

BAB I

PENDAHULUAN

1.1 Latar Belakang

Keberadaan media sosial saat ini sudah menyebar luas di Indonesia. Berbagai kalangan menggunakan sosial media untuk bersosialisasi, bertukar pikiran, dan mendapatkan informasi. Salah satu contohnya yaitu ulasan online. Ulasan online merupakan salah satu jenis sumber daya buatan pengguna yang populer dan penting sehingga semakin banyak digunakan dalam proses pengambilan keputusan, karena pengguna dapat mempublikasikan pengalaman dan pendapat mereka tentang suatu produk, acara, atau layanan [1]. Untuk itu, banyak para peneliti menganalisis sentimen dari teks yang diposting para pengguna untuk mendapatkan apakah memiliki banyak tanggapan positif, negatif, atau netral dari sebuah topik yang sedang diperbincangkan.

Analisis sentimen merupakan teknik untuk menganalisis opini, sentimen, penilaian, dan emosi terhadap suatu entitas seperti produk, jasa, kejadian atau yang lainnya dengan mengelompokkan teks, kalimat, atau dokumen kemudian menentukan teks, kalimat, atau dokumen tersebut yang termasuk kedalam sentimen yang positif, negatif, atau netral [2]. Analisis sentimen menjadi salah satu cara untuk mengungkap makna yang tersembunyi dalam pengungkapan informasi [3].

VADER (Valence Aware Dictionary and Sentiment Reasoner) merupakan salah satu metode analisis dari lexicon based yang memiliki 7.500 tokens yang mengandung kata berbahasa inggris, emoticon, serta akronim dan inisial yang akan dianalisis dan menghasilkan class sentiment berupa positif, negatif, atau netral dengan tambahan skor compound [4].

Pada penelitian mengenai analisis sentimen pengungkapan informasi manajemen menggunakan metode lexicon based dengan pendekatan VADER atau Valence Aware Dictionary and Sentiment Reasoner. Dasar pemilihan Vader ini dikarenakan nilai disetiap kata dalam teks murni penilaian manusia, bahkan vader juga mampu menangkap makna yang tersirat dari tanda baca yang ada di dalam

sebuah teks. Pemberian skor pada penelitian tersebut bervariasi dari -1 hingga +1, dengan nilai di atas 0,05 sebagai positif dan di bawah -0,05 dinyatakan negatif. Nilai di antara -0,05 hingga 0,05 dinyatakan netral. Hasil dari penelitian tersebut yaitu nilai positif dan negatif terbilang rendah dibanding sentimen yang netral, dapat diartikan bahwa sentimen yang cenderung netral karena pihak manajemen enggan untuk menyampaikan berita buruk yang dapat mempengaruhi reputasi perusahaan dan pengungkapan tersebut tanpa kata sifat dan kata keterangan sehingga sentimen yang didapatkan cenderung netral [3].

Dalam penelitian lain, pendekatan VADER digunakan untuk menganalisis sentimen pada twitter mengenai apex legends. Apex legends merupakan salah satu contoh game yang populer di dunia. Analisis sentimen ini dilakukan untuk mendapatkan informasi mengenai pendapat para pemain tentang permainan ini. Pada penelitian tersebut menggunakan data sebanyak 500 tweet. Hasil yang didapatkan dari penelitian tersebut yaitu mendapatkan nilai positif sebanyak 18%, nilai negatif sebanyak 4,8%, dan nilai netral sebanyak 77,2% [5].

Lalu dalam penelitian lainnya juga menggunakan pendekatan VADER, namun dalam penelitian ini pendekatan VADER digunakan untuk pelabelan data yang akan menilai kata yang dibawah 0 dinyatakan negatif, dan diatas 0 dinyatakan positif. Hasilnya adalah terdapat 469 sentimen positif dan 31 sentimen negatif [6].

Pada penelitian *sentimental analysis on audio and video using vader algorithm*, menggunakan ulasan produk dari para pengguna dari berbagai macam media seperti Facebook, Twitter, Youtube, dan yang lainnya dalam bentuk video dan audio yang diekstrak terlebih dahulu ke dalam bentuk teks kemudian diproses menggunakan algoritma vader [7].

