya yang diterbitkan maupun tidak diterbitkan dari penulis lain telah disebutkan dalam teks dan dicantumkan dalam Daftar Pustaka di bagian akhir skripsi ini.

Apabila di kemudian hari terbukti pernyataan saya tidak benar dan melanggar peraturan yang berlaku dalam karya tulis dan hak intelektual, maka saya bersedia ijazah yang telah saya terima untuk ditarik kembali oleh Universitas Maritim Raja Ali Haji dan menerima sanksi lainnya sesuai dengan peraturan yang berlaku. Dengan ini saya melimpahkan hak cipta dari karya tulis saya kepada Universitas Maritim Raja Ali Haji.

Tanjungpinang, 22 Juni 2026



Siska Wulndari
NIM 2101020015

The logo of Universitas Maritim Raja Ali Haji is a circular emblem. It features a blue outer ring with the university's name in yellow capital letters. Inside the ring is a yellow triangle with a green and blue vertical stripe. At the base of the triangle is a red and white striped banner. The entire logo is set against a light blue background.

© Hak Cipta Milik Universitas Maritim Raja Ali Haji, Tahun 2026

Hak Cipta Dilindungi Undang- Undang

Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan atau menyebutkan sumbernya. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik, atau tinjauan suatu masalah; dan pengutipan tersebut tidak merugikan kepentingan Universitas Maritim Raja Ali Haji.

Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Universitas Maritim Raja Ali Haji.

HALAMAN PERSETUJUAN

Judul : Analisis Sentimen Opini Publik terhadap Kesehatan Mental di Media Sosial Menggunakan Perbandingan Metode *Logistic Regression* dan *Support Vector Machine* (SVM)

Nama : Siska Wulandari

NIM : 2101020015

Program Studi : Teknik Informatika

Tanggal Persetujuan : 22 Juni 2026

Disetujui oleh,
Komisi Pembimbing


Nurul Havatv, S. T., M. Cs
(Ketua)


Nurfalinda, S. T., M. Cs
(Anggota)

HALAMAN PENGESAHAN

Judul : Analisis Sentimen Opini Publik terhadap Kesehatan Mental di Media Sosial Menggunakan Perbandingan *Metode Logistic Regression* dan *Support Vector Machine (SVM)*

Nama : Siska Wulandari

NIM : 2101020015

Telah berhasil dipertahankan di hadapan Komisi Penguji sebagai bagian persyaratan yang diperlukan dalam memperoleh gelar Sarjana Teknik pada Program Studi Teknik Informatika, Jurusan Teknik Elektro dan Informatika, Fakultas Teknik dan Teknologi Kemaritiman, Universitas Maritim Raja Ali Haji.

DEWAN PENGUJI

Ketua : Martaleli Bettiza, S.Si., M. Sc

Anggota I : Fortia Magfira, M. Kom

Anggota II : Ir. Muhamad Fadli, M. Kom

Anggota III : Nurul Hayaty, S.T., M. Cs

Anggota IV : Nurfalinda, S.T., M. Cs

(.....)

(.....)

(.....)

(.....)

(.....)

Mengetahui,
Ketua Jurusan Teknik Elektro dan Informatika

Hollanda Ariel Kusuma, S. IK., M. Si.

NIP. 198904012019031016

Tanggal Ujian: 13 Juli 2026

Tanggal Lulus:

PRAKATA

Puji dan syukur dipanjatkan kepada Allah *subhanahu wata'ala atas segala* karunia-Nya sehingga Laporan tugas akhir ini berhasil diselesaikan. Tema yang dipilih dalam tugas akhir yang akan dilaksanakan pada bulan Juli 2026 ini ialah analisis sentimen kesehatan mental, dengan judul **"Analisis Sentimen Opini Publik terhadap Kesehatan Mental di Media Sosial Menggunakan Perbandingan Metode *Logistic Regression* dan *Support Vector Machine*"**.

Terima kasih penulis ucapkan kepada Ibu Nurul Hayaty, S. T., M. Cs. dan Ibu Nurfalinda, S. T., M. Cs. yang telah banyak memberi saran. Di samping itu, Ungkapan terima kasih juga disampaikan kepada ayah, ibu, serta seluruh keluarga, atas segala doa dan kasih sayangnya.

Semoga Laporan tugas akhir ini bermanfaat.

