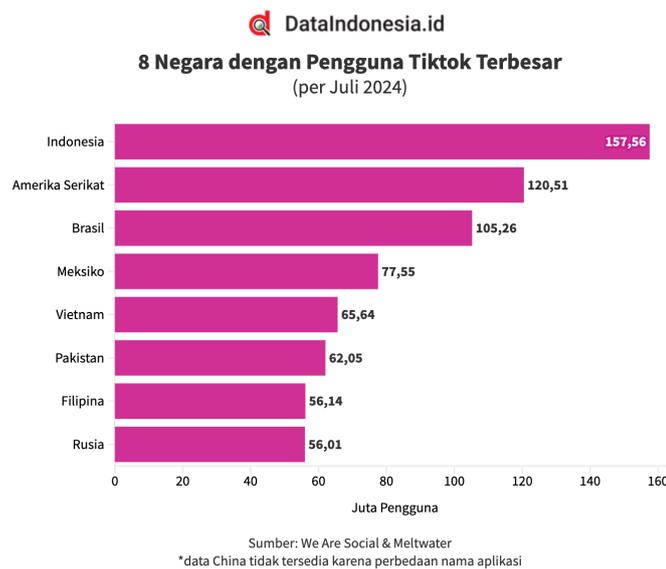


# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Di era digital saat ini, media sosial telah berkembang menjadi wadah bagi individu untuk berbagi informasi. Salah satu *platform* media sosial berbasis audio-visual yang populer saat ini adalah TikTok. Tiktok merupakan *platform* media sosial yang memungkinkan pengguna berbagi gambar, video, dan lainnya [1]. Aplikasi ini telah menarik minat jutaan pengguna dari berbagai kelompok usia, latar belakang, maupun budaya. Keunikan TikTok dibandingkan dengan *platform* media sosial lainnya terletak pada algoritma canggih yang menyesuaikan konten yang ditampilkan pada *feeds* pengguna sesuai dengan minat dan preferensi mereka [2]. Dengan data dari delapan negara dengan jumlah pengguna TikTok tertinggi di seluruh dunia, Gambar 1.1 menunjukkan bahwa Indonesia adalah negara dengan pengguna TikTok terbanyak di seluruh dunia.



Gambar 1. 1 Negara dengan Pengguna TikTok Terbanyak

sumber: dataindonesia.id [3]

Gambar 1.1 merupakan visualisasi yang menunjukkan jumlah pengguna TikTok di berbagai negara per Juli 2024. Dari gambar tersebut dapat dilihat bahwa Indonesia memiliki jumlah pengguna TikTok terbanyak dibandingkan

dengan negara lainnya dengan memiliki 157,56 Juta pengguna. Hal ini memperlihatkan betapa besar pengaruh TikTok di kalangan masyarakat di Indonesia. Setelah Indonesia, pada peringkat kedua terdapat Amerika Serikat dengan total 120,51 Juta pengguna, dan pada peringkat ketiga terdapat Brasil dengan total 105,26 Juta pengguna [3]. Keberagaman pengguna dari berbagai negara ini menunjukkan bahwa TikTok telah berhasil menjangkau audiens secara global.

Dari popularitasnya yang tinggi, TikTok dijadikan sebagai sumber informasi berita selain menjadi *platform* hiburan. Selain itu, TikTok juga memungkinkan penggunanya untuk berbagi pengetahuan, wawasan, dan berbagai informasi di berbagai bidang dalam masyarakat [4]. Salah satu jenis konten yang dapat ditemukan di TikTok adalah konten yang berkaitan dengan kesehatan mental. TikTok berhasil menciptakan pengalaman yang mengalir dan memberikan pemahaman baru kepada penggunanya melalui konten-konten yang edukatif dan inspiratif. Beragam fitur menarik yang dimiliki oleh TikTok memungkinkan para kreator untuk menyampaikan informasi dengan cara yang kreatif dan unik, sehingga banyak pengguna yang tertarik untuk mengikuti paparan terkait dengan kesehatan mental [5].

Kesehatan mental adalah kondisi dimana fungsi-fungsi psikologis bekerja secara seimbang, memiliki kemampuan dalam mengatasi masalah yang dihadapi, serta dapat merasakan bahagia dan memiliki pandangan positif mengenai dirinya [6]. Kesehatan mental merupakan elemen yang krusial dalam definisi kesehatan secara umum. Individu dengan kesehatan mental yang baik mampu melihat potensi mereka sendiri, juga dapat mengatasi tantangan hidup sehari-hari, bekerja dengan produktif, serta berperan aktif dalam masyarakat. Dengan demikian, gangguan kesehatan mental tidak dapat dianggap remeh, karena jumlah kasus yang ada saat ini masih cukup mengkhawatirkan [7].

