BAB I

PENDAHULUAN

1.1 Latar Belakang

Internet mengalami perkembangan yang pesat dari waktu ke waktu, bahkan hingga memegang peran penting dalam menyebarkan informasi yang dianggap penting bagi penggunanya. Perkembangan ini diikuti dengan pesatnya perkembangan sosial media karena meluasnya penggunaan di seluruh dunia (*Sharma* dan *Sharma*, 2020). Setiap hari pengguna sosial media terus meningkat. Tercatat pada tahun 2019, penggunaan sosial media diperkirakan naik hingga 2,77 miliar pengguna di dunia. Beragam tipe informasi di upload dan dibagikan di sosial media baik dalam bentuk teks, video, foto dan audio (*Drus* dan *Khalid*, 2019). Kebanyakan penggunaan sosial media digunakan untuk mengekspresikan perasaan, keseharian hingga pendapat mengenai berbagai macam hal (*Sharma* dan *Sharma*, 2020b).

Besarnya perkembangan Web 2.0 telah merubah cara seorang individu dalam mengekspresikan perspektif dan perasaan seseorang. Mulai dari komentar, pendapat, usul, pandangan bahkan hingga kritik yang berkaitan dengan suatu topik tertentu dibagikan dala jejaring sosial online (Alarifi dkk, 2020). Salah satu sosial media yang banyak dipakai adalah *twitter*. Penggunaan *twitter* di Indonesia mencapai 59% dan peringkat 5 sebagai sosial media yang sering digunakan pada tahun 2020 (Taofik dan Erry, 2021). *Twitter* banyak digunakan untuk menyampaikan keluh kesah dan opini masyrakat terhadap suatu pemikiran ataupun layanan (Rani dan Agus, 2020). Sosial media kaya akan data mentah dan belum diproses, namun meningkatnya teknologi khususnya dalam *machine learning* dan *artificial intelligence*

memungkinkan data untuk diproses ke dalam bentuk yang berguna (*Drus* dan *Khalid*, 2019b).

Data dapat memungkinkan untuk menafsirkan aspek seseorang pada topik tertentu dan informasi yang dapat dimanfaatkan untuk melakukan prediksi dalam penjualan produk, pasar saham, ataupun pemilihan topik, untuk mengambil keputusan secara tepat berdasarkan pertimbangan, analisis sentiment dianggap tepat (Pathak dkk, 2021). Analisis sentimen dapat digunakan untuk mengekstrak sentimen dan pendapat dari teks. Saat ini, analisis sentimen telah diakui. Tidak hanya dari peneliti tetapi juga perusahaan, pemerintah dan organisasi. Pendapat dan sentiment yang sering di *upload* di sosial media sangat sesuai dengan kehidupan sehari-hari dan karenanya perlu dianalisis untuk memantau opini publik dan membantu dalam pengambilan keputusan secara otomatis (*Brijali* dkk, 2021).

Analisis Sentimen dapat mendeteksi polaritas suatu opini ke dalam opini positif atau negatif dalam suatu teks, yang dapat berupa dokumen, paragraf, kalimat ataupun klausa. Analisis sentimen dapat digunakan dalam menganalisis sentimen pelanggan terhadap suatu produk, maupun melakukan analisis sentimen pada survei dan sosial media yang dapat membangun suatu merek atau produk dengan tujuan agar suatu perusahaan dapat menyesuaikan produk sesuai dengan kebutuhan pelanggan (*Priyadarshini* dan *Cotton*, 2022). Analisis sentimen tidak hanya digunakan untuk kepentingan bisnis, informasi yang tersedia dapat digunakan untuk kepentingan sosial, Pendidikan ataupun hiburan. Kesulitan yang muncul terletak pada melakukan analisis informasi, sehingga Analisis sentimen dapat melakukan analisis terhadap informasi yang dibutuhkan (*Kumar* dan *Garg*, 2020).

