

BAB I

PENDAHULUAN

1.1. Latar Belakang

Rohingya adalah sebutan bagi orang-orang yang bermigrasi dari Bangladesh ke Arakan. Rohingya merupakan sekelompok muslim di wilayah Frontier Mayu, yang saat ini bernama Buthidaung dan kota Arakan. Sistem Zamindary merupakan aturan kepemilikan tanah yang dibawah oleh Inggris, sistem Zamindary tersebut menyebabkan pertikaian antara umat Muslim dan Buddha di Myanmar. Sampai sekarang, kekerasan yang terjadi di Myanmar masih terjadi, sehingga umat Muslim di Arakan terus berusaha menghadapi gerakan militer Junta. Pada tahun 1962 pemerintahan Myanmar memberikan izin tinggal sementara, tetapi pada tahun 1990 pemerintahan Myanmar memberikan batasan sehingga status kewarganegaraannya tidak lagi di akui oleh negara (Wasalmi Wasalmi, 2023). Hal tersebut mengakibatkan kaum Rohingya melakukan migrasi ke negara-negara lain seperti Malaysia, Thailand bahkan Indonesia.

Pertengahan Mei tahun 2015 terdapat laporan nelayan mengetahui masuknya pengungsi dari Myanmar dan Bangladesh yang karam di pantai timur Aceh Indonesia (Faisal et al., 2023). Akhir-akhir ini ramai masuknya pengungsi Rohingya dan diberitakan (Zulkarnaini, 2023) pada Desember 2023 terdapat kapal yang mengangkut 315 pengungsi etnis Rohingya di Kabupaten Aceh Besar dan Kabupaten Pidie Provinsi Aceh. Terkait kejadian tersebut banyak timbulnya pro dan kontra terkait hadirnya rohingya di Indonesia, serta terciptanya berbagai opini yang beredar di Indonesia.

Di era saat ini, media sosial menjadi bagian dari kehidupan masyarakat untuk menyampaikan tanggapan, keluhan serta opini dari berbagai peristiwa yang terjadi (Wati & Ernawati, 2021). Opini yang berkembang di media sosial mengenai masuknya Rohingya berkembang sangat cepat, salah satunya di media sosial X. X merupakan media sosial yang sebelumnya bernama twitter dan banyak digunakan karena kemudahan dalam penggunaannya, media sosial tersebut memudahkan pengguna bebas mengeluarkan opini berupa teks. Oleh karena itu, banyak peneliti menggunakan data penelitian dari media sosial tersebut (Giovani et al., 2020). Saat ini, Indonesia adalah

negara dengan jumlah pengguna media sosial X terbesar ke 3 dengan 77% pengguna aktif yang aktif menulis tweet (Geofany & Liza, n.d., 2021).

Analisa sentimen adalah kegiatan melakukan analisa terkait sensasi, emosi, sentimen, opini, sudut pandang, penilaian, sikap, yang dimiliki seseorang terhadap berbagai hal yang berkaitan dengan peristiwa, produk, layanan, atau organisasi. Sentimen berupa opini publik, opini tersebut dapat berupa teks, video, dan audio. Analisa sentimen nantinya akan mendeteksi bagaimana prespektif publik terhadap opini-opini yang beredar di masyarakat apakah positif, negatif atau netral (D'Aniello et al., 2022). Pandangan dan opini masyarakat menjadi sumber daya untuk mengetahui bagaimana prespektif masyarakat terhadap masuknya Rohingya baik pro maupun kontra. Untuk mengetahui penilaian pro dan kontra masyarakat dapat dilakukan analisa sentimen sehingga dapat merangkum pandangan publik terkait peristiwa tersebut.

Sebelum dilakukan analisa sentimen terdapat proses *pre-processing* data, salah satunya pelabelan data. Proses pelabelan secara manual membutuhkan proses yang lama, sehingga keterbatasan waktu, kemampuan energi serta biaya yang besar dalam pelabelan tersebut (Chan et al., 2023). Maka dari itu dibutuhkan proses labeling otomatis untuk memberikan label positif, negatif atau netral terhadap data. Metode untuk melabeli yang dapat digunakan adalah *lexicon*. *Lexicon* adalah metode kamus untuk mencari bobot dari setiap kata pada kalimat dan menghitung gabungan kata tersebut lalu menentukan sentimen berdasarkan pembobotan kalimat yang telah dihitung (D'Aniello et al., 2022).

