

ABSTRAK

Konsumsi informasi berita yang semakin tinggi telah menjadi kebutuhan vital bagi masyarakat. Hal ini berdampak pada cepatnya perkembangan lalu lintas informasi di tengah masyarakat. Salah satu sumber informasi yang berperan penting dalam penyebaran berita adalah *tweet* yang dikirimkan melalui media sosial Twitter. Dengan banyaknya informasi yang beredar melalui *tweet* tentunya akan menyulitkan pembaca untuk mengetahui informasi berita spesifik yang tersebar. Pada penelitian ini penulis terfokus pada pengklasifikasian informasi berita ke dalam kategori berita dengan sumber data *tweet*. Teknik yang digunakan adalah algoritma *Naïve Bayes* dengan dua tahap klasifikasi. Tahap pertama digunakan untuk memisahkan antara *tweet* berita dan *tweet* bukan berita. Sedangkan tahap kedua digunakan untuk memisahkan *tweet* berita tersebut ke dalam kategori berita yang telah disiapkan. Hasil klasifikasi pada tahap pertama menghasilkan akurasi sebesar 77,35%, sedangkan hasil klasifikasi pada tahap kedua menghasilkan akurasi sebesar 61,21%. *Tweet* yang telah terkelompok di masing-masing kategori ditampilkan pada *dashboard user* untuk memudahkan pengguna dalam mengakses informasi berita. Dari penelitian ini menunjukkan bahwa algoritma *Naïve Bayes* layak digunakan untuk pengklasifikasian informasi berita dengan sumber data *tweet* berbahasa Indonesia.

Kata kunci: Klasifikasi, *Naïve Bayes*, Twitter, Berita

ABSTRACT

As the increase of news consumption has become a vital necessity of society. This condition affects on the rapid growth of information traffic in the community. One of information source that plays an important role in dissemination of news is tweet that sent via Twitter. With the amount of information circulating through Twitter will certainly make it difficult for readers to know the specific information that scattered. In this research author focused on the classification of news information into several categories with tweets as the data source. Author used Naïve Bayes algorithm with two-stage classifications. The first stage is used to separate tweets as news and non-news, while the second one is used to separate news tweets into categories that have been prepared. The result of the first stage shows the classification accuracy of 77,35%, while the second stage shows the classification accuracy of 61,21%. Tweets that have been classified in each category will displayed on user dashboard and allow user to access the news information comfortably. This research pointed that Naïve Bayes algorithm can be used for the classification of news information with tweets in Bahasa as the data source.

Keywords: Classification, Naïve Bayes, Twitter, News