

BAB I

PENDAHULUAN

1.1. Latar Belakang

Olahraga adalah bentuk kegiatan mengolah anggota tubuh yang dilakukan oleh manusia secara terstruktur dan beraturan guna menyehatkan dan membuat tubuh menjadi bugar. Olahraga sendiri juga dapat dilakukan dengan cara mandiri maupun berkelompok. Dalam undang-undang Republik Indonesia Nomor 3 Tahun 2005 tentang sistem keolahragaan nasional pada Bab 1 Pasal 1 berbunyi Keolahragaan adalah semua aspek yang bersangkutan dengan olahraga yang memerlukan pengaturan, pendidikan, pelatihan, pembinaan, pengembangan, dan pengawasan[1]. Olahraga memiliki banyak pemintat di dunia ini, oleh sebab itu ada banyak acara dan event olahraga yang dibuat untuk mempertandingkan olahraga tersebut, salah satu event yang paling terkenal di Asia Tenggara adalah SEA Games.

Southeast Asian Games (SEA Games) merupakan salah satu acara olahraga yang diselenggarakan di wilayah Asia Tenggara saja. Acara ini dilaksanakan setiap dua tahun sekali dan melibatkan sebelas negara anggota *Association of South East Asian Nation* (ASEAN). SEA Games memiliki banyak tujuan yaitu untuk mengeratkan kerja sama antar negara anggota ASEAN, menyatukan pemahaman dan mempererat hubungan antar negara anggota ASEAN di kawasan semenanjung Asia Tenggara[2].

Acara Sea Games ini sempat tertunda selama setahun karena pandemi Covid-19 yang melanda dunia. Karena hal ini ada banyak opini yang bertebaran di masyarakat terhadap kegiatan olahraga ini baik itu tanggapan positif ataupun negati, ada yang mendukung terlaksananya kegiatan ini dan ada juga yang menolak kegiatan ini diadakan untuk sekarang karena dunia masih dalam situasi *pandemic* covid-19. Salah satu media tempat orang beropini dan memberikan pendapat mereka adalah media sosial Twitter.

Twitter merupakan jenis media sosial *microblogging* yang memfasilitasi pengguna untuk menulis dan memublikasikan aktivitas serta atau pendapatnya. Secara historis, kehadiran dan munculnya media sosial Twitter yang menyediakan ruang tertentu atau maksimal 140 karakter. Sama seperti media sosial lainnya, di Twitter pengguna bisa menjalin jaringan dengan pengguna

lain, menyebarkan informasi, mempromosikan pendapat / pandangan pengguna lain, sampai membahas isu terhangat (trending topic) saat itu juga dan menjadi bagian dari isu tersebut dengan turut berkicau (tweet) menggunakan tagar (hashtag) tertentu[3]. Twitter digunakan sebagai media untuk melakukan penelitian sentiment analisis terhadap topik yang dinahas berdasarkan opini masyarakat dari twitter.

Sentiment analysis merupakan proses untuk memahami dan mengolah data tekstual secara otomatis untuk mendapatkan informasi sentimen yang terkandung dalam suatu kalimat opini. *Sentiment analysis* atau *opinion mining* mengacu pada bidang yang luas dari pengolahan bahasa alami, komputasi *linguistik* dan *text mining* yang bertujuan menganalisa pendapat, sentimen, evaluasi, sikap, penilaian dan emosi seseorang apakah pembicara atau penulis berkenan dengan suatu topik, produk, layanan, organisasi, individu, ataupun kegiatan tertentu [4]. Pada penelitian ini, peneliti menggunakan metode Naïve Bayes untuk proses analisis sentimen nya.

Naive Bayes adalah merupakan salah satu algoritma klasifikasi yang banyak digunakan pada *data mining* ataupun *text mining*. Algoritma Naive Bayes berdasarkan teorema Bayes bahwa semua kegiatan memberikan sebuah kontribusi yang sama penting atau saling bebas pada pemilihan kelas tertentu[5].