VADER "*Valence Aware Dictionary and sEntiment Reasoner*" adalah sebuah kamus yang memiliki aturan dalam menganalisa sebuah kalimat melalui pebobotan pada setiap katanya. Setiap kalimat akan dibobotkan dan jika bobot tersebut bernilai diatas 0.05 maka hasil sentimennya positif, jika dibawah -0.05 maka sentimennya negatif, dan jika diantara 0.05 sampai -0.05 maka bernilai netral (Biswas et al., 2023). Selain VADER, terdapat juga InSet *lexicon*. InSet *lexicon* adalah kamus lexion dengan bahasa indonesia yang dikembangkan dengan data dari media sosial X dengan isi 3609 kata positif dan 6609 kata negatif. InSet *lexicon* memiliki pembobotan kata antara 5 sampai -5 yang memiliki akurasi sebesar 65.78% (Koto & Rahmaningtyas, 2017).

Dalam beberapa penelitian, dilakukan penggabungan metode yang bertujuan mengefisiensi dari segi waktu dan biaya untuk pelabelan. Penggabungan dalam penelitian telah dilakukan oleh (Isnain et al., 2023) dengan menggabungkan pelabelan

lexicon jenis VADER dictionary dengan SVM “*support vector machine*” menghasilkan F1 sebesar 80%. Penelitian lain dilakukan (Muhammadi et al., 2022) dengan melakukan pelabelan *lexicon* InSet yang digabungkan dengan SVM dengan judul “Combination of Support Vector Machine and Lexion Based Algorithm in Twitter Sentiment Analysis” menghasilkan akurasi sebesar 78.85%

Dengan adanya hal tersebut terjadi perbedaan akurasi antara penggunaan pelabelan *lexicon* InSet dan *lexicon* VADER sehingga menjadikan peneliti tertarik membuat perbandingan pelabelan *lexicon* pada analisa sentimen Rohingya untuk mengetahui performa yang lebih baik dalam pelabelan. Maka dari itu, peneliti menyimpulkan penelitian dengan judul “Perbandingan Performa Labeling *Lexicon* Inset dan *Lexicon* Vader pada Analisa Sentimen Rohingya di Aplikasi X dengan *Support Vector Machine*”

1.2. Rumusan Masalah

1. Bagaimana analisa sentimen rohingya di sosial media X menggunakan *lexicon* InSet dan *lexicon* VADER?
2. Bagaimana performa pelabelan *lexicon* InSet dan *lexicon* Vader dengan SVM dalam melakukan analisa sentimen Rohingya di sosial media X?
3. Bagaimana pengaruh komposisi data latih dan uji performa antara pelabelan *lexicon* vader dan *lexicon* inset dengan SVM pada analisa sentimen Rohingya di sosial media X?

1.3. Tujuan

1. Mengetahui analisa sentimen rohingya pada sosial media X dengan *lexicon* InSet dan *lexicon* VADER
2. Menghitung performa yang dihasilkan dalam analisa sentimen Rohingya di sosial media X dengan pelabelan *lexicon* vader dan *lexicon* inset dengan SVM
3. Mengetahui perngaruh komposisi data latih dan uji antara pelabelan *lexicon* vader dan *lexicon* inset dengan SVM pada analisa sentimen Rohingya di sosial media X.

1.4. Manfaat

1. Memberikan informasi sentimen yang dihasilkan mengenai masuknya Rohingnya di Indonesia di sosial media X

2. Mengetahui performa pelabelan *lexicon* InSet dan Vader terhadap kombinasi SVM pada analisa sentimen Rohingya pada sosial media X
3. Sebagai pengetahuan baru mengenai penggunaan pelabelan pada analisa sentimen

1.5. Batasan Masalah

1. Metode yang digunakan dalam pelabelan adalah *lexicon* Vader dan *lexicon* InSet
2. Metode dalam klasifikasi menggunakan support vector machine
3. Data yang diambil selama 14 hari dimulai dari 29 Desember 2023 – 11 Januari 2024.
4. Data yang digunakan adalah data dari media sosial X dengan bahasa Indonesia dengan penggunaan kata kunci Rohingya.