Machine learning merupakan metode yang digunakan dalam penerapan analisis sentimen (Primandani dan Retno, 2021). Sistem dilatih berdasarkan kumpulan data pelatihan, dan setelah pelatihan, sistem mencoba membuat klasifikasi otomatis atau dapat membuat prediksi. Efisiensi sistem diuji berdasarkan nilai prediksi yang benar dan kumpulan data uji (*Rameshwer* dan *Rajeshwar*, 2021). Beberapa metode yang sering digunakan adalah *Naïve Bayes*, K-NN, *Random Forest* dan *Support Vector Machine* (Primandani dan Retno, 2021). Metode dalam latar

belakang ini yang dibahas dibatasi pada *Naïve Baye*s dan *Support Vector Machine* (SVM) saja.

Metode *Naive Bayes* merupakan algoritma pemantauan *machine learning* yang mendasar dan terkenal berdasarkan teorema Bayesian dan anggapan kondisi karakteristik otonom. Metode ini mendalilkan aksesibilitas pada tahap dokumen dengan tampilan dan label yang telah ditentukan sebelumnya, sedangkan SVM *konvensional* membutuhkan satu set informasi input dan memprediksi kelas yang mungkin sebagai output untuk setiap item yang ditentukan (*Kanika* dan *Rajni*, 2021).

Penelitian sebelumnya terdapat penelitian mengenai analisis sentimen menggunakan metode SVM, *Naïve Bayes* dan *Logistic* dalam melihat sentimen pengguna aplikasi Tokopedia. Penelitian ini menggunakan 3125 review yang berasal dari pengguna Tokopedia. Penelitian ini menggunakan metode *web scraping* dalam mengumpulkan atau ekstraksi data semi-terstruktur dari *website*. *Review* dari penelitian ini diklasifikasikan menjadi sentimen positif dan negatif. Hasil dari penelitian ini adalah terbaginya 3125 data *review* menjadi sentimen positif sebanyak 2598 *review* dan sentimen negatif sebanyak 527 *review*. Hasil yang didapat, di mana sentiment positif berkaitan dengan kemudahan penggunaan aplikasi Tokopedia dan sentiment negatif berkaitan dengan sistem aplikasi *error* dan sering terjadi *logging* atau *bug* (Meishita dan Iqbal, 2022).

Metode SVM dan *Naïve Bayes* juga digunakan untuk melihat sentimen terhadap *MyIndiHome*. Penelitian ini menggunakan data yang dikumpulkan dalam rentang waktu 1 November 2020 hingga 15 Desember 2020 dengan total 2.539, diperoleh dari aplikasi *Google Play*. Hasil dari penelitian ini menunjukkan bahwa tingkat *error* yang terjadi pada Indihome masih terbilang tinggi. Rata-rata total akurasi metode SVM sebesar 86,54% lebih baik dari *Naïve bayes* (Sulton dkk, 2021).

Analisis Sentimen terhadap Opini mengenai vaksinasi *Covid*-19 menggunakan metode *Naïve Bayes* dan SVM. Penelitian dilakukan utnuk melihat opini masyarakat mengenai Covid-19, di mana pada saat itu masyarakat banyak menuangkan opininya ke sosial media. Penelitian ini menggunakan 1000 data yang

berkisar dari tahun 2020 yang berasal dari *twitter*. Hasil dari penelitian ini untuk melihat perbandingan anatara *Naïve Bayes* dan SVM. SVM memiliki performa yang lebih baik dari pada *Naïve Bayes* (Frizka dkk, 2021).

Analisis sentimen juga dapat digunakan untuk menganalisis sentimen terhadap politik, seperti penelitian yang dilakukan Akhmad Mazuki dan Arita Witanti yang meneliti tentang analisis sentiment masyarakat di *twitter*. Penelitian ini menggunakan 1200 yang diambil dari *twitter* yang kemudian akan dianalisis menggunakan metode *Naïve Bayes* dan *Support Vector Machine*. Hasil analisis yang dilakukan didapatkan hasil bahwa penggunaan *Naïve Bayes* ternyata lebih baik ketimbang *Support Vector Machine* dengan tingkat performasi untuk klasifikasi *tweet* 92,2%. Disimpulkan berdasarkan tingkat akurasi, *Naïve Bayes* lebih baik. Berdasarkan penelitian ini didapatkan kesimpulan bahwa dalam masa Covid-19 antusiasme masyarakat dapat diukur melalui analisis sentimen (Muzaki dan Witanti, 2021).