Tanjungpinang, 22 Juni 2026



Siska Wulandari

DAFTAR ISI

HALAMAN PERSETUJUAN	v
HALAMAN PENGESAHAN	vi
ABSTRAK	vii
ABSTRACT	viii
HALAMAN PERSEMBAHASAN	ix
HALAMAN MOTO	xi
PRAKATA	xii
DAFTAR ISI	xiii
DAFTAR TABEL	xvi
DAFTAR GAMBAR	xix
BAB I PENDAHULUAN	1
A. Latar Belakang	1
B. Rumusan Masalah	3
C. Batasan Penelitian	3
D. Tujuan Penelitian	3
E. Manfaat Penelitian	4
BAB II KAJIAN LITERATUR	5
A. Tinjauan Pustaka	5
B. Landasan Teori	14
1. Analisis Sentimen	14
2. Kesehatan Mental	15
3. YouTube	15
4. <i>Text Mining</i>	15
5. <i>Web Scraping</i>	16
6. <i>Data Preprocessing</i>	16
7. <i>Term Frequency-Inverse Document Frequency (TF-IDF)</i>	18

8.	<i>Google Colaboratory</i>	19
9.	Python	19
10.	Indonesia Sentiment Lexicon (<i>InSet Lexicon</i>)	20
11.	<i>Logistic Regression</i>	21
12.	<i>Support Vector Machine</i> (SVM)	22
13.	Evaluasi	25
BAB III METODE PENELITIAN		27
A.	Waktu dan Tempat Penelitian	27
B.	Alat atau Instrumen Penelitian	27
C.	Tahapan Penelitian	27
D.	Pengumpulan Data	28
E.	Analisis dan Perancangan	30
1.	Data Preprocessing	32
2.	Data Set	40
3.	Pembobotan Kata	41
F.	Perancangan <i>Logistic Regression</i>	49
G.	Perancangan <i>Support Vector Machine</i>	66
H.	Perancangan Sistem	82
BAB IV HASIL DAN PEMBAHASAN		87
A.	Pelatihan Model	87
1.	Pelatihan Model <i>Logistic Regression</i>	88
2.	Pelatihan Model <i>Support Vector Machine</i> (SVM)	88
B.	Pengujian Model	89
1.	Pengujian Model <i>Logistic Regression</i>	89
2.	Pengujian Model <i>Support Vector Machine</i>	91
C.	Analisis Perbandingan Kinerja Model	92
D.	Visualisasi	95
E.	Implementasi Antarmuka Aplikasi	98
BAB V SIMPULAN DAN SARAN		103
A.	Simpulan	103

