

1. Pendahuluan

Jejaring sosial memudahkan masyarakat agar tetap saling terhubung melalui pertukaran informasi secara daring. Twitter menjadi salah satu jejaring sosial yang populer di Indonesia. Jumlah pengguna Twitter di Indonesia menduduki peringkat 5 terbanyak di dunia yaitu sekitar 19.5 juta orang [1]. Pengguna dapat mengirim teks dan multimedia yang dikenal dengan sebutan *tweet*. Layanan mikroblog ini juga memiliki fitur *trending topic* yang berlangsung secara *real-time* agar pengguna bisa mengetahui tren terkini. Namun tidak semua *tweet* yang tersebar di Twitter mengandung fakta. Banyak informasi belum diketahui kebenarannya dibagikan ke sesama pengguna yang dikenal sebagai rumor [2].

Dalam keadaan darurat seperti bencana alam, Twitter banyak digunakan sebagai media komunikasi dan informasi karena penyebaran yang berlangsung sangat cepat. Di tahun 2011, gempa disertai Tsunami di Jepang mematikan jaringan telepon di beberapa daerah namun akses Internet tetap berjalan secara normal [3]. Jejaring sosial yang salah satunya adalah Twitter dimanfaatkan untuk berbagi informasi terkini terkait bencana alam karena layanan komunikasi yang dibatasi [3]. Kehadiran rumor saat terjadi situasi darurat seperti bencana alam atau krisis ekonomi dapat sangat membahayakan stabilitas sosial [4]. Pertukaran informasi yang fleksibel di Twitter memberikan lahan subur bagi penyebar rumor untuk menyebarkan desas-desus yang dapat menimbulkan kekacauan [5]. Oknum ini juga tidak segan menyebarkan ketakutan atau kebencian [6]. Sehingga pada penelitian ini dibangun sebuah deteksi rumor Bahasa Indonesia menggunakan algoritma J48 dan metode TF-IDF.

J48 adalah implementasi dari algoritma C4.5 yang membangun sebuah *decision tree* juga sebagai pengembangan dari algoritma ID3 [7]. Algoritma ini juga mampu memproses data diskrit dan kontinu serta menangani jika terdapat atribut yang mengandung *missing value* [7]. Metode pembobotan TF-IDF dilibatkan pada penelitian ini untuk mengetahui pengaruh kata yang terkandung pada *tweet* terhadap proses klasifikasi selain penggunaan fitur dasar seperti *mention*, *number of followings*, dan *number of followers*.

Penelitian ini menawarkan fitur baru yaitu *number of digits in display name* (angka pada *display name*), *number of emoticons in display name* (*emoticon* pada *display name*), dan *number of digits in username* (angka pada *username*). Tambahan fitur ini berguna untuk mengetahui ciri-ciri pengguna yang menyebarkan rumor secara lebih mendetil. Adapun tujuan dari penelitian ini adalah untuk mengimplementasikan metode yang telah dipaparkan di atas serta mengetahui kinerja sistem dalam mendeteksi kehadiran rumor berbahasa Indonesia.