

## 1. Pendahuluan

### Latar Belakang

Produktivitas masyarakat dan kualitas hidup individu dipengaruhi secara signifikan oleh kesehatan mental. Menurut WHO, depresi memengaruhi lebih dari 264 juta orang di seluruh dunia, angka yang terus meningkat setiap tahunnya. Sayangnya, stigma sosial masih menjadi penghalang utama untuk mencari bantuan profesional. Oleh karena itu, penelitian di bidang kesehatan mental sangat penting untuk merumuskan solusi preventif dan responsif terhadap gangguan psikologis [1].

Dengan lebih dari 450 juta pengguna aktif bulanan, X merupakan salah satu jaringan media sosial terbesar yang menyediakan data real-time dalam bentuk teks pendek dan media visual. Banyak pengguna yang secara terbuka membagikan perasaan, pengalaman pribadi, dan bahkan keinginan untuk bunuh diri melalui cuitan mereka, menjadikan X sebagai sumber data yang potensial untuk deteksi dini gangguan mental. Setelah pandemi COVID-19, tren penggunaan X untuk mengekspresikan masalah kesehatan mental telah meningkat, seperti yang terlihat dari tagar populer seperti *#MentalHealthAwareness* dan *#Depression* [2], [3].

Cuitan di X tidak hanya berisi informasi verbal, tetapi juga visual melalui gambar yang menyertainya. Analisis multimodal yang menggabungkan teks dan gambar telah terbukti lebih unggul daripada pendekatan unimodal karena memberikan konteks yang lebih lengkap. Sebagai contoh, teks yang netral dapat menunjukkan tanda-tanda depresi ketika digabungkan dengan gambar yang menyedihkan. Model seperti CNN-BiLSTM dan penyematan FastText telah terbukti meningkatkan akurasi prediksi kondisi mental dalam pendekatan ini [4].

Pengembangan kerangka kerja multimodal yang menggabungkan *Hybrid CNN-BiLSTM* dan TF-IDF adalah kontribusi utama dari penelitian ini. Kerangka kerja ini disempurnakan melalui FastText untuk perluasan fitur dalam ekstraksi fitur tekstual dan untuk prediksi sentimen pada subjek kesehatan mental, VGG-16 untuk ekstraksi fitur visual. Sementara itu, strategi fusi menengah menggabungkan karakteristik visual dan tekstual, dan output dari model teks saja dan gambar saja digabungkan pada fusi tingkat keputusan. Para penulis mengklaim bahwa belum ada penelitian sebelumnya yang meneliti analisis sentimen multimodal dengan ekstraksi TF-IDF dalam model prediksi sentimen. Oleh karena itu, tujuan dari penelitian ini adalah untuk membuat kerangka kerja seperti itu karena, menurut hasil penelitian sebelumnya, hal tersebut dapat meningkatkan akurasi.

### Topik dan Batasannya

Kesehatan mental merupakan masalah besar yang sangat memengaruhi kualitas hidup individu dan produktivitas masyarakat. Tetapi stigma sosial sering menghalangi orang untuk mencari bantuan. Sebaliknya, media sosial, terutama X, telah berkembang menjadi platform utama bagi pengguna untuk mengungkapkan opini dan emosi yang berkaitan dengan kesehatan mental. X menyediakan banyak data yang dapat digunakan untuk mengidentifikasi lebih baik psikologi pengguna dengan fitur teks pendek dan fitur multimodal seperti gambar.

### Perumusan Masalah

1. Bagaimana cara mengimplementasikan analisis sentimen multimodal menggunakan metode CNN-BiLSTM dengan ekspansi fitur FastText untuk meningkatkan akurasi klasifikasi sentimen?
2. Bagaimana performansi sistem analisis sentimen multimodal yang dibangun menggunakan model CNN-BiLSTM dan ekspansi fitur FastText dalam mengklasifikasi sentimen?

### Tujuan

Tujuan utama penelitian ini adalah untuk membuat sistem multimodal yang dapat menganalisis opini publik tentang kesehatan mental di X. Untuk mendapatkan pola semantik teks yang lebih baik, teknik yang digunakan termasuk model *Convolutional Neural Network* (CNN) dan *Bidirectional Long Short-Term Memory* (BiLSTM), yang didukung oleh embedding FastText. Penelitian ini membuat pendekatan multimodal yang memberikan pemahaman data yang lebih luas dengan menggabungkan analisis teks dan gambar.

Berikut beberapa tujuannya:

1. Mengimplementasikan analisis sentimen multimodal dengan menggunakan metode CNN-BiLSTM dan ekspansi fitur FastText untuk meningkatkan akurasi klasifikasi sentimen.
2. Mengukur performansi sistem analisis sentimen multimodal dengan menggunakan metode CNN-BiLSTM dan ekspansi fitur FastText.