

BAB 1

PENDAHULUAN

1.1 Latar Belakang Masalah

Kesehatan mental remaja merupakan salah satu isu global yang semakin mendesak di era digital saat ini. Media sosial telah menjadi bagian yang tak terpisahkan dari kehidupan sehari-hari remaja. Namun penggunaannya yang intensif sering dikaitkan dengan peningkatan gejala depresi dan kecemasan [2], [3]. Jumlah penderita depresi di dunia meningkat sebesar 18,4% antara tahun 2005 hingga 2015, dengan lebih dari 332 juta orang (4,4% populasi global) mengalami depresi pada tahun 2015 [1]. Remaja yang aktif menggunakan *platform* seperti *Twitter*, *Facebook*, dan *Instagram* cenderung menunjukkan peningkatan *distress* mental, perilaku *self-harm*, serta *suicidality*, terutama pada pola penggunaan pasif dan *problematic use* [2], [3], [6].

Pola serupa juga terlihat dari analisis data besar media sosial. *Dataset* yang dibangun dari *platform Reddit* menunjukkan bahwa pola depresi dan kecemasan cukup umum di kalangan remaja [4]. Selain itu, survei longitudinal terhadap 2.456 remaja di Inggris menemukan bahwa tipe pengguna yang sering membagikan konten memiliki risiko kesehatan mental terburuk satu tahun kemudian [5]. Berbagai penelitian terkini semakin memperkuat temuan tersebut, di antaranya melalui pendekatan *multimodal* yang berhasil meningkatkan akurasi deteksi depresi hingga 5% [12] serta analisis berbasis *explainable AI* yang menunjukkan kekuatan pola bahasa negatif sebagai indikator kuat *ideasi suicidal* [9]. Saat ini, media sosial berperan sebagai cermin yang merefleksikan kondisi kesehatan mental remaja secara *real-time* [17].

Untuk mendeteksi dampak tersebut secara dini, pendekatan komputasional berbasis *machine learning* menjadi sangat relevan dan diperlukan. Sebagian besar penelitian yang telah dilakukan menggunakan metode supervised learning, di mana *Support Vector Machine* (SVM) termasuk salah satu algoritma yang paling sering digunakan karena efektivitasnya dalam mengolah data teks media sosial [10]. Tinjauan komprehensif menunjukkan kemajuan signifikan teknik *machine learning* dan *deep learning* dalam deteksi depresi di media sosial [11]. Penelitian ini memanfaatkan dataset publik “*Sentiment Analysis for Mental Health*” dari

Kaggle yang berisi ribuan pernyataan teks berlabel sentimen untuk menganalisis pola sentimen negatif yang mencerminkan dampak media sosial terhadap kesehatan mental remaja [4], [10].

1.2 Rumusan Masalah

Berdasarkan latar belakang tersebut, rumusan masalah dalam penelitian ini adalah sebagai berikut:

1. Bagaimana pola sentimen pada postingan media sosial remaja mencerminkan dampak negatif terhadap kesehatan mental?
2. Bagaimana performa algoritma Support Vector Machine (SVM) dalam mengklasifikasikan sentimen dari data teks media sosial untuk mendeteksi gejala kesehatan mental remaja?
3. Apa faktor-faktor linguistik dan perilaku di media sosial yang paling berpengaruh dalam analisis sentimen dampak kesehatan mental remaja?

1.3 Tujuan Penelitian

Tujuan utama penelitian ini adalah untuk menganalisis pola sentimen pada postingan media sosial remaja yang mencerminkan dampak negatif terhadap kesehatan mental, mengukur performa algoritma *Support Vector Machine* (SVM) dalam mengklasifikasikan sentimen dari data teks media sosial untuk deteksi gejala kesehatan mental remaja, serta mengidentifikasi faktor-faktor linguistik dan perilaku di media sosial yang paling berpengaruh dalam analisis sentimen tersebut.