Pada negara berkembang, terdapat sekitar 85% orang dengan gangguan kesehatan mental parah yang belum mendapatkan perawatan yang memadai. Kondisi serupa juga dialami di Indonesia, di mana faktor-faktor seperti biologis, psikologis, sosial, serta keragaman populasi berpotensi dalam meningkatkan

jumlah kasus gangguan mental [8]. Gambar 1.2 merupakan data gangguan kesehatan mental yang dikhawatirkan masyarakat Indonesia pada tahun 2024.



Gambar 1. 2 Gangguan Mental Paling Dikhawatirkan di Indonesia (2024)

Sumber: GoodStats [9]

Gambar 1.2 merupakan data gangguan kesehatan mental yang dikhawatirkan oleh masyarakat Indonesia pada tahun 2024. Berdasarkan *Asia Care Survey 2024* yang melibatkan lebih dari 1.000 responden, *stress* dan *burnout* menjadi gangguan kesehatan mental yang paling dikhawatirkan oleh masyarakat, dengan 56% responden yang menyebutkan bahwa hal ini merupakan kekhawatiran pertama. Pada urutan kedua, terdapat gangguan tidur dengan total 42.6% responden yang merasa cemas mengenai hal tersebut. Selanjutnya terdapat 28.2% responden yang khawatir mengenai kecemasan, sementara itu terdapat 24.9% lainnya yang mengkhawatirkan kesepian. Depresi juga menjadi perhatian dengan total 20.7% responden yang mengaku khawatir mengalaminya, dan gangguan kognitif mencatatkan angka 9.1% [9].

Kekhawatiran masyarakat Indonesia terhadap berbagai masalah kesehatan mental, seperti yang ditunjukkan dalam gambar 1.2, semakin memperkuat pentingnya akses terhadap informasi yang tepat dan juga akurat. Menurut penelitian sebelumnya, penggunaan media sosial dapat berdampak negatif terhadap kesehatan mental, terutama dalam meningkatkan kecemasan dan *stress* [10][11]. Ketika informasi mengenai kesehatan mental banyak disajikan di media sosial seperti TikTok tanpa pengawasan yang memadai, risiko misinformasi

menjadi tidak dapat dihindarkan. TikTok menggunakan algoritma yang menyesuaikan konten berdasarkan interaksi penggunanya untuk meningkatkan keterlibatan. Namun, jika seseorang terus menerus melihat konten yang sama, hal ini dapat menciptakan ruang yang tertutup yang memperkuat pandangan tertentu dan berpotensi memperburuk masalah mengenai kesehatan mental [12]. Adanya informasi yang disebarluaskan melalui TikTok mengenai kesehatan mental ini seringkali tidak disertai dengan sikap yang bijak dari pengguna TikTok. Banyak orang yang cenderung menerima informasi begitu saja tanpa memeriksa kebenarannya dan kredibilitas sumbernya. Akibat dari perilaku tersebut, tindakan mendiagnosis penyakit mental menjadi hal yang dianggap biasa [13]. Tindakan tersebut sering dikenal sebagai *self-diagnose*.

*Self-diagnose* merujuk pada situasi di mana individu menilai kondisi kesehatannya sendiri berdasarkan informasi yang mereka ketahui [14]. *Self-diagnose* dapat berdampak pada kesehatan mental dengan memicu kecemasan yang tidak perlu. Kecemasan ini seringkali muncul akibat mempercayai informasi yang tidak valid atau tidak berasal dari sumber yang kompeten. Menilai kondisi diri sendiri tanpa bantuan profesional dapat beresiko karena setiap penyakit atau gangguan memerlukan perawatan yang sesuai. Jika penanganannya salah, kondisi kesehatan dapat semakin memburuk [15]. Selain menyebabkan kecemasan yang berlebihan dan adanya penanganan yang salah, *self-diagnose* dapat menyebabkan kesulitan dalam mengidentifikasi kondisi kesehatan yang sebenarnya, serta timbulnya pemahaman yang keliru terkait dengan gangguan kesehatan mental [16]. Dengan demikian, fenomena *self-diagnose* yang berkembang di TikTok menjadi semakin berisiko, mengingat banyaknya informasi yang tidak selalu akurat dan kredibel. Tanpa disertai pengawasan yang memadai, pengguna TikTok dapat dengan mudah terjebak dengan kesalahpahaman mengenai kondisi kesehatan mental mereka, yang pada akhirnya hanya dapat memperburuk kondisi tersebut jika tidak ditangani dengan benar.