B. Saran	103
DAFTAR PUSTAKA	105
Lampiran 1. Hasil Validasi Data Komentor Drs.Suhardi, M.Pd	108

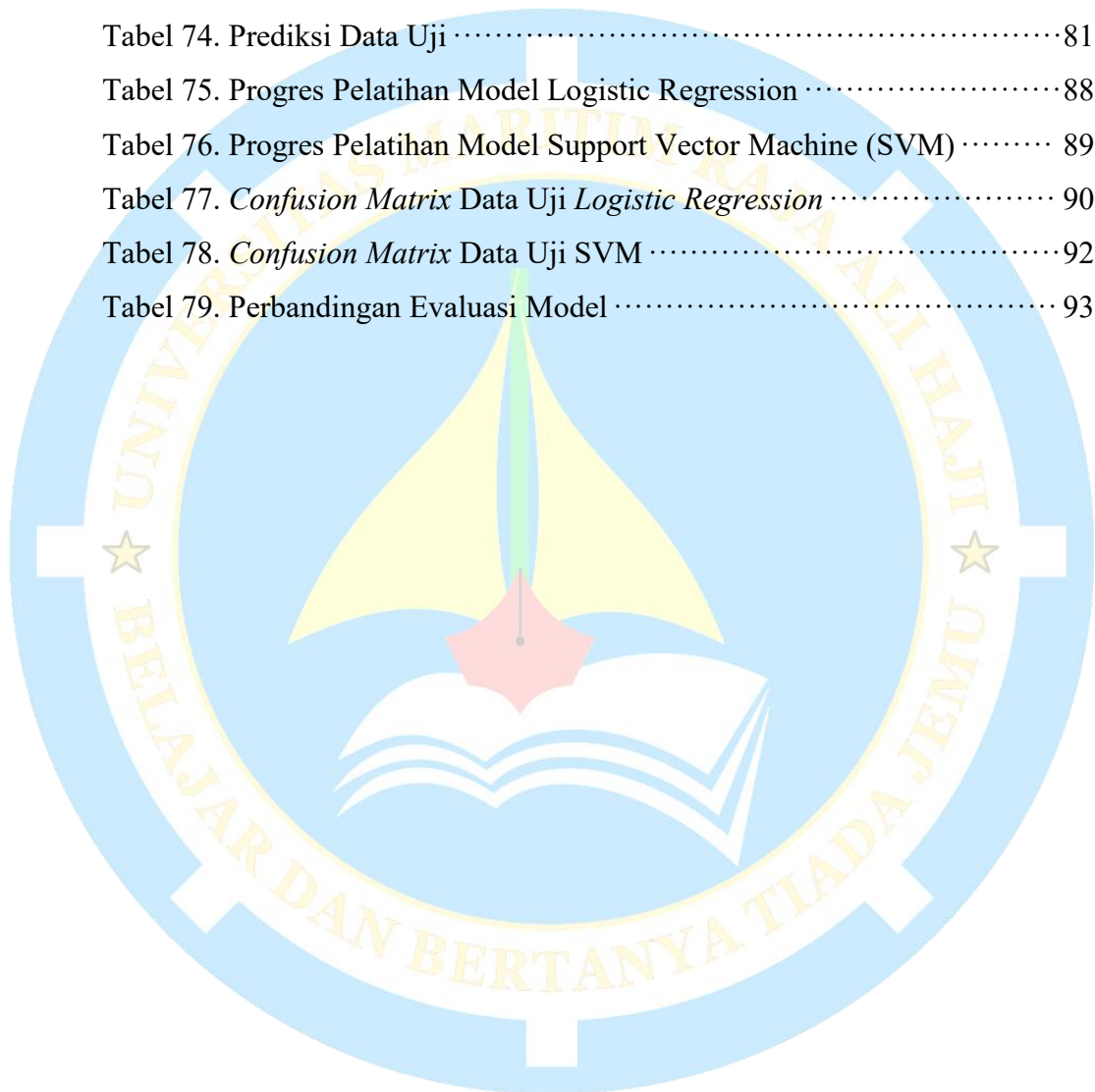


DAFTAR TABEL

Tabel 1. Penelitian Yang Pernah Dilakukan	8
Tabel 2. Penelitian Yang Pernah Dilakukan (Lanjutan)	11
Tabel 3. <i>Confusion matix</i>	25
Tabel 4. Data Latih Case Folding	33
Tabel 5. Data Uji Case Folding	33
Tabel 6. <i>Cleansing</i>	34
Tabel 7. Data Uji Cleansing	35
Tabel 8. Tokenizing	35
Tabel 9. Data Uji Tokenizing	36
Tabel 10. Filterisasi	37
Tabel 11. Data Uji Filterisasi	38
Tabel 12. Stemming	38
Tabel 13. Data Uji Stemming	39
Tabel 14. Data <i>Feature Term</i>	40
Tabel 15. Data Uji	40
Tabel 16. Pembobotan Kata	41
Tabel 17. Perhitungan IDF	42
Tabel 18. Perhitungan TF-IDF	44
Tabel 19. Data UJI	45
Tabel 20. Hitung TF Data UJI	45
Tabel 21. Nilai DF Data Uji	46
Tabel 22. Nilai IDF Data Uji	47
Tabel 23. Hasil TF-IDF data Uji	48
Tabel 24. Parameter	51
Tabel 25. Bobot dan Bias Iterasi D1	51
Tabel 26. Iterasi D1	52
Tabel 27. Bobot & Bias Akhir Iterasi D1	53
Tabel 28. Bobot dan Bias Awal Iterasi D2	54
Tabel 29. Iterasi D2	54
Tabel 30. Bobot & Bias Akhir Iterasi D2	56
Tabel 31. Bobot dan Bias Awal Iterasi D3	57
Tabel 32. Iterasi D3	57

