

BAB I

PENDAHULUAN

1.1 Latar Belakang Masalah

Perkembangan teknologi digital saat ini yang terus berkembang memberikan kemudahan bagi manusia untuk melakukan aktivitas sehari-hari. Teknologi juga membuat segala sesuatu berjalan dengan mudah, praktis, dan cepat. Penggunaan teknologi itu sendiri sudah banyak diterapkan untuk melakukan pengelolaan data, penyajian informasi, dan mempermudah komunikasi antar manusia, dikarenakan daya efektivitas dan efisiensinya serta keakuratan yang terkomputerisasi.

Kemajuan ini membuat manusia mengalami perubahan terlebih dalam hal berkomunikasi, terutama berkomunikasi jarak jauh. Munculnya berbagai aplikasi media sosial yang memiliki ciri khas masing-masing dan selalu menyajikan inovasi-inovasi terbaru dalam penyajian informasi membuat masyarakat terus memanfaatkannya.

Media sosial merupakan sebuah media di mana memungkinkan penggunaanya untuk bersosialisasi dan berinteraksi satu sama lain yang dilakukan secara *online* tanpa dibatasi ruang dan waktu. Seiring perkembangannya media sosial tidak hanya digunakan untuk sekedar memposting gambar dan bertukar kabar. Salah satu pemanfaatan media sosial lainnya yaitu media sosial digunakan untuk mengungkapkan dan berbagi pendapat, pandangan, usul, bahkan kritik terhadap suatu topik, isu, berita dan produk yang dibagikan ke ruang publik. Masyarakat bebas memberikan respon dan opininya di berbagai media sosial dan menjadikan media sosial tempat untuk mencurahkan segala hal (Fitriana dkk, 2021).

Salah satu media sosial yang banyak digunakan untuk menyuarakan pendapat adalah Twitter. Per Januari 2023, Indonesia berhasil menjadi negara di peringkat kelima dengan pengguna aktif Twitter sebanyak 24 juta pengguna (Hootsuite We Are Social, 2023). Hal ini bisa disebabkan oleh salah satu fitur yang ditawarkan oleh Twitter yaitu, fitur tagar (hashtag) dan trending topic yang dimilikinya. Fitur ini membantu pengguna untuk melihat dan membantu pengguna untuk mengikuti topik, berita, fenomena, peristiwa dan percakapan yang mereka minati ataupun yang sedang ramai diperbincangkan.

Seperti salah satu topik yang mencuat belakangan ini adalah terkait dengan Childfree. Childfree adalah sebuah keputusan yang diambil oleh seseorang atau pasangan untuk tidak memiliki keturunan dalam pernikahan mereka. Banyak hal yang dijadikan alasan oleh seseorang untuk memilih Childfree di dalam kehidupannya, seperti kekhawatiran dalam perkembangan sang anak, masalah pribadi, ekonomi, bahkan faktor lingkungan sekitarnya (Siswanto & Nurhasanah, 2022).

Topik ini menjadi ramai diperbincangkan oleh masyarakat dikarenakan salah satu pegiat media sosial, Gita Savitri dan suaminya yang memutuskan untuk Childfree dan memberikan pandangan terkait keputusan tersebut ke ruang publik. Hal ini memunculkan berbagai komentar pro kontra di masyarakat sehingga cukup banyak masyarakat menjadikan topik ini sebagai bahasan atau diskusi di berbagai media sosial. Berdasarkan opini masyarakat tersebut, dapat dilakukan sebuah analisis, yaitu analisis sentimen. Analisis sentimen adalah proses menganalisis teks digital untuk menentukan apakah nada emosional pesan tersebut positif, negatif, atau netral.

Analisa sentimen mendeteksi suatu polaritas, yaitu opini positif atau negatif di dalam suatu teks, yang dapat berupa dokumen, paragraf, kalimat, atau klausa (Priyadarshini & Chase, 2021). Informasi yang terdapat pada sentimen dapat membantu publik untuk merasakan opini suatu produk, peristiwa politik, isu, kebijakan, bisnis, dan lainnya (Kumar & Geetanjali, 2019).

