

## 1. Pendahuluan

### 1.1 Latar Belakang

Pada masa perkembangan informasi yang pesat ini banyak bermunculan platform atau wadah yang menunjang interaksi manusia, media sosial sendiri merupakan platform media yang bertujuan untuk memfasilitasi manusia dalam beraktivitas maupun berkolaborasi [1]. Twitter adalah situs media sosial gratis yang memungkinkan pengguna terdaftar untuk berkomunikasi dengan orang lain secara *real-time* menggunakan 140 pernyataan karakter. Aplikasi seperti Twitter tidak hanya menjadi tempat berbagi tulisan dan konten multimedia tetapi juga menawarkan penggunaannya untuk mengekspresikan perasaan, emosi dan sentimen mereka tentang suatu masalah. Sehingga dengan hal tersebut sering ditemukan pengguna Twitter membuat tulisan yang menunjukkan bagaimana perilaku pengguna termasuk masalah mental yang dialami pengguna seperti gejala gangguan depresi, kecemasan dan stres.

Depresi merupakan gangguan mental dengan kesamaan konseptual dan teoritis dan terkadang tumpang tindih, terutama pada generasi muda [2]. Hanya sekitar separuh kasus depresi yang dapat dideteksi oleh dokter atau ahli lainnya, hal ini dikarenakan sampai saat ini, diagnosis depresi dimulai dari laporan pasien, keluarga, atau teman dekat pasien, atau juga dimulai dari hasil tes tertentu seperti kuesioner. Oleh karena itu, diperlukan penelitian lebih lanjut untuk mengetahui tingkat gangguan depresi seseorang. Beck Anxiety Inventory (BAI) dan Beck Depression Inventory (BDI) telah mengembangkan metode untuk menilai tingkat kecemasan dan depresi, tetapi mereka tidak membahas komponen stres sebagai reaksi tubuh. Untuk mengatasi hal ini, Lovibond PF dan Lovibond SH mengembangkan skala Depresi, Kecemasan, dan Stres (DASS) untuk mendefinisikan, memahami, dan mengukur besarnya tiga keadaan emosi negatif. DASS yang asli adalah kuesioner 42 item dalam bahasa Inggris yang berisi 14 pertanyaan masing-masing untuk menilai depresi, kecemasan, dan stres [3].

Beberapa penelitian telah dilakukan untuk memprediksi depresi dengan menggunakan berbagai macam algoritma. Salah satunya adalah penelitian yang dilakukan oleh Rodrigo Martinez Castano dan timnya menggunakan algoritma BERT yang merupakan salah satu representasi dari Unsupervised Learning Text untuk memprediksi depresi dari media sosial Reddit. Dari penelitian tersebut menunjukkan bahwa dari konten yang dibuat oleh pengguna Reddit dapat memprediksi gejala depresi pengguna tersebut. Dalam proses klasifikasi, telah menerapkan kinerja algoritma BERT (XLM-RoBERTa-base language model). Dikarenakan dalam penelitian tersebut menggunakan algoritma BERT yang membuat model terpisah untuk memprediksi hasil setiap pertanyaan menyebabkan terdapat kelemahan pada penelitian tersebut yang mana algoritma BERT tidak bisa mendapatkan konteks dari kedua sisi sehingga hasil prediksi tidak sesuai dugaan[4].

Representasi Unsupervised Learning Text selain algoritma BERT terdapat algoritma XLNet. Dengan kemampuan memodelkan konteks dua arah, BERT mencapai kinerja yang lebih baik daripada pendekatan pretrain berdasarkan pemodelan bahasa autoregresif. Algoritma BERT bekerja dengan cara merusak input yang menggunakan simbol buatan dan mengabaikan ketergantungan antara posisi simbol buatan dengan perbedaan pretrain-fine tune[5]. Sedangkan XLNet, metode pretrain autoregresif umum yang memungkinkan pembelajaran konteks dua arah dengan memaksimalkan kemungkinan yang diharapkan atas semua permutasi dari urutan faktorisasi. XLNet mengintegrasikan ide-ide dari Transformer-XL, model autoregressive yang canggih ke dalam pretrain sehingga mengatasi keterbatasan BERT berkat autoregresifnya[6].

### 1.2 Topik dan Batasannya

Berdasarkan latar belakang, topik utama dalam penelitian ini yaitu bagaimana memprediksi gangguan depresi berdasarkan DASS-42 di Twitter menggunakan metode klasifikasi XLNet serta bagaimana tingkat akurasi performa model klasifikasi XLNet dalam memprediksi gangguan depresi berdasarkan DASS-42 di Twitter.

Batasan yang terdapat dalam penelitian ini adalah dataset yang digunakan berasal dari cuitan pengguna Twitter yang telah mengisi kuesioner berbahasa Indonesia dengan responden sebanyak lebih dari 100 responden. Kuesioner berupa pertanyaan yang diajukan untuk kalangan umum berdasarkan pada kuesioner DASS-42 yang berisi 42 butir pertanyaan.

### 1.3 Tujuan

Tujuan penelitian ini adalah untuk membangun model yang memprediksi gangguan depresi dari cuitan pengguna Twitter berdasarkan DASS-42 sebagai penilaian skala depresi menggunakan metode klasifikasi XLNet, serta mengetahui bagaimana performa dari model sistem yang dibangun. Tujuan lainnya adalah model dengan performa terbaik digunakan untuk memprediksi gangguan depresi pada pengguna Twitter.

### 1.4 Organisasi Tulisan

Organisasi penulisan pada penelitian ini adalah pada bagian pertama yaitu pendahuluan berisi latar belakang, topik dan batasannya, tujuan, serta organisasi tulisan. Pada bagian kedua adalah studi terkait topik penelitian ini kemudian rancangan serta teori sistem yang dibangun dibahas pada bagian ketiga, pembahasan dan hasil penelitian di bagian keempat, terakhir kesimpulan penelitian pada bagian kelima.