

ABSTRAK

Di tengah semakin pesatnya perkembangan teknologi digital hingga hari ini, berpendapat ke ruang terbuka publik bukanlah hal yang sulit dilakukan. Media sosial yang menawarkan kebebasan lebih untuk aktivitas secara digital kemudian menjadi wadah yang cukup diminati orang-orang sebagai tempat untuk mengemukakan pendapatnya. Salah satu topik yang cukup sering diperbincangkan di ruang publik adalah gangguan mental. Di era keterbukaan informasi seperti dewasa ini, gangguan mental sudah tidak lagi menjadi topik bahasan tabu di tengah masyarakat. Ini kemudian menimbulkan efek dikotomis. Di satu sisi media sosial dapat membantu meningkatkan kesadaran masyarakat atas gangguan mental. Akan tetapi di sisi lain, karena informasi yang salah lalu media sosial malah menjadi sarana menumbuh-suburkan persepsi buruk dan miskonsepsi sebagian masyarakat terhadap gangguan mental itu sendiri. Analisis sentimen merupakan suatu teknik dalam *natural language processing* yang dapat digunakan untuk mengungkap tendensi emosi di balik sebuah pendapat seseorang secara tekstual yang membahas gangguan mental; apakah itu ke arah emosi positif, netral, atau malah sebaliknya. Penelitian ini kemudian dilakukan untuk melihat unjuk kerja teknik tersebut dalam menganalisis sentimen masyarakat terhadap dua tipe gangguan mental yang umum dikenali masyarakat Indonesia, yakni depresi dan bipolar, berdasarkan pendapatnya yang dikemukakannya di media sosial Twitter. Dengan menggunakan sampel data sejumlah 3409 twit, model sistem analisis sentimen untuk kasus ini dibangun berfondasikan sebuah metode klasifikasi *support vector machine*. Melalui dua tahapan pelatihan, *support vector machine* dimodelkan sedemikian rupa untuk menemukan jenis kernel beserta parameter pengontrolnya sehingga sistem analisis sentimen mencapai kinerja terbaiknya, diukur dari tingkat akurasi klasifikasi yang didapatkan. Berdasarkan pengukuran performa, ditemukan bahwa model terbaik yaitu kernel *radial basis function* dengan nilai $C=10$ dan $\gamma=0,1$ untuk *support vector machine* yang mana datanya diekstraksi menggunakan metode *term frequency-inverse document frequency*. Pengujian atas model terbaik ini memperoleh akurasi sebesar 74,1 persen dan *f1-score* 72,5 persen.

Kata kunci: analisis sentimen, depresi, bipolar, *support vector machine*, twitter

ABSTRACT

In the midst of the rapid development of digital technology to this day, speaking out in public is not a difficult thing to do. Social media, which offers more freedom for digital activities, has become a platform that people are quite interested in as a place to express their opinions. One topic that is often discussed in the public sphere is mental disorders. In this era of information openness, mental disorders are no longer taboo topics in society. This then creates a dichotomous effect. On the one hand, social media can help raise public awareness of mental disorders. But on the other hand, due to misinformation, social media has become a means of fostering bad perceptions and misconceptions about mental disorders. Sentiment analysis is a technique in natural language processing that can be used to uncover the emotional tendencies behind a textual person's opinion on mental disorders, whether they are positive, neutral, or otherwise. This study was then conducted to see the performance of the technique in analyzing people's sentiments towards two commonly recognized types of mental disorders in Indonesia, namely depression and bipolar, based on their opinions expressed on Twitter. Using a data sample of 3409 tweets, the sentiment analysis system model for this case was built based on a support vector machine classification method. Through two training stages, the support vector machine is modeled in such a way as to find the kernel type and its controlling parameters so that the sentiment analysis system achieves its best performance, measured by the classification accuracy rate obtained. Based on the performance measurement, it is found that the best model is the radial basis function kernel with $C=10$ and $\gamma=0,1$ for the support vector machine, where the data is extracted using the term frequency-inverse document frequency method. Tests on this best model obtained an accuracy of 74,1 percent and an f1-score of 72,5 percent.

Keywords: sentiment analysis, depression, bipolar, support vector machine, twitter