Pada penelitian *a complete VADER-based sentiment Analysis of Bitcoin (BTC) Tweets during the Era of COVID-19*, menggunakan tweet dari para pengguna twitter untuk mencari informasi mengenai perkiraan harga bitcoin selama pandemi covid-19 yang menggunakan skor vader dari strategi preprocessing yang lebih kompleks karena menurutnya hasil preprocessing yang baik akan meningkatkan korelasi skor sentimen dan skor polaritas dengan harga bitcoin, serta

akan meningkatkan akurasi yang lebih baik dengan harga yang sebenarnya [8].

Penelitian GerVADER – a German adaptation of the VADER sentiment analysis tool for social media texts, merupakan sebuah penelitian yang membangun sebuah alat analisis sentimen berbasis leksikon berbahasa Jerman yang diadaptasi dari VADER yang dinamakan Ger-VADER [9].

Dari beberapa penelitian sebelumnya ada banyak sekali penelitian yang menggunakan pendekatan VADER, namun penelitian tersebut langsung memakai library VADER yang sudah disediakan. Pada penelitian ini, penulis ingin membangun sebuah prototype aplikasi berbasis website, dimana nantinya para user akan dapat langsung menginputkan data yang sudah di crawling ke dalam aplikasi tersebut dan nantinya akan langsung menghasilkan sentimen dengan pendekatan VADER. Prototype website tersebut akan digunakan untuk meneliti lebih lanjut berapa tingkat akurasi yang dihasilkan VADER menggunakan Naïve Bayes Classifier.

1.2 Rumusan Masalah

Berdasarkan pada permasalahan yang telah dijelaskan pada bagian latar belakang, maka rumusan masalah dapat disusun sebagai berikut:

1. Berapa tingkat akurasi pelabelan Vader pada uji coba analisis sentimen?
2. Apakah pelabelan Vader tersebut dapat diimplementasikan sebagai prototype aplikasi pelabelan berbasis website?

1.3 Batasan Masalah

Agar tidak meluasnya pembahasan, maka skripsi ini dibatasi dengan hal sebagai berikut:

1. Hasil yang didapatkan dari uji coba analisis sentimen menggunakan pelabelan vader ini adalah tingkat akurasi dari pelabelan vader.
2. Website yang dibangun hanya sebagai prototype.
3. Hanya dapat menggunakan dataset yang memiliki tiga kelas sentimen yaitu positif, netral, dan negatif.
4. Dataset harus melewati preprocessing terlebih dahulu.

1.4 Tujuan Penelitian

Tujuan penelitian yang ingin dicapai oleh penulis dalam penelitian ini adalah sebagai berikut:

1. Mengetahui tingkat akurasi dari metode pelabelan vader.
2. Menghasilkan sebuah prototype aplikasi berbasis website.

1.5 Manfaat Penelitian

Manfaat yang didapatkan dari skripsi ini yaitu:

1. Hasil tingkat akurasi yang didapatkan dapat menjadi bahan pertimbangan untuk menggunakan pelabelan VADER dalam penelitian analisis sentimen lainnya.
2. Prototype aplikasi yang dihasilkan dapat membantu pelabelan VADER menjadi lebih mudah digunakan.

1.6 Sistematika Penulisan

.Sistematika penulisan dalam skripsi ini sebagai berikut:

BAB I PENDAHULUAN, berisi Latar belakang masalah, rumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian, dan sistematika penelitian.

BAB II TINJAUAN PUSTAKA, berisi tinjauan pustaka dan dasar-dasar teori yang digunakan.

BAB III METODE PENELITIAN, didalamnya terdapat tinjauan umum tentang objek penelitian, alur penelitian, serta alat dan bahan.

BAB IV HASIL DAN PEMBAHASAN, bab ini merupakan tahapan yang penulis lakukan dalam merancang prototype website, pengumpulan data, preprocessing data, pelabelan vader menggunakan prototype website, dan klasifikasi naïve bayes.

BAB V PENUTUP, berisi kesimpulan dan saran yang dapat peneliti rangkum selama proses penelitian.