

ABSTRAK

Seiring berkembangnya teknologi informasi dan komunikasi, masyarakat dapat dengan mudah dalam menyampaikan tanggapan dan opini mereka terhadap jasa pelayanan transportasi *online* melalui jejaring sosial khususnya Instagram. Berdasarkan data *We Are Social 2020* dikatakan ada 79% dari populasi masyarakat Indonesia aktif menggunakan sosial media Instagram, dengan banyaknya pengguna aktif tersebut pastinya menghasilkan banyak data seperti data penilaian masyarakat terhadap jasa layanan transportasi *online* yang banyak disampaikan melalui kolom komentar pada Instagram, dengan beragam sentimen positif dan negatif. Sentimen masyarakat ini tentunya dapat menjadi informasi bagi pelaksana program untuk mengevaluasi kebijakan dan layanannya.

Pada penelitian ini dilakukan pengelompokan sentimen pada masing – masing data sentimen positif negatif, dan netral menggunakan algoritma DBSCAN (*Density-Based Spatial Clustering of Applications With Noise*), Tujuan utama dari *clustering* ini untuk mengelompokkan opini masyarakat yang berdasar pada kesamaan karakteristik atau makna dalam penulisan di antara opini-opini tersebut untuk menentukan positif, negatif, dan netral berdasarkan komentar pada media sosial instagram. Dengan melakukan tahapan *preprocessing* seperti *tokenize*, *stopword*, dan *stemming*, kemudian dilakukan pembobotan kata dengan menggunakan TF- IDF untuk dapat melakukan pengelompokan opini. Dari hasil *Clustering* didapatkan hasil dari pengujian *dataset* positif, negatif, dan netral masing masing diuji coba dengan *range* nilai min sampel dari 10-50 dan nilai *epsilon* dari 0,1-1,0 dengan menghasilkan nilai *silhouette coefficient*nya berbeda beda. Namun untuk nilai terbaik dari ketiga *dataset* didapatkan pada nilai inputan $eps=1,0$ dan inputan nilai min sampel = 10, untuk hasil *dataset* positif nilai *silhouette coefficient*-nya adalah 0.7800973549904059, untuk hasil *dataset* netral nilai *silhouette coefficient*-nya adalah 0.7526159947007542, untuk hasil *dataset* negatif nilai *silhouette coefficient*-nya adalah 0.8047251594403672. Kemudian visualisasi data hasil *clustering* topik tersebut akan ditunjukkan pada perangkat lunak berbasis web yang juga dirancang pada penelitian Tugas Akhir ini.

Kata Kunci: *Clustering, TF-IDF, Preprocessing, silhouette coefficient.*