

1. Pendahuluan

Depresi adalah gangguan kesehatan mental yang menyebabkan pasien mengalami gejala berat yang mempengaruhi aktivitas sehari-harinya seperti makan, tidur, bekerja, dan bagaimana perasaan atau pemikirannya [1]. Menurut WHO, depresi mempengaruhi 3,8% dari populasi manusia di seluruh dunia, dengan 5,0% orang dewasa dan 5,7% orang dewasa di atas 60 tahun. Sekitar 280 juta orang di seluruh dunia menderita akibat depresi. Depresi dapat menyebabkan seseorang sangat menderita dan menunjukkan kinerja yang buruk dalam aktivitas sehari-hari; bahkan dapat menyebabkan bunuh diri. Orang dengan depresi seringkali salah didiagnosis, sedangkan orang yang tidak depresi seringkali diresepkan antidepresan [2].

Dengan berkembangnya teknologi, manusia seringkali mengekspresikan dirinya melalui postingan di media sosial. Oleh karena itu, pada penelitian [3] dilakukan untuk mengumpulkan data dengan kata kunci yang terindikasi gangguan depresi pada *platform Twitter* dengan melibatkan psikiater untuk melabeli dataset yang terindikasi depresi atau tidak. Berdasarkan penelitian tersebut, kita dapat mengidentifikasi apakah seseorang terindikasi depresi atau tidak melalui media sosial, khususnya *Twitter*.

Media sosial adalah *platform online* untuk bersosialisasi antara pengguna dengan minat, latar belakang, atau aktivitas serupa yang memungkinkan pengguna berinteraksi tanpa batasan. Media sosial memungkinkan manusia untuk saling berkomunikasi di manapun mereka berada dan kapanpun mereka mau [4]. Menurut *Kepios*, per April 2022, 58,7% manusia di seluruh dunia memiliki akun media sosial [5]. *Twitter* adalah media sosial untuk menghubungkan dan berkomunikasi melalui pertukaran pesan yang cepat dan sering. Pengguna dapat memposting *tweet* yang berisi teks, foto, video, dan tautan. Selain itu, *tweet* akan ditampilkan di profil dan dapat dilihat oleh pengikut atau dapat dicari di *Twitter* [6]. *Statistica Research Department* menunjukkan bahwa pada Januari 2022, *Twitter* memiliki 342,75 juta pengguna aktif harian yang dapat dimonetisasi di seluruh dunia, dengan Indonesia berada di peringkat kelima [7].

Batasan pada penelitian ini yaitu jumlah data yang relatif sedikit dan sulit untuk didapatkan. Hanya terdapat sebanyak 157 baris data. Berbeda dengan penelitian-penelitian sejenis yang menggunakan hingga ribuan data. Kami memiliki keterbatasan data dikarenakan untuk mendapatkan data pada penelitian ini dibutuhkan responden yang bersedia mengisi kuesioner dan bersedia untuk tidak mem-privasi akun *Twitter*-nya.

Tujuan dari penelitian ini yaitu untuk mendeteksi apakah seseorang terindikasi depresi atau tidak berdasarkan postingan *tweet*-nya. Misalkan kita dapat mendeteksi apakah seseorang menunjukkan depresi melalui media sosialnya. Dalam hal ini dapat diberikan penanganan lebih lanjut, baik secara profesional maupun bantuan moral dari orang terdekat, sebelum ditangani lebih lanjut. Jadi, mempelajari sistem yang dapat mendeteksi apakah seseorang menunjukkan depresi atau tidak berdasarkan *tweet* mereka dapat membantu dalam memberikan perawatan bagi orang yang menunjukkan depresi. Selain itu juga untuk melihat pengaruh parameter *maximum features* pada TF-IDF dan parameter *maximum depth* pada *Decision Tree* terhadap hasil sistem. Untuk memenuhi tujuan ini, dilakukan pengujian skenario *split data* dan pemilihan jumlah *maximum features* pada TF-IDF untuk mencari rasio *split data* dan jumlah fitur terbaik yang akan menjadi *baseline*. Kemudian, setelah mendapatkan *baseline*, dilakukan *hyperparameter tuning* pada *Decision Tree* untuk mendapatkan hasil terbaik. Terakhir, dilakukan perbandingan model yang telah dibangun dengan algoritma-algoritma *Decision Tree* lainnya, yaitu *AdaBoost*, *Gradient Boosted Decision Tree*, dan *Random Forest*.