

## 1. Pendahuluan

Pada era digital saat ini, perkembangan teknologi yang semakin maju membuat masyarakat semakin dimudahkan dalam berbagai aspek aktivitas, salah satunya adalah media sosial. Media sosial merupakan tempat masyarakat mengekspresikan kegiatan secara online. Banyak platform media sosial yang menarik perhatian masyarakat, salah satunya adalah *TikTok*. *TikTok* merupakan aplikasi dimana pengguna dapat mengekspresikan melalui video berdurasi pendek dalam berbagai topik [1]. Fitur – fitur yang terdapat pada *TikTok* mendorong kreativitas penggunaannya dalam membuat video [2]. Sehingga banyak masyarakat yang tertarik untuk menggunakan aplikasi ini. Populernya *TikTok* di Indonesia tidak lepas dari jejak diblokir oleh pemerintah karena terdapat pelanggaran konten [3]. Berdasarkan data yang diperoleh dari SensorTower [4], *TikTok* merupakan aplikasi yang paling sering diunduh dan menjadikan *TikTok* sebagai aplikasi Non-Facebook (Instagram, Facebook, dan Whatsapp) yang berhasil mencapai 3 Miliar total pengunduhan dalam App Store dan Google Play.

Dengan demikian, banyak pengguna memberikan ulasan terhadap *TikTok* pada App Store dan Google Play, ulasan tersebut dapat berupa positif atau negatif. Namun, ada banyak ekspresi yang digunakan pengguna dalam memberikan ulasan [5]. Ini dapat menyebabkan ulasan menjadi ambigu atau tidak pasti dikarenakan ulasan yang diberikan mengandung makna positif dan negatif. Analisis sentimen berbasis aspek menjadi solusi dalam menganalisa sentimen yang tidak mampu mengolah sentimen ke dalam positif atau negatif [6] dan menyimpulkan faktor apa saja yang dibahas oleh pengguna [7]. Maka dari itu, studi ini akan membahas analisis terhadap review *TikTok* untuk mendapatkan nilai pasti terhadap *review* tersebut.

Penelitian yang dilakukan Aufer dkk [8], membandingkan metode Decision Tree dengan Random Forest terhadap komentar dalam media sosial yaitu Youtube. Hasil yang diperoleh cukup baik dari kedua metode tersebut, akurasi Decision Tree 89.4% dan Random Forest 88.2%. Penelitian [9] membandingkan metode SVM, Naïve Bayes, Random Forest dan Hybrid. Hasil yang diperoleh metode Hybrid menghasilkan rata – rata akurasi 84%. Penelitian [10] menggunakan metode Random Forest terhadap komentar pada Twitter dan memperoleh akurasi sebesar 75%. Penelitian Word2Vec [11,12] dengan model Skip-Gram menghasilkan akurasi lebih baik dibandingkan dengan model CBOW sebesar 85.9% dan 72%. Aplikasi *TikTok* dipilih untuk kasus analisis sentimen karena termasuk media sosial yang populer saat ini sesuai dengan data [4]. *Word2Vec* dipilih sebagai model ekstraksi fitur karena memiliki kemampuan dalam mengelompokkan kata yang bermakna sama dalam bentuk vektor sehingga dapat memprediksi kata lebih baik. Pada penelitian [11,12], penggunaan *Word2Vec* memiliki hasil akurasi yang baik. Metode klasifikasi digunakan dengan *Random Forest Classifier* karena pada penelitian [13] menghasilkan akurasi yang lebih baik dibandingkan model klasifikasi yang lain.

Oleh karena itu, motivasi pada studi ini akan membahas proses mengolah ulasan dari aplikasi *TikTok* pada *Google Play* dengan menggunakan metode pemodelan *Word2Vec* dan *Random Forest* sebagai klasifikasi. Dengan melakukan eksperimen untuk menentukan keefektifan akurasi dari penggabungan metode tersebut.

Topik dan Batasan masalah dalam penelitian ini adalah melakukan pengujian berdasarkan skenario yang telah ditetapkan untuk mengetahui performansi dari akurasi model yang dirancang. Batasan masalah pada penelitian ini adalah dilakukan *Multi-Aspect* analisis sentimen menggunakan *Word2Vec* sebagai pemodelan data dan *Random Forest Classifier* sebagai klasifikasi. Data yang digunakan adalah ulasan aplikasi *TikTok* berbahasa Indonesia dari *Google Play* yang berjumlah 10.000. Data tersebut akan dibagi ke dalam 3 aspek yaitu aspek fitur, bisnis dan konten. Setiap aspek memiliki nilai 1 untuk positif, 0 untuk netral dan -1 untuk negatif.

Tujuan dari penelitian ini adalah mengetahui performansi dan akurasi menggunakan *Word2Vec* dan *Random Forest Classifier* dalam membandingkan model *Skip-gram* dan CBOW serta mengoptimalkan parameter *Random Forest Classifier* dalam *Multi-Aspect sentiment analysis*.

Pada bab 2 membahas studi terkait penelitian yang berhubungan dengan penelitian ini, bab 3 membahas perancangan sistem yang akan dibangun, bab 4 membahas hasil skenario pengujian dan bab 5 membahas kesimpulan dan saran yang dilakukan dalam penelitian ini.