

1. Pendahuluan

1.1. Latar Belakang

Penggunaan media sosial dewasa ini menjadi marak. Pasalnya, media sosial telah menjadi sarana utama manusia dalam mengakses segala bentuk komunikasi[1]. Media sosial hadir sebagai medium baru bagi kita untuk mengemukakan pendapat serta berbagi sudut pandang dan pemikiran baru mengenai berbagai isu dan topik hangat. Salah satu media sosial yang populer dalam berkomunikasi menggunakan teks adalah Twitter. Twitter merupakan media sosial bertipe *micro-blogging* dengan maksimum 280 karakter. Dengan ini, twitter dapat menciptakan ruang publik di mana orang dapat berkumpul, berdiskusi, dan menuntut akuntabilitas yang dapat mengarah pada perubahan sosial yang positif[2]. Salah satu isu atau topik hangat yang sedang diperbincangkan adalah pemindahan Ibu Kota Negara Indonesia di Nusantara.

Wacana pemindahan ibu kota seperti menjadi agenda rutin pada tiap periode kepresidenan. Namun, dibawah pemerintahan Jokowi – Ma’ruf akhirnya direalisasikan dengan ditetapkannya UU No.3 Tahun 2022 mengenai Ibu Kota Negara[3]. Hal ini kemudian menjadi polemik di tengah masyarakat Indonesia. Tidak sedikit masyarakat mengemukakan pendapat dan pandangannya mengenai pemindahan Ibu Kota Negara ini pada media sosial Twitter. Kemudahan akses Twitter dalam penyampaian pendapat bisa memberi peluang penilaian dan evaluasi terhadap ibu kota negara Indonesia yang baru. Untuk menghasilkan informasi dari data opini yang ada, dilakukan analisis sentimen untuk memisahkan opini ke dalam kelas sentimen positif atau negatif dan menentukan faktor mana yang sering dibahas dalam opini tersebut.[4].

Analisis Sentimen atau biasa disebut *opinion mining* merupakan sebuah bidang studi yang menganalisis opini, penilaian, dan emosi terhadap suatu hal diantaranya jasa, produk, dan bahkan isu-isu atau topik[5]. Hasil analisis tersebut dapat menjadi pendukung utama penentuan sebuah topik bersifat positif atau negatif. Opini sendiri telah menjadi hal penting dalam pengambilan sebuah keputusan[6]. Sentimen analisis dapat dilakukan dengan berbagai metode diantaranya adalah *Naïve Bayes*, *K-Nearest Neighbor* (KNN), *Support Vector Machine* (SVM), dan lain sebagainya. Salah satu penelitian oleh P. Arsi *et al.*[7] melakukan analisis sentimen menggunakan metode SVM dan pembobotan fitur *tf.idf* menghasilkan akurasi sebesar 96,68%. Adapun tahapan – tahapan dasar dalam melakukan analisis sentimen diantaranya adalah pengambilan data, pelabelan data, *preprocessing*, penambahan pembobotan fitur, klasifikasi, lalu melakukan analisis.

Pembobotan fitur bertujuan untuk memberikan bobot untuk masing-masing kata pada suatu dokumen atau kalimat. Relevansi istilah apa pun dalam model ini berubah tergantung pada teksnya. Akibatnya, menambahkan bobot (nilai) ke setiap frasa sangat penting untuk merepresentasikan dokumen teks[8]. Beberapa diantara pembobotan fitur atau *-term weighting* yang sering digunakan adalah *tf*, *tf.idf*, *tf.rf*, dan lain sebagainya. Pun tidak sedikit pihak yang berupaya untuk memodifikasi pembobotan fitur yang telah ada untuk menambah performansi serta akurasi model. Seperti penelitian oleh Deqing Wang *et al.*[9] yang memperkenalkan *inverse category frequency* ke dalam skema pembobotan fitur dan mengusulkan skema pembobotan *tf.icf* dan *icf-based* atau *tf.bin-icf* yang menggabungkan *inverse category frequency(icf)* dan *relevance frequency(rf)*. Hasilnya mengungkapkan bahwa skema pembobotan *tf.bin-icf* mencapai kinerja yang lebih baik dibandingkan tujuh skema pembobotan lainnya.

Berangkat dari latar belakang tersebut, penelitian ini akan melakukan perbandingan *term weighting* atau pembobotan fitur diantaranya *tf.idf*, *tf.rf*, serta *tf.bin-icf*. Topik Ibu Kota Negara dipilih dalam analisis sentimen ini karena topik tersebut masih hangat dan ramai diperbincangkan masyarakat dengan berbagai opini mereka. Salah satu penelitian yang dilakukan oleh Zhiejie Liu[10] membuktikan bahwa metode klasifikasi SVM lebih unggul dibandingkan metode lain terlebih pada bidang atau topik politik dengan akurasi 97,31% dan F1-score mencapai 93,78%. Penelitian ini mengimplementasikan klasifikasi SVM dengan linear kernel dan *K-fold cross validation*, bersama dengan metode evaluasi *Confusion Matrix* untuk mengukur akurasi. Kernel Linear SVM sering digunakan ketika data dapat dipisahkan secara linear[11] yang sangat cocok untuk *binary classification task*.

1.2. Topik dan Batasannya

Penelitian ini membahas perbandingan metode pembobotan fitur dengan metode klasifikasi *Support Vector Machine* pada cuitan *-tweet* masyarakat Indonesia yang membahas topik Ibu Kota Negara Baru. Pembobotan fitur yang digunakan antara lain adalah *tf.idf*, *tf.rf*, dan *tf.bin-icf*. Adapun data yang digunakan merupakan *tweet* berbahasa Indonesia dengan kata kunci “ibu kota negara” dan “ikn” yang diambil dari media sosial Twitter.

1.3. Tujuan

Tujuan dilakukannya penelitian ini adalah untuk mengetahui performansi dari kombinasi metode klasifikasi *Support Vector Machine* dengan pembobotan fitur diantara *tf.idf*, *tf.rf*, serta skema pembobotan baru yaitu *tf.bin-icf* dalam analisis sentimen dengan topik Ibu Kota Negara Baru.

1.4. Organisasi Tulisan

Pada laporan penelitian ini, akan dibahas studi literatur yang terkait dan berhubungan serta studi yang menjadi acuan penelitian pada Bab 2. Kemudian sistem yang dibangun pada penelitian ini dapat dilihat pada Bab 3. Hasil dari penelitian yang telah dilakukan tertera pada Bab 4. Serta kesimpulan dari penelitian dapat dilihat pada Bab 5.