

1. Pendahuluan

Salah satu media sosial terbesar di dunia adalah Twitter. Twitter biasanya digunakan masyarakat untuk menyampaikan pesan dan pendapat berupa opini tentang bahasan tertentu dan bahasan yang menjadi *trending topic* terkini [1]. Salah satu bahasan yang sempat ramai diperbincangkan adalah kabar bahwa Pak Erick Thohir mengangkat sekaligus mengumumkan Pak Ahok sebagai Komisaris Utama salah satu perusahaan BUMN yaitu PT Pertamina (Persero). Kabar tersebut merupakan salah satu bahasan yang pernah menjadi topik hangat masyarakat terutama di media sosial dan sempat menjadi *trending topic*.

Dalam media sosial Twitter, opini-opini yang dituangkan dalam bentuk *tweets* yang dapat berupa opini positif maupun opini negatif. Namun, penilaian sentimen secara manual terbilang sulit karena jumlah data yang besar sehingga perlu diterapkannya otomatisasi yang dapat memudahkan hal tersebut dengan menggunakan metode *Feature Selection Query Expansion Ranking* dan metode klasifikasi *K-Nearest Neighbor*. *Query Expansion Ranking* adalah sebuah metode seleksi fitur yang berguna untuk mengurangi kompleksitas komputasi tanpa mengurangi kualitas dari analisis sentimen (Fauzi et al., 2017). Kemudian hasil dari *Query Expansion Ranking* akan diklasifikasikan dengan *K-Nearest Neighbor* untuk mendapatkan hasil akhir. *K-Nearest Neighbor* (KNN) adalah sebuah metode untuk mengklasifikasikan objek berdasarkan data pembelajaran yang jaraknya paling dekat dengan objek tersebut. Berdasarkan k buah data latih yang jaraknya paling dekat dengan objek tersebut. Syarat nilai k adalah tidak lebih besar dari jumlah data latih, dan nilai k harus ganjil serta lebih dari satu. Dekat atau jauhnya jarak data latih, yang paling dekat dengan objek yang akan diklasifikasi dapat dihitung berdasarkan jarak *Euclidean*.

Pada penelitian tugas akhir ini penulis menggunakan penggabungan metode klasifikasi KNN dengan *feature selection Query Expansion Ranking* (QER), yang akan melalui proses *preprocessing*, pembobotan, dan QER lalu setelah itu diklasifikasikan dengan KNN.

Pendekatan yang digunakan pada penelitian ini berasal dari teks *tweet* dari sosial media Twitter dengan mengambil *trending topic* #AhokDirutMafiaCemberut. Batasan pada tugas akhir ini adalah *dataset* yang digunakan adalah data *tweets* yang berasal dari opini pengguna Twitter. Dan *dataset* yang digunakan adalah sebanyak 10000 *tweets*.

Tujuan yang ingin dicapai dari pengerjaan tugas akhir ini adalah mengetahui pengaruh penggunaan *Feature Selection* pada metode *K-Nearest Neighbor* dan mengetahui hasil dari performa algoritma *K-Nearest Neighbor* dengan menggunakan *Feature Selection Query Expansion Ranking* pada data Twitter. Pada penulisan tugas akhir ini, terdapat bab 1 yang menjelaskan latar belakang, tujuan, rumusan masalah serta batasan masalah yang diambil. Kemudian bab 2, menjelaskan landasan teori dan metode yang terpakai. Selanjutnya bab 3, menjelaskan isi dari rancangan sistem yang akan dijalankan pada penelitian. Lalu bab 4, menjelaskan hasil uji beserta analisisnya. Dan terakhir bab 5, menjelaskan kesimpulan yang didapat pada penelitian.