Analisa sentimen juga dapat disebut dengan text mining. Text mining merupakan suatu kegiatan menambang data yang berupa teks, di mana sumber data biasanya di dapatkan dari sebuah dokumen dan tujuannya untuk mencari kata-kata yang dapat mewakili isi dari dokumen sehingga dapat dilakukan analisa keterhubungan antar dokumen (Harieby dkk, 2022).

Text mining merupakan salah satu bagian dari data mining, di mana data mining adalah sebuah studi tentang mengumpulkan, membersihkan, memproses, menganalisis, dan memperoleh manfaat berupa pengetahuan atau wawasan dari data (Aggarwal, 2015). Data mining memiliki bentuk yang lebih terukur dan terstruktur, dibandingkan dengan text mining yang memiliki bentuk tekstual yang tidak terstruktur.

Salah satu penerapan analisis sentimen dilakukan untuk menganalisa opini masyarakat terhadap vaksin Covid-19 pada Twitter menggunakan metode Support Vector Machine dan Naïve Bayes. Hasil penelitian ini menunjukkan bahwa algoritma SVM memiliki performa lebih baik dibandingkan dengan algoritma Naïve Bayes. Algoritma SVM dengan nilai akurasi 90,47%, presisi 90,23%, dan *recall* 90,78%. Algoritma Naïve Bayes dengan nilai akurasi 88,64%, presisi 87,32%, dan *recall* 88,13%. Selanjutnya untuk hasil sentimen, sentimen analisis netral memperoleh nilai 8,76%, negatif 42,92%, dan positif 48,32% untuk Naïve Bayes dan netral 10,56%, negatif 41,28%, dan positif 48,16% untuk Support Vector Machine (Fitriana dkk, 2021).

Perbandingan beberapa metode juga dilakukan pada analisis sentimen terhadap opini masyarakat terkait dengan isu kebocoran data kartu identitas ponsel dengan menggunakan kata kunci Twitter yaitu, sim card, data sim, bocor data, dan kominfo. Kemudian diklasifikasikan dengan empat metode, yaitu Random Forest, Logistic Regression, Support Vector Machine, dan model IndoBERT. Hasil penelitian menunjukkan algoritma SVM memiliki performa terbaik dilihat dari nilai *f1-score* yang paling tinggi dibandingkan dengan algoritma lainnya. Random Forest memiliki *f1-score* yaitu 0.78, Logistic Regression 0.74, Support Vector Machine 0.81, dan IndoBERT 0.76 (Amal dkk, 2022).

Penelitian lainnya dilakukan untuk menganalisis sentimen pada pariwisata di tiga wilayah Thailand, yaitu Bangkok, Phuket, dan Chiang Mai selama pandemi Covid-19. Dengan membandingkan tiga algoritma, yaitu CART, Random Forest, dan Support Vector Machine. Hasil penelitian menunjukkan bahwa algoritma SVM memberikan hasil terbaik dalam melakukan analisis sentiment dengan akurasi maksimal, yaitu 77.4%. (Leelawat dkk, 2022).

Tidak hanya untuk sektor pariwisata, analisa sentimen dilakukan untuk melihat antusiasme publik tentang dirilisnya sebuah produk dipasaran, serta untuk melihat apa saja pertimbangan masyarakat untuk menilai suatu produk terbaru dari Apple dan

apakah mendapatkan respon yang baik atau tidak. Penelitian dilakukan menggunakan metode Support Vector Machine. Hasil penelitian menunjukkan algoritma SVM dapat memberikan hasil yang memuaskan dengan nilai akurasi, presisi dan *f1-score*, yaitu 89,21%, 92,43%, dan 93,95% (Bourequat & Hassan, 2021).