Paparan yang semakin intens terhadap konten kesehatan mental di *platform* TikTok telah memunculkan kekhawatiran baru, khususnya terkait pola perilaku pengguna muda. Salah satu tren yang berkembang adalah meningkatnya praktik *self-diagnose*, yakni penilaian kondisi psikologis secara mandiri tanpa melibatkan

tenaga profesional. Salah satu studi survei menunjukkan adanya korelasi antara intensitas paparan remaja terhadap video bertema gangguan mental di TikTok dengan tingkat kepercayaan mereka terhadap informasi yang disajikan. Semakin sering remaja mengakses konten semacam ini, semakin tinggi pula kecenderungan mereka untuk menyimpulkan kondisi psikologisnya secara mandiri tanpa melalui proses diagnosis yang sah secara medis [17]. Temuan serupa juga diperoleh dari hasil analisis konten yang menunjukkan bahwa video dengan tagar seperti *#anxiety* dan *#depression* yang menampilkan narasi pengalaman pribadi cenderung mendapatkan tingkat keterlibatan yang tinggi dari pengguna. Meskipun hal ini mencerminkan tingginya minat terhadap isu kesehatan mental, jenis konten semacam itu dapat memunculkan interpretasi yang keliru, terutama karena tidak disertai dengan arahan dari pihak profesional. Kondisi ini menunjukkan bahwa meskipun TikTok dapat berfungsi sebagai sarana penyebaran kesadaran akan isu kesehatan mental, di sisi lain *platform* ini juga memiliki risiko dalam membentuk pemahaman yang menyimpang, terutama di kalangan pengguna yang belum memiliki pengetahuan psikologis yang memadai [18].

Di samping itu, persebaran konten mengenai ADHD yang menjadi viral di TikTok turut dikaitkan dengan meningkatnya diskusi terkait *self-diagnose* di berbagai *platform* media sosial lainnya, seperti Twitter. Bahkan, tidak sedikit pengguna yang menyatakan bahwa mereka merasa memiliki gejala ADHD setelah menonton video tersebut [19]. Fenomena ini tidak semata-mata dipengaruhi oleh kontennya saja, melainkan juga didorong oleh berbagai faktor psikologis seperti kecenderungan bias konfirmasi, kebutuhan untuk mengenali jati diri, serta keinginan untuk mendapatkan validasi sosial. Kesesuaian juga ditemukan dalam hasil studi eksperimental yang menunjukkan bahwa meskipun konten bertema kesehatan mental di TikTok tidak secara langsung memperkuat stigma terhadap individu dengan gangguan psikologis, namun justru dapat mendorong munculnya *self-pathologizing*. Kondisi ini merujuk pada kecenderungan seseorang untuk meyakini bahwa dirinya mengalami gangguan mental, meskipun belum pernah mendapatkan diagnosis resmi dari ahli [20].

Berbagai penelitian terdahulu menunjukkan bahwa penting untuk menelaah bagaimana masyarakat memaknai dan merespons konten kesehatan mental,

khususnya yang tersebar melalui platform media sosial seperti TikTok. Salah satu pendekatan yang dinilai tepat untuk menjawab kebutuhan tersebut adalah analisis sentimen. Metode ini memungkinkan peneliti untuk mengidentifikasi opini serta emosi yang terkandung dalam teks tidak terstruktur, seperti komentar pengguna, dalam jumlah besar. Kemampuan ini menjadikan analisis sentimen sebagai alat yang efektif dalam mengungkap pola persepsi publik terhadap isu-isu psikologis di dunia digital [21]. Salah satu hasil survei menunjukkan bahwa sebagian besar responden, yakni mencapai 89%, mengaku pernah melakukan *self-diagnose* setelah menyaksikan konten psikologi di TikTok. Fakta ini memperkuat urgensi dilakukannya kajian yang lebih mendalam terhadap komentar-komentar pengguna, sebagai cerminan dari respons emosional mereka. Pendekatan ini diperlukan untuk memahami sejauh mana paparan terhadap konten tersebut dapat memengaruhi persepsi individu terhadap kondisi psikologisnya sendiri [22].

