

ABSTRAK

Twitter merupakan salah satu sosial media yang paling berpengaruh di dunia dengan jumlah penggunanya sebanyak 15,7 juta pengguna di Indonesia dan menjadi negara peringkat ke 6 sebagai negara dengan pengguna *Twitter* terbanyak. Selain digunakan sebagai media sosial, *Twitter* juga digunakan sebagai media untuk membaca ataupun mengirim suatu berita. Informasi berita yang terus meningkat menyebabkan pengguna akan mengalami kesulitan dalam mencari informasi-informasi berita tertentu. Salah satu solusi yang dapat diterapkan untuk dapat mengatasi masalah tersebut adalah melalui *clustering* terhadap informasi berita yang ada di *Twitter*. Pada penelitian ini, peneliti menggunakan jenis penelitian secara kuantitatif dengan metode *K-Means Clustering* yaitu salah satu metode *clustering* yang digunakan dalam mengelompokkan data. Data yang digunakan pada penelitian ini adalah dataset yang diambil dari *UCI Machine Learning Repository* yaitu *Health News in Twitter Data Set* sebanyak 14.970 data *tweet*. Hasil penelitian menunjukkan penentuan klaster terbaik dengan metode *Elbow* pada dataset tersebut menghasilkan bukti empiris bahwa klaster terbaik adalah $K=3$. Hasil pengelompokkan berita kesehatan pada sosial media *Twitter* dengan metode *K-Means Clustering* dengan $K=3$ menghasilkan jumlah *cluster* yaitu C1 sebanyak 4.991 data *tweets*, C2 sebanyak 4.482 data *tweets*, dan C3 sebanyak 5.497 *tweets*.

Kata Kunci: Pengelompokkan Berita Kesehatan, Sosial Media *Twitter*, Metode *Elbow*, Metode *K-Means Clustering*

ABSTRACT

Twitter is one of the most influential social media in the world with 15.7 million users in Indonesia and is ranked 6th as the country with the most Twitter users. Besides being used as social media, Twitter is also used as a medium to read or send news. News information that continues to increase causes users to have difficulty in finding certain news information. One solution that can be applied to overcome this problem is through clustering of news information on Twitter. In this study, the researcher used quantitative research with the K-Means Clustering method, which is one of the clustering methods used in grouping data. The data used in this study is a dataset taken from the UCI Machine Learning Repository, namely Health News in Twitter Data Set as many as 14,970 tweet data. The results showed that the determination of the best cluster using the Elbow method on the dataset resulted in empirical evidence that the best cluster was $K=3$. The results of grouping health news on Twitter social media using the K-Means Clustering method with $K=3$ resulted in the number of clusters, namely C1 as many as 4,991 data tweets, C2 as many as 4,482 tweets data, and C3 as many as 5,497 tweets.

Keywords: Health News Grouping, Twitter Social Media, Elbow Method, K-Means Clustering Method