Analisa sentimen juga dapat dilakukan untuk melihat respon publik terhadap kebijakan yang dibuat oleh pemerintah India. Sentimen diklasifikasikan menggunakan algoritma Naïve Bayes, Decision Tree, Random Forest, dan Support Vector Machine. Hasil penelitian ini menunjukkan bahwa algoritma Random Forest memiliki nilai akurasi tertinggi dalam pengklasifikasian di antara algoritma lainnya. Algoritma Random Forest mendapatkan nilai presisi, *recall* dan *f1-score*, yaitu 0.99, 0.92, dan 0.95 untuk sentimen negatif, untuk sentimen netral, yaitu 0.95, 0.99, dan 0.97, serta sentimen positif 0.98, 0.95, 0.96, dengan menggunakan pembobotan *Bag of Words*. Sedangkan dengan menggunakan pembobotan TF-IDF, Algoritma Random Forest mendapatkan nilai presisi, *recall* dan *f1-score*, yaitu 0.96, 0.93, dan 0.94 untuk sentimen negatif, untuk sentimen netral, yaitu 0.97, 0.98, dan 0.97, serta sentimen positif 0.95, 0.96, dan 0.96 (Neogi dkk, 2021).

Analisis sentimen dengan metode Support Vector Machine juga dilakukan untuk melihat bagaimana respon pengguna Gojek terhadap pelayanan yang di berikan oleh Gojek di media sosial Twitter. Hasil penelitian menunjukkan bahwa *tweets* pengguna Gojek dengan pelabelan secara manual cenderung bersentimen positif, dengan 1.142 *tweets* bersentimen positif dan 358 *tweets* bersentimen negatif. Sedangkan pelabelan dengan *sentiment scoring* belum bisa dijadikan acuan karena terdapat 172 *tweets* memperoleh label yang tidak tepat. Hasil klasifikasi sentimen dengan pelabelan data secara manual menggunakan algoritma SVM memiliki tingkat akurasi keseluruhan terbaik sebesar 79,19% dan akurasi kappa terbaik sebesar 16,52%. Sedangkan hasil klasifikasi sentimen dengan pelabelan data *sentiment scoring* menggunakan algoritma SVM memiliki tingkat akurasi keseluruhan terbaik sebesar 79,19% dan akurasi kappa terbaik sebesar 21% (Fitriyah dkk, 2020).

Dalam penelitian ini dilakukan analisis sentimen dengan tujuan untuk melihat fenomena apa yang sedang terjadi di tengah masyarakat terkait Childfree kemudian bagaimana kecenderungan opini publik terhadap isu Childfree yang dituangkan pada media sosial Twitter. Nantinya opini masyarakat terkait Childfree akan dianalisa kemudian diklasifikasi ke dalam tiga sentimen, yaitu sentimen positif, negatif, dan netral dengan menggunakan metode Support Vector Machine.

Metode Support Vector Machine digunakan karena memiliki konsep yang lebih matang dan jelas secara matematis dibandingkan dengan teknik klasifikasi lainnya serta memiliki tingkat akurasi yang lebih baik dan melakukan prediksi lebih cepat.

1.2 Perumusan Masalah

Berdasarkan latar belakang permasalahan yang telah diuraikan sebelumnya, maka didapatkan perumusan masalah sebagai berikut:

1. Bagaimana menerapkan metode Support Vector Machine dalam melakukan analisis sentimen terhadap opini Childfree?
2. Bagaimana mengetahui kecenderungan opini publik terhadap isu Childfree yang dituangkan pada media sosial Twitter?
3. Bagaimana menguji tingkat akurasi penerapan metode Support Vector Machine dalam melakukan analisis sentimen terhadap opini Childfree?