Salah satu penelitian berskala besar yang menggunakan data sebanyak 1,6 juta tweet menunjukkan bahwa algoritma XGBoost mampu mencapai skor *f1 score* tertinggi, yaitu 0.966, sementara SVM unggul dalam aspek precision dengan nilai 0.963. Temuan ini menunjukkan bahwa XGBoost lebih unggul dalam menyeimbangkan aspek *precision* dan *recall*, khususnya dalam pengolahan data berukuran besar [11]. Pada studi lain, membuktikan bahwa kombinasi SVM dengan metode representasi teks TF-IDF mampu mencapai akurasi hingga 89% serta precision sebesar 90% dalam mendeteksi opini publik seputar pandemi, menjadikannya sangat relevan untuk konteks teks pendek di media sosial [23]. Sementara itu, pada penelitian selanjutnya menunjukkan bahwa XGBoost memiliki performa terbaik dalam mendeteksi gejala depresi, baik dari sisi akurasi maupun nilai *f1-score*. Dalam isu sosial seperti perundungan digital, SVM juga menunjukkan keandalannya dengan *recall* mencapai 96%, membuktikan kemampuannya dalam mengenali konten yang cenderung ekstrem [24]. Selain itu, XGBoost juga sangat efektif untuk menangani data yang tidak seimbang dengan capaian *f1-score* mencapai 97% dalam klasifikasi gangguan depresi berbasis biomarker [25].

Berdasarkan temuan dari sejumlah penelitian sebelumnya yang membuktikan keandalan algoritma SVM dan XGBoost dalam mengolah data teks, kedua

algoritma tersebut dipilih sebagai metode utama dalam penelitian ini. Dengan mengambil data berupa komentar dari *platform* TikTok, penelitian ini mencoba melihat bagaimana tanggapan pengguna terhadap konten kesehatan mental, sekaligus mengamati adanya kecenderungan melakukan *self-diagnose*. Tujuan utama dari penelitian ini adalah untuk mengetahui bagaimana pengguna memaknai dan merespons informasi yang mereka temui, serta apakah hal tersebut memicu penilaian terhadap kondisi psikologis mereka sendiri tanpa bantuan tenaga ahli. Selain itu, penelitian ini juga dimaksudkan untuk mengevaluasi tingkat akurasi dari kedua algoritma dalam mengklasifikasikan sentimen yang muncul dari komentar pengguna. Diharapkan, hasil penelitian ini tidak hanya memberikan gambaran tentang pola interaksi pengguna terhadap konten kesehatan mental, tetapi juga menjadi pijakan awal dalam mendorong masyarakat agar lebih selektif dan bijak dalam menyikapi informasi yang mereka terima di media sosial, khususnya TikTok.

## 1.2 Rumusan Masalah

Dari latar belakang yang telah diuraikan di atas, maka diperoleh beberapa rumusan masalah, antara lain:

1. Bagaimana hasil pelabelan komentar pengguna TikTok mengenai konten kesehatan mental berdasarkan indikasi *self-diagnose*?
2. Bagaimana performa algoritma *Support Vector Machine* (SVM) dan XGBoost dalam melakukan klasifikasi sentimen (positif, netral, negatif) dari komentar TikTok terkait konten kesehatan mental?
3. Bagaimana pola distribusi terkait komentar *self-diagnose* terhadap waktu, panjang komentar, serta interaksi pengguna (*likes* dan *replies*)?

## 1.3 Batasan Masalah

Agar penelitian ini tetap berada dalam ruang lingkupnya, penelitian ini memiliki sejumlah batasan masalah, yaitu:

1. Konten TikTok yang berkaitan dengan kesehatan mental, yaitu konten yang menggunakan *hashtag* *#mentalhealth*, *#anxiety*, dan *#depression*, menjadi subjek penelitian ini.

2. Data penelitian berasal dari komentar yang dapat diakses publik di aplikasi TikTok. Informasi pribadi pengguna tidak akan digunakan.
3. Data yang diambil yaitu data komentar dari konten kesehatan mental pada TikTok di tahun 2024.
4. Analisis sentimen dalam penelitian ini akan menggunakan algoritma SVM dan XGBoost.
5. Penelitian ini tidak mengevaluasi keakuratan medis dari konten kesehatan mental pada TikTok, melainkan hanya berfokus pada sentimen yang muncul dari interaksi pengguna.
6. Penelitian terbatas pada data yang dikumpulkan dalam rentang waktu tertentu.
7. Penelitian ini tidak membahas dampak psikologis langsung dari fenomena *self-diagnose*, melainkan hanya mengidentifikasi dan menganalisis pola perilaku dari data yang diperoleh.
8. Skema pelabelan dalam penelitian ini mencakup tiga kategori: positif, netral, dan negatif. Label positif diberikan pada komentar yang secara eksplisit menyatakan bahwa pengguna merasa mengalami gangguan mental setelah melihat konten, tanpa melalui diagnosis medis. Label netral mencakup komentar yang membahas topik serupa namun tidak menunjukkan pengakuan atau penolakan secara jelas, serta yang berpotensi mengarah pada *self-diagnose* secara implisit. Sementara itu, label negatif mencakup komentar yang menolak *self-diagnose*, bersifat umum, tidak relevan, atau menyatakan kondisi mental berdasarkan hasil konsultasi profesional.
9. Penelitian ini mengacu pada tahapan *Cross Industry Standard Process for Machine Learning* (CRISP-ML) dalam pelaksanaan proses analisis, mulai dari pemahaman bisnis hingga evaluasi model. Namun, fase *monitoring* dan *maintenance* tidak dilakukan karena penelitian ini ditujukan untuk kepentingan akademis dan tidak dilanjutkan ke tahap implementasi sistem secara berkelanjutan di lingkungan produksi.