Tabel 33. Bobot dan Bias Akhir Iterasi D3	59
Tabel 34. Bobot dan Bias awal Iterasi D4	59
Tabel 35. Iterasi D4	60
Tabel 36. Bobot dan Bias akhir Iterasi D4	62
Tabel 37. Bobot dan Bias awal Iterasi D5	62
Tabel 38. Iterasi D5	63
Tabel 39. Bobot dan Bias Akhir Iterasi D5	65
Tabel 40. Hasil Prediksi Akhir	65
Tabel 41. Inisialisasi Parameter	68
Tabel 42. Label Data Classifier (Ovr)	68
Tabel 43. Perhitungan Iterasi D1	69
Tabel 44. Update bobot dan bias D1	70
Tabel 45. Perhitungan iterasi D2	70
Tabel 46. Update bias iterasi D2	71
Tabel 47. Perhitungan iterasi D3	71
Tabel 48. Update bias iterasi D3	71
Tabel 49. Perhitungan iterasi D4	72
Tabel 50. Update bobot dan bias iterasi D4	72
Tabel 51. Perhitungan iterasi D5	72
Tabel 52. Update bias iterasi D5	73
Tabel 53. Perhitungan Iterasi D1	73
Tabel 54. Update bias Iterasi D1	73
Tabel 55. Perhitungan Iterasi D2	74
Tabel 56. Update bias Iterasi D2	74
Tabel 57. Perhitungan Iterasi D3	75
Tabel 58. Update bias Iterasi D3	75
Tabel 59. Perhitungan Iterasi D4	75
Tabel 60. Update bias Iterasi D4	76
Tabel 61. Perhitungan Iterasi D5	76
Tabel 62. Update bias Iterasi D5	76
Tabel 63. Perhitungan Iterasi D1	77
Tabel 64. Update bias Iterasi D1	77
Tabel 65. Perhitungan Iterasi D2	78
Tabel 66. Update bias Iterasi D2	78

Tabel 67. Perhitungan Iterasi D3	78
Tabel 68. Update bias Iterasi D3	79
Tabel 69. Perhitungan Iterasi D4	79
Tabel 70. Update bias Iterasi D4	79
Tabel 71. Perhitungan Iterasi D5	80
Tabel 72. Update bias Iterasi D5	80
Tabel 73. Rekapitulasi parameter akhir ketiga classifier	81
Tabel 74. Prediksi Data Uji	81
Tabel 75. Progres Pelatihan Model Logistic Regression	88
Tabel 76. Progres Pelatihan Model Support Vector Machine (SVM)	89
Tabel 77. <i>Confusion Matrix</i> Data Uji <i>Logistic Regression</i>	90
Tabel 78. <i>Confusion Matrix</i> Data Uji SVM	92
Tabel 79. Perbandingan Evaluasi Model	93



DAFTAR GAMBAR

Gambar 1. YouTube	15
Gambar 2. Python	20
Gambar 3. Prosedur Penelitian	28
Gambar 4. Mengidentifikasi sumber data	29
Gambar 5. Persiapan Sistem	29
Gambar 6. Menginstal Library	30
Gambar 7. Menjalankan Program	30
Gambar 8. Menyimpan Data dalam Format CSV	30
Gambar 9. Arsitektur sistem	31
Gambar 12. Flowchart <i>Logistic Regression</i>	50
Gambar 13. Flowchart SVM	67
Gambar 14. Use Case Diagram	83
Gambar 15. Activity Diagram	84
Gambar 16. <i>User Interface/UI</i>	86
Gambar 17. <i>Confusion Matrix Logistic Regression</i>	90
Gambar 18. <i>Confusion Matrix SVM</i>	91
Gambar 19. Grafik Perbandingan Evaluasi Model	93
Gambar 20. <i>Wordcloud Sentimen Positif Logistic Regression</i>	96
Gambar 21. <i>Wordcloud Sentimen Netral Logistic Regression</i>	96
Gambar 22. <i>Wordcloud Sentimen Negatif Logistic Regression</i>	96
Gambar 23. <i>Wordcloud Sentimen Positif (SVM)</i>	97
Gambar 24. <i>Wordcloud Sentimen Netral (SVM)</i>	97
Gambar 25. <i>Wordcloud Sentimen Negatif (SVM)</i>	98
Gambar 26. Halaman Beranda	98
Gambar 27. Halaman Klasifikasi Sentimen	99
Gambar 28. Halaman Analisis Perbandingan	99
Gambar 29. Halaman Analisis Perbandingan <i>Preprocessing</i>	100
Gambar 30. Halaman InSet Lexicon	100
Gambar 31. Halaman Analisis Perbandingan Prediksi	101
Gambar 32. Halaman Analisis Perbandingan Visualisasi	101
Gambar 33. Halaman Analisis Perbandingan Evaluasi Model	102