1.3 Batasan Masalah

Batasan masalah bertujuan untuk membantu dalam mengidentifikasi masalah, membatasi ruang lingkup dan menjadi gambaran terkait hal yang akan diteliti sehingga penelitian menjadi lebih terarah dan tidak menyimpang dari pokok bahasan dan permasalahan. Untuk membatasi ruang lingkup penelitian, maka batasan masalah dibatasi sebagai berikut:

1. Data yang digunakan merupakan data yang bersumber dari media sosial Twitter.
2. Data yang akan digunakan dalam penelitian ini merupakan data dari Tweet yang diposting pada media sosial Twitter dengan menggunakan kata kunci, yaitu Childfree.
3. Tweet yang digunakan hanya tweet yang menggunakan Bahasa Indonesia.
4. Metode yang digunakan pada penelitian ini adalah metode Support Vector Machine.

1.4 Tujuan Penelitian

Tujuan penelitian merupakan suatu hasil akhir atau sesuatu yang ingin dicapai dalam suatu penelitian. Diperlukan suatu perencanaan dan pelaksanaan penelitian dengan matang supaya tujuan penelitian dapat dicapai. Dalam penelitian ini ada beberapa tujuan yang hendak dicapai, yaitu:

1. Memahami penerapan metode Support Vector Machine dalam melakukan analisis sentimen terhadap opini Childfree serta menganalisa kecenderungan opini publik terhadap isu Childfree yang dituangkan pada media sosial Twitter.
2. Merancang metode Support Vector Machine untuk mengetahui kecenderungan opini publik terhadap isu Childfree yang dituangkan pada media sosial Twitter.
3. Mengimplementasikan serta menguji tingkat akurasi metode Support Vector Machine dalam melakukan analisis sentimen terhadap opini Childfree.

1.5 Manfaat Penelitian

Suatu penelitian yang dilakukan diharapkan dapat memberi manfaat baik bagi kepentingan peneliti, organisasi, maupun ilmu pengetahuan bagi peneliti selanjutnya. Manfaat penelitian merupakan jawaban atas tujuan penelitian yang dilakukan. Adapun manfaat yang akan didapatkan dalam penelitian ini adalah sebagai berikut :

1. Dapat menjadi bahan ilmu pengetahuan baru serta sebagai perbandingan dan sumber acuan untuk bidang kajian yang sama.
2. Dapat memberikan gambaran bagaimana fenomena dan pembahasan mengenai Childfree terus berkembang di Indonesia.

1.6 Sistematika Penulisan

Sistematika penulisan merupakan sebuah kerangka atau urutan penelitian agar lebih terstruktur. Terdapat enam BAB untuk mekanisme penelitian ini dengan gambaran umum sebagai berikut:

BAB I : PENDAHULUAN

Bab ini berisikan pendahuluan dan latar belakang masalah dalam melakukan penelitian. Adapun yang akan dibahas pada bab ini, yaitu perumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian yang diakhiri dengan sistematika penulisan.

BAB II : LANDASAN TEORI

Bab ini menjelaskan tentang beberapa hal yang mendasar tentang masalah yang akan dibahas, yang akan digunakan sebagai landasan dasar dalam melakukan pemecahan masalah pada penelitian yang dilakukan.

BAB III : METODOLOGI PENELITIAN

Pada bab ini berisikan tentang langkah atau kerangka kerja dan prosedur dalam melakukan analisis sentimen dengan menggunakan metode Support Vector Machine.

BAB IV : ANALISA DAN PERANCANGAN

Bab ini memuat tentang analisa data, pengolahan data Twitter mengenai opini terhadap Childfree. Selain itu bab ini juga memuat hasil dari penggunaan metode Support Vector Machine.

BAB V : IMPLEMENTASI DAN PENGUJIAN

Pada bab ini berisi penerapan metode Support Vector Machine menggunakan bahasa pemrograman Python sebagai aplikasi pengklasifikasian data yang diambil dari twitter.

BAB VI : KESIMPULAN DAN SARAN

Bab ini menjelaskan tentang kesimpulan dari hasil penelitian dan saran-saran untuk pihak lain dalam melakukan penelitian dan pengembangan untuk selanjutnya.