

1. Pendahuluan

Kebijakan publik adalah kebijakan atau serangkaian tindakan yang dialokasikan secara sah oleh pemerintah yang mempengaruhi setiap orang di suatu negara. Pemerintah yang berwenang dalam pembuatan kebijakan publik, diharapkan dapat membuat suatu kebijakan yang menyelesaikan permasalahan dan membawa kebaikan bagi berbagai pihak terutama masyarakat. Namun, sebagai sebuah fakta, beberapa kebijakan publik yang dibuat tidak hanya memberikan efek atau sisi positif, melainkan juga sisi negatif. Akibat hal tersebut, muncul berbagai respon pro maupun kontra yang diutarakan oleh pengguna jejaring media sosial seperti di salah satu *web microblog* yang terkenal, yaitu Twitter. *Microblog* adalah web yang memungkinkan penggunaannya untuk bertukar elemen kecil dari konten seperti kalimat pendek, gambar individu, atau tautan video [1]. Twitter menyebut elemen ini dengan istilah *tweet*. Dalam mengunggah sebuah *tweet*, Twitter memberikan batasan jumlah panjang kalimat yang dapat ditulis sebanyak 280 karakter, hal ini dapat meningkatkan penggunaan variasi kata dan *emoticon* sehingga memungkinkan terjadinya ketidakcocokan kosakata.

Penelitian mengenai analisis sentimen Twitter saat ini menjadi topik yang populer untuk penelitian [2]. Analisis sentimen adalah sebuah proses otomatis untuk menentukan apakah segmen teks berisi konten objektif atau opini, dan selanjutnya menentukan polaritas sentimennya [3]. Analisis sentimen merupakan instrumen yang sangat bermanfaat untuk mengumpulkan dan mengklasifikasikan opini publik dengan menganalisis data sosial yang besar. Pengklasifikasian sentimen Twitter bertujuan untuk mengklasifikasikan sentimen tersebut apakah polaritas *tweet* termasuk positif, negatif, atau netral [4]. Pada penelitian [5], penggunaan Word2Vec dengan *Support Vector Machine* (SVM) untuk analisis sentimen ulasan produk berbahasa Indonesia mencapai nilai akurasi sebesar 70%. Penelitian oleh Kurniawan [6], membandingkan *Continuous Bag-of-Words* (CBOW) dan *Skip-Gram* dengan berbagai dimensi klasifikasi SVM yang menghasilkan nilai tertinggi *precision* sebesar 64,4%, *recall* 58%, dan *F1-score* 61,1% pada metode *Skip-Gram* dengan 100 dimensi. Pada penelitian yang dilakukan Erwin [7], menghasilkan akurasi sebesar 98,81% pada *Logistic Regression* dengan penerapan metode *hybrid* yang menggabungkan *basic feature* dan *feature expansion* untuk analisis sentimen di Twitter. Pada penelitian [8] dengan metode Doc2Vec dan SVM mengenai klasifikasi sentimen pada dokumen ulasan film berbahasa Inggris, menghasilkan nilai *F1-Measure* terbaik sebesar 51,38% pada dimensi 500 dan nilai akurasi tertinggi sebesar 54,19% didapatkan dengan menggunakan kernel *polynomial*. Lalu, pada penelitian [9], digunakan sebuah pengembangan metode bernama *feature expansion* menggunakan *word embedding* Word2Vec untuk klasifikasi topik sebuah *tweet* menunjukkan, bahwa metode yang diusulkan mempengaruhi dan meningkatkan nilai akurasi saat menggunakan klasifikasi *Logistic Regression*.

Berdasarkan berbagai penelitian yang telah dijelaskan, tujuan pada penelitian ini adalah mengembangkan sebuah sistem untuk analisis sentimen Twitter dengan menerapkan *feature expansion* menggunakan Word2Vec. Alasan pemilihannya, diharapkan metode tersebut dapat mengatasi permasalahan ketidakcocokan kosakata dalam kalimat terbatas seperti *tweet* di Twitter. Untuk perbandingan, digunakan nilai performansi sistem dari metode pembobotan kata TF-IDF dan penggunaan dua jenis algoritma klasifikasi berbeda, yaitu *Support Vector Machine* (SVM) dan *Logistic Regression* (LR).

Batasan penelitian dalam Tugas Akhir ini, yaitu data yang digunakan adalah data sentimen Bahasa Indonesia sebanyak 16.597 bertopik kebijakan publik di Indonesia yang diambil dari *tweet* pengguna media sosial di Twitter, proses pelabelan sentimen dilakukan secara manual menjadi dua kategori, yaitu positif dan negatif, nilai matriks performansi yang digunakan adalah akurasi dan *F1-Score*, serta *word embedding* yang digunakan adalah Word2Vec.

Jurnal tugas akhir ini disusun dengan struktur sebagai berikut, pada bab 2 akan membahas studi/literatur yang mendukung atau berkaitan erat dengan penelitian ini. Bab 3 membahas teori terkait penelitian dan pemodelan sistem yang dibangun. Pada bab 4 membahas hasil, analisis, dan evaluasi model penelitian. Lalu, pada bab 5 menjelaskan kesimpulan dan saran untuk penelitian selanjutnya.