BAB I

PENDAHULUAN

A. Latar Belakang

Kualitas hidup seseorang sangat dipengaruhi oleh kondisi kesehatan mentalnya. Individu yang secara mental sehat umumnya mampu berpikir dengan jernih, mengelola emosi dengan baik, serta tetap produktif dalam menjalani relasi sosialnya. Sebaliknya, gangguan seperti depresi, kecemasan, atau stres kerap mengganggu aktivitas sehari-hari dan berdampak buruk pada kesejahteraan individu maupun lingkungan di sekitarnya. Perhatian publik terhadap isu kesehatan mental di Indonesia terus meningkat seiring semakin banyaknya kasus yang terungkap melalui media massa maupun platform digital. *YouTube* misalnya, tidak lagi sekadar menjadi sarana hiburan, tetapi juga wadah bagi masyarakat untuk berekspresi. Melalui kolom komentar, siapa pun dapat menyuarakan pandangannya secara bebas, mulai dari bentuk dukungan hingga stigma terhadap penderita gangguan mental. Kondisi ini memperlihatkan bahwa opini yang beredar di media sosial mencerminkan keragaman kesadaran dan sikap masyarakat, baik yang suportif maupun yang justru bersifat diskriminatif [1] [2].

Keragaman opini publik tersebut penting untuk dipahami lebih dalam. Sebagian warganet mendorong keterbukaan dalam membicarakan kesehatan mental, sementara sebagian lainnya masih menunjukkan sikap negatif yang justru memperkuat stigma yang ada. Perbedaan persepsi semacam ini berpotensi memperlebar jarak antara individu dengan gangguan mental dan lingkungan sekitarnya. Analisis sentimen merupakan salah satu pendekatan yang dapat digunakan untuk membaca pola opini publik tersebut. Melalui metode *text mining*, opini masyarakat dapat dikelompokkan ke dalam kategori positif, netral, dan negatif, sehingga peneliti dapat mengolah data komentar dalam jumlah besar secara kuantitatif dan memperoleh gambaran yang lebih objektif mengenai sikap publik terhadap isu kesehatan mental [3].

Secara umum, analisis sentimen merupakan teknik otomatis untuk menilai apakah suatu teks mengandung pandangan yang positif, negatif, atau netral terhadap suatu objek tertentu [4]. Hasil pengolahan dan analisis tersebut divisualisasikan agar dapat mendukung pengambilan keputusan strategis, baik untuk keperluan akademis

maupun praktis, misalnya dalam menyusun strategi komunikasi kesehatan mental di ruang publik digital.

Di antara berbagai metode analisis data, *Logistic Regression* dan *Support Vector Machine (SVM)* tergolong populer dan masing-masing memiliki keunggulan tersendiri. *Logistic Regression* merupakan model statistik yang digunakan untuk menilai pengaruh variabel independen terhadap variabel dependen biner dan menghasilkan probabilitas antara 0 dan 1 melalui fungsi *sigmoid*, sehingga cocok untuk prediksi yang bersifat probabilistik. SVM di sisi lain adalah metode pembelajaran terbimbing yang mampu mengenali pola dalam data untuk keperluan klasifikasi maupun regresi, dengan membangun *hyperplane* optimal guna memisahkan kelas secara maksimal. Perpaduan antara kekuatan probabilistik *Logistic Regression* dan kemampuan pemisahan kelas pada SVM menjadikan kedua metode ini relevan dan saling melengkapi dalam berbagai aplikasi analisis data dan prediksi [5].

Salah satu penelitian terdahulu yang relevan dengan topik ini [11] berfokus pada analisis sentimen kemungkinan depresi dan kecemasan menggunakan data dari platform Twitter sebanyak 2.701 data dengan dua kelas, yaitu positif dan negatif. Metode *Support Vector Machine (SVM)* yang digunakan pada penelitian tersebut menghasilkan akurasi sebesar 82,5%.