Melihat tingkat kepuasan konsumen terhadap suatu aplikasi dapat menggunakan anlisis sentiment, seperti penelitian yang membahas mengenai sentimen pengguna terhadap aplikasi Ruangguru. Analisis ini menggunakan dataset sebanyak 1629 dengan bentuk *text review* dalam rentang waktu 01 Maret sampai 30 Maret 2020. Menggunakan metode *Naïve Bayes, Random Forest dan Support Vector Machine* dataset akan diuji. Berdasarkan hasil Analisa dan hasil pengujian dari beberapa metode didapatkan hasil akurasi tertinggi pada metode *Random Forest* dengan jumlah akurasi 97,16%, *Support Vector Machine* dengan akurasi 96,01%, dan nilai akurasi terendah pada pengujian *Naïve Bayes* dengan nilai akurasi 94,16%. Kesimpulan dari penelitian ini adalah berdasarkan table hasil *preprocessing* dapat disimpulkan bahwa komentar positif lebih banyak ketimbang komentar negatif (Fitri dkk, 2020).

Hampir sama dengan penelitian mengenai sentimen terhadap aplikasi Ruangguru, analisis sentimen terhadap ulasan tekstual di *Google Play Store* yang diambil dari penelitian sebelumnya menggunakan metode *Suppport Vector Machine* dan *Naïve Bayes*. Data yang digunakan yaitu komentar yang terdapat pada *Google Play Store* dengan total 1818 komentar yang terdiri dari 606 komentar untuk

sentimen positif dan 606 komentar untuk sentiment negatif dan 606 komentar untuk sentimen crash. Berdasarkan hasil pengujian menggunakan 3-*fold cross-validation* didaptkan bahwa *Support Vector Machine* memiliki tingkat akurasi yang lebih besar yaitu 81,46% dibandingkan metode *Naïve Bayes* yang mendapatkan akurasi sebesar 75,41%. Kesimpulan dari penelitian ini adalah hasil akurasi dari *Support Vector Machine* memiliki nilai lebih tinggi jika dibandingkan dengan *Naïve Bayes* untuk mengklasifikasikan ulasan tekstual berbahasa Indonesia di *Google Play Store* (Ilmawan dan Mude, 2020)

Feminis atau feminisme membahas mengenai Isu kesetaraan gender yang mecoba untuk menyetarakan peran laki-laki dan perempuan. Feminisme pertama kali muncul di *Prancis* karena tradisi turun menurun yang seringkali merendahkan kaum Wanita. Gerakan ini menuntut kesamaan dan keadilan secara hak dengan pria dalam bidang sosial, budaya, pergerakan politik, ekonomi dan filosofi moral. Meski berbicara mengenai perjuangan akan hak perempuan. Feminisme tak lepas dari kontra dan kritik mengenai perkembangannya, seperti perbedaan pendapat tokoh feminis dengan kajian Islam. Sebuah pemikiran dari *Amina Wadud* tokoh feminis yang menafsirkan ayat al-Qur'an melalui kacamata feminis beranggapan bahwa al-Qur'an yang ditafsirkan oleh ulama klasik mengusung bias patriarki (Dzakiyyah, 2022). Feminisme telah menghasilkan banyak pencapaian emansipatoris yang luar biasa. Berhasil mengubah hukum dan praktik kelembagaan dalam standar diskriminasi gender, pekerjaan dan partisipasi perempuan di ruang publik (*Alexeyeff*, 2020).

Melihat besarnya pertumbuhan feminisme yang semakin marak disuarakan terlebih di Indonesia, penulis melakukan penelitian mengenai sentimen terhadap opini publik mengenai feminisme dengan judul "**Analisis Sentimen Terhadap Opini Feminisme Menggunakan Metode *Support Vector machine* dan *Naïve Bayes*".**

1.2 Perumusan Masalah

Rumusan masalah dirumuskan berdasarkan uraian latar belakang dalam penelitian sebagai acuan peneliti dalam mencari jawaban pertanyaan yang dikemukakan. Penelitian ini ditetapkan beberapa rumusan masalah yang dapat dirumuskan sebagai berikut:

1. Bagaimana Menerapkan *Naïve Bayes* dan *Support Vector Machine* dalam melakukan analisis sentimen terhadap opini feminisme?
2. Bagaimana perbandingan antara metode *Naïve Bayes* dan *Support Vector Machine* dalam melakukan analisis sentimen terhadap opini feminisme?
3. Bagaimana menguji kedua metode dalam melakukan analisis sentiment terhadap opini feminisme?

