

## 1. Pendahuluan

### Latar Belakang

Tragedi yang akan dibahas dalam tugas akhir ini adalah tragedi di Stadion Kanjuruhan, Kabupaten Malang, pada pertandingan Arema melawan Persebaya tanggal 1 Oktober 2022. Tragedi tersebut mengakibatkan 135 orang meninggal dunia dan 500 lebih orang terluka.[1] Potongan-potongan video dan gambar serta opini masyarakat memenuhi linimasa sosial media. Terdapat beragam tanggapan masyarakat mengenai tragedi ini, seperti kecaman terhadap kepolisian, mencurahkan empati kepada korban tragedi, memprotes kelalaian federasi sepak bola Indonesia (PSSI), menyinggung aksi anarkis suporter, dan lain sebagainya.

Penelitian terkait tentang analisis sentimen sudah sering dilakukan. Berbagai variasi metode telah digunakan untuk menganalisis sentimen yang didapat dari media sosial atau media *online* lainnya. Penelitian Hermanto, Kuntoro, *et al.*[2] tentang analisis sentimen *review* di Google Play Store yang menggunakan Naïve Bayes dan Support Vector Machine (SVM) dengan keduanya menggunakan teknik SMOTE<sup>1</sup>. Disimpulkan bahwa performa metode SVM lebih tinggi dibandingkan dengan Naïve Bayes. Marcio Guia *et al.*[3] membandingkan hasil klasifikasi dengan SVM, Decision Tree, dan Random Forest pada ulasan telepon seluler di Amazon dengan pembobotan TF-IDF. Dari perbandingan ketiga *classifier*, mereka mengklaim bahwa SVM dengan kernel linear mempunyai nilai matriks: accuracy, precision, recall, dan F1-Score paling tinggi.

Menurut Al-Saqqa dan Awajan [4], Word2Vec adalah salah satu pemodelan ekstraksi fitur yang umum digunakan. Mereka melakukan survei pada 12 literatur tentang penggunaan Word2Vec dan mendapat dampaknya signifikan pada klasifikasi data. Penggunaan W2V dengan model skip-gram pada SVM [5]–[8], menghasilkan nilai matriks akurasi di atas 0,7.

Rizki pada sentimen vaksin Covid-19 dengan TF-IDF[9] dan Nooraeni *et al* pada sentimen RUU KPK dengan pelabelan Lexicon[10], keduanya menggunakan dataset dari Twitter dan menggunakan Support Vector Machine dengan kernel rbf sebagai *classifier*. Mereka berdua mengklaim mereka menghasilkan nilai matriks yang memuaskan.

Tujuan dari tugas akhir ini adalah untuk menganalisis opini masyarakat tentang tragedi stadion sepak bola terburuk kedua di dunia[11] tersebut melalui sosial media. Metode yang digunakan dalam analisis ini adalah word2vec dan support vector machine. Word2vec digunakan untuk mengubah kata-kata dalam teks menjadi vektor numerik, yang kemudian digunakan oleh support vector machine untuk melakukan klasifikasi sentimen, positif, netral, dan negatif. Hasil dari analisis ini diharapkan dapat memberikan informasi yang berguna bagi pembaca untuk mengetahui bagaimana sentiment mengenai tragedi ini secara luas.

### Topik dan Batasannya

Berdasarkan topik yang dibahas dan penelitian terdahulu, metode dan rancangan yang digunakan untuk menganalisis sentimen pada Tragedi Kanjuruhan pada Twitter memiliki rumusan masalah sebagai berikut:

- Bagaimana metode word2vec dan Support Vector Machine (SVM) dapat digunakan untuk menganalisis opini masyarakat di Twitter?
- Berapa akurasi dari metode yang digunakan?

Supaya penelitian terarah dan mempermudah dalam pengerjaan, batasan masalah ditetapkan sebagai berikut:

- Data hanya didapatkan dari Twitter.
- Analisis yang dilakukan hanya pada opini yang diunggah sejak 2 Oktober 2022 hingga 29 Desember 2022.
- Hanya menggunakan klasifikasi Support Vector Machine dengan kernel rbf dan linear
- *Tweet* yang diambil hanya berbahasa Indonesia.
- Kategori *tweet* (positif dan negatif) mungkin saja terdapat bias dari pelabel.

### Tujuan

Tujuan dari penelitian ini adalah sebagai berikut:

- Mengetahui bagaimana sentimen masyarakat mengenai tragedi di media sosial Twitter.
- Mengetahui performa word2vec dan Support Vector Machine