Mengacu pada uraian di atas, penelitian ini membangun sebuah model analisis sentimen berbasis *machine learning* untuk mengklasifikasikan opini publik terkait kesehatan mental pada kolom komentar *YouTube* ke dalam tiga kelas, yaitu positif, negatif, dan netral. Komentar dikumpulkan dari video *YouTube* yang membahas isu kesehatan mental, kemudian melalui tahap pra-pemrosesan teks sebelum dijadikan data latih dan data uji. Selanjutnya, dua metode klasifikasi, yaitu *Logistic Regression* dan *Support Vector Machine (SVM)*, diterapkan dan dibandingkan kinerjanya menggunakan metrik evaluasi berupa akurasi, *precision*, *recall*, dan *F1-score* guna menentukan metode yang paling optimal dalam melakukan analisis sentimen.

Selama proses penyusunan penelitian ini, peneliti menghadapi berbagai tantangan, baik dalam memahami isu kesehatan mental secara komprehensif maupun dalam mengelola komentar *YouTube* yang gaya bahasanya sangat beragam. Analisis sentimen yang dilakukan tidak hanya bertujuan mengelompokkan opini publik, tetapi juga diarahkan menjadi landasan penyusunan upaya edukasi yang membantu masyarakat memahami pentingnya kesehatan mental dan mengurangi stigma yang

masih muncul di ruang digital. Penyesuaian metodologi, revisi yang berkelanjutan, serta pengelolaan data dalam jumlah besar menjadi bagian dari proses pembelajaran sekaligus memperkuat motivasi peneliti untuk menghasilkan karya yang tidak hanya bernilai akademis, tetapi juga memberi dampak nyata bagi masyarakat.

B. Rumusan Masalah

Berdasarkan latar belakang yang telah dipaparkan, maka rumusan masalah dalam penelitian ini adalah bagaimana sentimen opini publik terhadap isu kesehatan mental pada komentar di kanal YouTube Raditya Dika, Rory Asyari, Helmy Yahya Bicara, dan Tirta PengPengPeng, serta bagaimana perbandingan performa metode *Logistic Regression* dan *Support Vector Machine* (SVM) dalam mengklasifikasikan sentimen komentar tersebut?

C. Batasan Penelitian

Untuk menyederhanakan permasalahan dalam usulan penelitian ini diperlukan adanya batasan-batasan yaitu:

1. Data yang digunakan berasal dari komentar pengguna pada platform Youtube yang membahas isu kesehatan mental.
2. Sumber video yang dianalisis pada kategori podcast dari kanal Raditya Dika, Rory Asyari, Helmy Yahya Bicara, dan Tirta PengPengPeng dengan tahun publikasi 2024-2025 membahas isu kesehatan mental dan berjumlah 9 video, sehingga tidak mencakup seluruh video dalam kanal tersebut.
3. Metode analisis yang digunakan yang mencakup dua algoritma klasifikasi yaitu *Logistic Regression* dan *Support Vector Machine* (SVM).
4. Hasil klasifikasi sentimen dibatasi pada tiga kategori, yaitu sentimen positif, negatif dan sentimen netral.

D. Tujuan Penelitian

Tujuan penelitian ini adalah untuk mengetahui sentimen opini publik terhadap isu kesehatan mental pada komentar di kanal Youtube Raditya Dika, Helmy Yahya Bicara, Rory Asyari, dan Tirta PengPengPeng, serta membandingkan kinerja metode *Logistic Regression* dan *Support Vector Machine* (SVM) dalam mengklasifikasikan sentimen komentar tersebut.

E. Manfaat Penelitian

1. Memberikan kontribusi akademik yang relevan dalam pengembangan analisis sentimen pada isu kesehatan mental di media sosial, khususnya dengan memanfaatkan komentar Youtube yang masih jarang digunakan sebagai sumber data penelitian.
2. Menyediakan gambaran sentimen dan persepsi publik terhadap kesehatan mental, yang dapat dijadikan dasar dalam penyusunan upaya edukasi untuk meningkatkan pemahaman masyarakat serta mengurangi stigma yang masih muncul di ruang digital.
3. Meningkatkan pemahaman tentang penerapan metode *Logistic Regression* dan *Support Vector Machine (SVM)* dalam analisis data teks, khususnya dalam menganalisis sentimen pada komentar- komentar online.