1.3 Batasan Masalah

Batasan masalah berguna dalam membantu mengidentifikasi masalah, membatasi ruang lingkup serta memberikan gambaran terkait hal yang akan diteliti supaya penelitian lebih terarah. Berdasarkan perumusan masalah di atas diberikan batasan ruang lingkup penelitian sebagai berikut:

1. Metode yang digunakan dalam penelitian ini adalah *Naïve Bayes* dan *Support Vector Machine*
2. Data yang digunakan merupakan data yang diambil dari *platform Twitter*
3. Data yang akan digunakan dalam penelitian ini merupakan data dari *tweet* yang di posting pada sosial media *twitter* dengan hashtag sebagai berikut : #feminisme #feminism
4. Bahasa yang digunakan menggunakan Bahasa Indonesia

1.4 Tujuan Penelitian

Hasil akhir dalam suatu penelitian merupakan tujuaan dari penelitian. Diperlukan suatu perencanaan dan pelaksanaan penelitian dengan matang supaya tujuan penelitian dapat dicapai. Penelitian ini ada beberapa tujuan yang hendak dicapai dan diringkas menjadi beberapa poin, yaitu:

1. Menerapkan *Naïve Bayes* dan *Support Vector Machine* dalam melakukan analisis sentimen terhadap opini feminisme
2. Perbandingan antara metode *Naïve Bayes* dan *Support Vector Machine* dalam melakukan analisis sentimen terhadap opini feminisme
3. Menguji kedua metode dalam melakukan analisis sentiment terhadap opini feminisme

1.5 Manfaat Penelitian

Suatu penelitian dilakukan dengan harapan dapat memeberikan manfaat baik bagi kepentingan peneliti, organisasi, maupun ilmu pengetahuan bagi peneliti sesudahnya. Manfaat penelitian ini merupakan jawaban dari tujuan penelitian yang dilakukan. Berikut manfaat penelitian yang akan didapatkan dalam penelitian ini sebagai berikut :

1. Manfaat bagi peneliti adalah dapat mengimplementasikan ilmu yang di dapat saat mengikuti perkuliahan.
2. Manfaat bagi pembaca diharapkan dapat memeberikan gambaran bagaimana tingkat perkembangan feminisme, terutama di Indonesia di mana isu mengenai feminisme dan bagiannya terus berkembang.

1.6 Sistematika Penulisan

Pada suatu penelitian dilakakukan pendokumentasian ditiap tahapan dalam bentuk tertulis. Sistematika penulisan meruapkan kerangka penelitian agar lebih terstruktur. Terdapat enam BAB untuk mekanisme penelitian ini dengan gambaran umum sebagai berikut :

BAB I : PENDAHULUAN

Bab I berisi pendahuluan dan latar belakag masalah dalam penelitian. Adapun yang akan dibahas pada bab ini adalah 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 mana nantinya akan digunakan sebagai landasan dasar dalam melakukan pemecahan masalah pada penelitian yang dilakukan.

BAB III : METODOLOGI PENELITIAN

Bab ini berisikan tentang langkah atau kerangka kerja dalam melakukan analisis sentiment dengan menggunakan metpde *Naïve Bayes* dan *Support Vector Machine*.

BAB IV : ANALISA DAN HASIL

Bab ini memuat tentang analisa data, pengolahan data twitter mengenai opini terhadap feminisme. Selain itu bab ini juga memuat hasil dari penggunaan metode *Naïve Bayes* dan SVM

BAB V : IMPLEMENTASI DAN PENGUJIAN

Pada bab ini berisi penerapan metode *Naïve Bayes* dan SVM menggunakan Bahasa pemograman phyton dan Rapid Miner 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.