

ABSTRAK

Pada era saat ini, dengan maraknya penggunaan media sosial *X* berbasis teks ini seringkali masyarakat membuat *tweet* berupa opini terhadap produk atau barang yang digunakannya sehingga memberikan pengaruh terhadap pengguna lain bagi yang membaca *tweet* tersebut. Karena hal tersebut, pada penelitian ini akan membahas mengenai sentimen analisis yang bertujuan memberikan suatu keputusan dari berbagai macam opini masyarakat di *X* terhadap produk *sunscreen* yang marak digunakan saat ini yang nantinya bisa menjadi peluang untuk produsen dalam membuat inovasi produk yang dibutuhkan oleh masyarakat saat ini. Dari data sentimen analisis ini data diolah menggunakan algoritma *K-Medoids* dan *K-Means* yang dianalisis yang bertujuan untuk mengetahui tingkat keakurasian dari masing-masing metode algoritma. Adapun untuk tahapan untuk penelitian ini dimulai dari pengumpulan data dari platform sosial media *X* sebanyak sekitar 10.000 *tweet* yang berkaitan dengan kata kunci *sunscreen*, selanjutnya dilakukan *pre-processing* untuk mencari kata-kata yang sering muncul pada *tweet*, lalu data yang sudah di *pre-processing* diproses menggunakan metode sentimen analisis dengan menghasilkan 3 kelompok yaitu *tweet* dengan kelompok positif, netral, dan negatif. Dari hasil kelompok tersebut, kelompok *tweet* dengan hasil positif dihitung jumlah frekuensi kemunculannya sehingga dihasilkan jenis *sunscreen* dengan keunggulan tertentu yang dapat dijadikan inovasi produk selanjutnya. Dari hasil data sentimen analisis sebelumnya, data diproses menggunakan metode algoritma *clustering K-Medoids* dan *K-Means* untuk mengetahui hasil akurasi dari masing-masing algoritma dengan diperoleh hasil algoritma *K-Means* dengan indeks hasil uji tertinggi yaitu 0,609 pada *cluster* 5 sedangkan pada algoritma *K-Medoids* dengan indeks hasil uji tertinggi sebesar 0,689 berada pada *cluster* 3. Sehingga untuk algoritma yang baik dalam pengolahan data sentimen ini adalah algoritma *K-Medoids*.

Kata kunci : *sentimen analisis, tweet, k-medoids, k-means*

ABSTRACT

In today's era, with the widespread use of X-based text social media, people often tweet their opinions about the products or items they use to influence other users who read those tweets. Because of this, this research will discuss sentiment analysis aimed at providing a decision from various opinions of the X community regarding sunscreen products that are currently widely used, which could later become an opportunity for manufacturers to innovate products needed by the community today. From this sentiment analysis data, the data is processed using the K-Medoids and K-Means algorithms, which are analyzed to determine the accuracy level of each algorithm. The steps for this research begin with data collection from the X social media platform, totaling about 10,000 tweets related to the keyword "sunscreen". Next, pre-processing is carried out to find frequently occurring words in tweets. Then, the pre-processed data is processed using sentiment analysis methods, resulting in 3 groups: tweets with positive, neutral, and negative sentiments. From these groups, tweets with positive results are counted for frequency of occurrence, thus producing types of sunscreen with specific advantages that can be used for further product innovation. From the previous sentiment analysis data, the data is processed using the K-Medoids and K-Means clustering algorithms to determine the accuracy of each algorithm. The K-Means algorithm yielded the highest test result index of 0.609 in cluster 5, while the K-Medoids algorithm yielded the highest test result index of 0.689 in cluster 3. Thus, the algorithm that performs well in processing this sentiment data is the K-Medoids algorithm.

Keywords: sentiment analysis, tweet, k-medoids, k-means