Penelitian ini juga menggunakan proses *text mining*, dimana *text mining* merupakan bagian dari *data mining*, namun proses *text mining* memerlukan lebih banyak tahapan dibandingkan dengan *data mining* karena data teks memiliki karakteristik yang lebih kompleks dari pada data biasa bahkan data yang sudah terstruktur. Berdasarkan ketidakaturan struktur data teks, maka proses *text mining* memerlukan beberapa tahap awal yang pada intinya adalah mempersiapkan agar teks dapat diubah menjadi lebih terstruktur.

Preprocessing adalah tugas dan langkah penting dalam *Text Mining*, *Natural Language Processing (NLP)* dan *Information Retrieval (IR)*. Di bidang *Text Mining*, data *preprocessing* digunakan untuk mengekstraksi pengetahuan yang menarik dan penting serta dari data teks yang tidak terstruktur. Berdasarkan ketidak aturan struktur data teks, maka proses sistem temu kembali informasi ataupun *text mining* memerlukan beberapa tahap awal yang pada intinya adalah mempersiapkan agar teks dapat diubah menjadi lebih terstruktur. Tahap *Text Preprocessing* adalah tahapan dimana aplikasi melakukan seleksi data yang akan diproses pada setiap dokumen. Proses *preprocessing* ini meliputi (1) *case folding*, (2) *tokenizing*, (3) *filtering*, dan (4) *stemming*[6].

Berdasarkan penjelasan yang telah dipaparkan diatas, maka penelitian ini bertujuan untuk mengetahui komentar masyarakat tentang *SEA Games 2021* pada Twitter yang akan dituangkan dalam laporan Tugas Akhir yang berjudul “Analisis Sentimen Pengguna Twitter Terhadap *SEA Games Vietnam 2021* Menggunakan Metode Naïve Bayes”. Alasan pengambilan judul ini yaitu peneliti berharap dari penelitian yang dilakukan terhadap data yang ada di Twitter untuk mengetahui tanggapan pengguna twitter terhadap *SEA Games 2021* secara akurat dengan menggunakan dua metode tersebut.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan diatas, maka rumusan masalah yang terdapat pada penelitian ini adalah belum adanya hasil penelitian pengelompokan pandangan positif, negatif, dan netral terhadap *SEA Games 2021*

1.3 Pertanyaan Penelitian

Berapa nilai akurasi perhitungan sentiment positif, negatif, dan netral berdasarkan perhitungan dari proses *Naïve Bayes* dalam analisis sentimen pengguna twitter terhadap *SEA Games 2021* ?

1.4 Tujuan Penelitian

Mengetahui nilai akurasi perhitungan sentiment positif, negatif, dan netral dari penggunaan *Naïve Bayes* dalam mengklasifikasikan sentimen pengguna twitter terhadap *SEA Games 2021*.

1.5 Batasan Masalah

Berdasarkan rumusan masalah dan tujuan penelitian, maka diperoleh batasan-batasan masalah penelitian sebagai berikut:

1. Data hanya berupa klasifikasi dari tweet atau teks yang ada di twitter
2. Tweet berisi kata kunci yaitu, *Sea Games 2021, SEA Games Vietnam*.
3. Tweet dengan konten promosi produk diabaikan dan kolom tweet tidak diaktifkan fitur “Lindungi Tweet Anda” atau hidden oleh pengguna Twitter.
4. Tweet diambil pada tanggal 6 sampai 23 Mei 2022 dari pengguna twitter Indonesia dan tweet berbahasa Indonesia
5. Metode yang digunakan untuk klasifikasi adalah *Naïve Bayes*.

1.6 Manfaat Penelitian

Manfaat dari penelitian ini adalah :

1. Dapat menjadi acuan pada penelitian lainnya yang sesuai dengan penelitian yang dilakukan ini yaitu sentiment analisis.
2. Sebagai tambahan ilmu pengetahuan tentang sentiment analisis bagi yang ingin memahami tentang sentiment analisis sebagai referensi yang dapat dibaca.