

Abstrak

Twitter adalah salah satu media social, yang mempunyai pengguna sangat banyak, dengan banyaknya pengguna, *tweet* yang dihasilkan Twitter setiap harinya sangatlah banyak, dan dapat dikelompokkan menjadi *trending topic* yang menggambarkan berita yang sedang hangat dibicarakan pada saat ini. Namun terdapat masalah karena masih banyak *tweet* yang menggunakan *hashtag* tentang kejadian yang hangat dengan isi *tweet* yang berbeda dari *hashtag* yang dimaksudkan, sehingga membuat pengguna harus membaca satu persatu isi dari setiap *tweet*. Penelitian ini bertujuan untuk mengatasi masalah dengan membagi berbagai kategori pada *trending topic* yang ada pada Twitter. Penelitian ini dilakukan dengan menggunakan metode klasifikasi *Support Vector Machine* (SVM) dengan menambahkan metode pembobotan *Term Frequency-Inverse Document Frequency* (TF-IDF). Dari hasil percobaan didapatkan akurasi tertinggi sebesar 81.4% dengan menggunakan skenario *dataset* dan *ratio* 90:10, dan didapatkan akurasi tertinggi sebesar 81.47% pada *ratio* 90:10, dengan skenario n-gram pada bagian unigram + bigram + trigram.