1.4 Batasan Penelitian

Ruang lingkup penelitian ini dibatasi sebagai berikut:

1. Penelitian ini hanya menggunakan dataset sekunder berupa “Sentiment Analysis for Mental Health” dari Kaggle yang berisi data teks berlabel sentimen, sehingga tidak melakukan pengumpulan data primer di lapangan.
2. Analisis dibatasi pada algoritma Support Vector Machine (SVM) tanpa membandingkan dengan algoritma machine learning lainnya.
3. Penelitian ini hanya mendeteksi pola sentimen negatif yang mencerminkan dampak media sosial terhadap kesehatan mental remaja, tidak sampai pada diagnosis klinis atau intervensi medis.

4. Data yang digunakan berbahasa Inggris dan berfokus pada konteks remaja berdasarkan literatur yang ada, sehingga hasil penelitian tidak dapat digeneralisasi ke semua kelompok usia atau bahasa lain.

1.5 Keterbaruan

Penelitian ini memiliki keterbaruan dibandingkan penelitian terdahulu, sebagaimana berikut:

1. Raihan et al. [4] membangun dataset *MentalHelp* berukuran 14 juta *instance* menggunakan Flan-T5, *DistilBERT*, dan *Mental-BERT* untuk mendeteksi berbagai gangguan mental, sementara penelitian ini fokus pada analisis sentimen remaja menggunakan SVM yang lebih ringan dan data terbatas.
2. Liu et al. [1] melakukan *systematic review* terhadap 17 studi *machine learning* untuk deteksi depresi dari media sosial, sedangkan penelitian ini menerapkan SVM secara spesifik pada pola sentimen remaja untuk deteksi dini gejala depresi dan kecemasan.
3. Khalaf et al. [2] meninjau dampak negatif media sosial pada remaja seperti peningkatan *self-harm*, *suicidality*, dan gangguan tidur melalui *systematic review*, namun tanpa model prediksi komputasional.
4. Min et al. [11] menyusun direktori *dataset* publik nasional untuk kesehatan mental remaja berusia di bawah 18 tahun dengan fokus pada survei longitudinal, penelitian ini berbeda dengan memanfaatkan data *real-time* media sosial untuk analisis sentimen menggunakan SVM.
5. Winstone et al. [5] mengidentifikasi empat tipe pengguna media sosial remaja melalui *latent class analysis*, di mana *Broadcasters* berisiko tertinggi depresi.
6. Perez et al. [10] menggunakan SVM untuk deteksi depresi berdasarkan pola aktivitas *web* dan kata-kata pada mahasiswa, dengan fokus pada *first-person pronouns* dan emosi negatif.
7. Kim et al. [7] melakukan *bibliometric study* terhadap 565 publikasi (2015-2020) tentang *machine learning* untuk kesehatan mental di media sosial.
8. Orben [6] melakukan *narrative review of reviews* tentang remaja, layar, dan media sosial dengan kesimpulan bahwa hubungan bersifat kompleks dan bukti masih terbatas, sedangkan penelitian ini memberikan pendekatan komputasional SVM yang konkret untuk menganalisis pola sentimen negatif pada data remaja secara langsung.

9. Tahir et al. [11] memberikan tinjauan komprehensif teknik *machine learning* dan *deep learning* dalam deteksi depresi di media sosial.
10. Yazdavar et al. [12] berhasil meningkatkan akurasi deteksi depresi melalui pendekatan multimodal.
11. Mobin et al. [17] menekankan bahwa media sosial dapat berfungsi sebagai cermin yang merefleksikan kondisi kesehatan mental melalui analisis linguistik.
12. Salas-Zárate et al. [14] menemukan bahwa *Twitter* merupakan platform yang paling banyak diteliti untuk deteksi tanda depresi.
13. Helmy [15] membandingkan performa model *machine learning* untuk deteksi depresi pada teks bahasa Inggris dan Arab.
14. Qasim et al. [16] mengeksplorasi penggunaan *transformer-based models* untuk mengklasifikasikan tingkat keparahan depresi.