## 1.4 Tujuan dan Manfaat Penelitian

### 1.4.1 Tujuan Penelitian

Berikut ini adalah tujuan dari penelitian ini berdasarkan beberapa rumusan masalah yang ada, yaitu:

1. Melakukan pelabelan otomatis terhadap komentar TikTok yang berkaitan dengan konten kesehatan mental berdasarkan indikasi *self-diagnose* menggunakan metode *rule-based labeling* berbasis kata kunci dan struktur kalimat tertentu. Untuk memastikan akurasi hasil pelabelan, dilakukan validasi oleh pakar dengan latar belakang psikologi.
2. Membangun dan mengevaluasi model klasifikasi sentimen menggunakan Support Vector Machine (SVM) dan XGBoost dengan mengukur performa berdasarkan metrik evaluasi berupa akurasi, presisi, *recall*, dan *f1-score*.
3. Menganalisis pola distribusi komentar *self-diagnose* berdasarkan waktu komentar, panjang komentar, serta interaksi pengguna (*likes* dan *replies*).

### 1.4.2 Manfaat Penelitian

Penelitian ini juga memiliki sejumlah manfaat yang berguna bagi beberapa pihak. Manfaat teoritis dari penelitian ini, yaitu:

1. Penelitian ini memberikan kontribusi dalam pengembangan keilmuan di bidang analisis sentimen dan kecerdasan buatan, khususnya dalam mengklasifikasikan teks pendek yang berasal dari interaksi pengguna di media sosial mengenai isu kesehatan mental.
2. Penelitian ini juga memperluas literatur mengenai penerapan algoritma *Support Vector Machine* dan *Extreme Gradient Boosting* dalam menangani data tidak seimbang serta mengidentifikasi kecenderungan *self-diagnose*, suatu topik yang masih relatif jarang dieksplorasi dengan pendekatan *machine learning* dalam studi-studi sebelumnya.

Selain memiliki manfaat teoritis, penelitian ini juga memiliki manfaat praktis, yaitu:

1. Memberikan gambaran berbasis data mengenai perilaku pengguna TikTok terhadap konten kesehatan mental yang dapat dimanfaatkan untuk keperluan edukasi, riset, maupun intervensi sosial.
2. Meningkatkan kesadaran masyarakat terhadap kecenderungan *self-diagnose* yang teridentifikasi melalui komentar pengguna TikTok, agar lebih kritis dan bijak dalam menerima serta menyikapi informasi kesehatan mental.

## **1.5 Sistematika Penulisan**

### **BAB I PENDAHULUAN**

Pada bagian ini memberikan gambaran umum tentang latar belakang penelitian, rumusan masalah, kendala, tujuan, manfaat, serta menjelaskan metodologi penulisan yang digunakan dalam penelitian ini.

### **BAB II LANDASAN TEORI**

Pada bagian ini terdapat berbagai teori yang mendasari topik dari penelitian yang dilakukan.

### **BAB III METODELOGI PENELITIAN**

Pada bagian ini membahas mengenai berbagai metode yang diterapkan pada penelitian ini

### **BAB IV ANALISIS DAN HASIL PENELITIAN**

Pada bagian ini membahas mengenai tahapan penelitian berdasarkan metodologi CRISP-ML, yang dimulai dari tahap *business and data understanding* hingga implementasi akhir pada tahap *deployment*. Selain itu, pada bagian ini juga menyajikan evaluasi hasil akurasi dan performa model yang diterapkan yaitu algoritma SVM dan XGBoost.

### **BAB V SIMPULAN DAN SARAN**

Pada bagian ini, terdapat rangkuman hasil dari penelitian, dan disertakan rekomendasi untuk arah penelitian di masa depan.