

## BAB I

### PENDAHULUAN

#### 1.1. Latar Belakang Masalah

Bahan Bakar Minyak (BBM) sudah menjadi kebutuhan pokok bagi masyarakat Indonesia. Peran BBM saat ini sama pentingnya dengan energi listrik karena banyak sektor usaha, rumah tangga, industri, dan transportasi yang sangat bergantung pada keberadaan BBM. Penggunaan Bahan Bakar Minyak (BBM) di Indonesia terus meningkat seiring dengan dengan penambahan jumlah usaha yang membutuhkan pasokan bahan bakar ini. Pemerintah menyebut peningkatan kebutuhan energi BBM di Indonesia mencapai 8 persen per tahun. Masyarakat dapat memperoleh Bahan Bakar Minyak (BBM) ini di Stasiun Pengisian Bahan-bakar Umum (SPBU). Jenis bahan bakar yang tersedia di SPBU biasanya adalah bensin, solar, dan minyak tanah [1].

Ketergantungan terhadap penggunaan bensin saat ini mengakibatkan tingkat penjualan selalu meningkat sedangkan pasokan stok tetap sama. Pemerintah bersama Pertamina sudah memberi harga subsidi, namun penyebaran subsidi tersebut tidak tepat sasaran. Pemerintah melalui Badan Usaha Milik Negara (BUMN) bekerja sama dengan Pertamina untuk mencari solusi yang dapat mengatasi masalah penyebaran subsidi tersebut. My Pertamina adalah salah satu solusi yang diberikan pemerintah berupa aplikasi *e-payment* dari PT Pertamina (Persero) yang dapat digunakan dengan mudah untuk seluruh pelanggan Pertamina. My Pertamina berfungsi sebagai *cashless payment* yang memiliki sembilan sistem pembayaran non tunai, sarana pelanggan untuk mendapatkan poin dan *reward* serta fitur pemakaian *e-voucher*. Pertamina lebih mengutamakan promosi - promosi yang berhubungan dengan bahan bakar Pertamax. Hal ini sesuai dengan tujuan Pertamina membuat masyarakat untuk lebih memilih atau mengambil keputusan untuk menggunakan bahan bakar Pertamax [2].

Baru-baru ini, dikeluarkan aturan baru berkaitan dengan penggunaan aplikasi My Pertamina dalam transaksi Bahan Bakar Minyak (BBM) di SPBU. Tentu saja hal ini mendapat reaksi beragam dari banyak pihak. Banyak orang mulai mengomentari kebijakan ini di beberapa jejaring sosial, termasuk *Twitter*. *Twitter* adalah salah satu media sosial yang bertipe *microblogging* sebagai layanan interaksinya. *Twitter* menjadi salah satu layanan media sosial yang paling terkenal di dunia dengan lebih dari 200 juta pengguna aktif dan lebih dari 10,6 milyar *tweet* yang telah dihasilkan [3]. Pengguna *twitter* berbagi banyak komentar dan pendapat tentang penetapan aturan baru pada penggunaan aplikasi My Pertamina ini. Komentar yang mereka berikan dapat berupa komentar setuju (*pro*) maupun tidak setuju (*kontra*) terhadap kebijakan ini.

Tulisan dari pengguna *twitter* ini dapat digunakan sebagai bahan pendukung keputusan, evaluasi dan atau pengembangan dari pihak pembuat keputusan. Untuk mengetahui mayoritas pendapat dari pengguna dapat dilakukan proses analisis sentimen terhadap komentar dari para pengguna terhadap kebijakan baru terkait penggunaan aplikasi My Pertamina. Analisis sentimen merupakan salah satu penelitian yang tergolong dalam *Natural Language Processing* (NLP), komputasi linguistik, dan *text mining*. Analisis sentimen atau biasa dikenal dengan *opinion mining* merupakan salah satu cabang penelitian dari *text mining* yang bertujuan untuk menentukan persepsi atau subjektivitas publik (khalayak) terhadap suatu topik pembahasan, kejadian, ataupun permasalahan [11].

Data - data yang diperoleh akan dianalisis sehingga mendapatkan hasil berupa sentimen masyarakat terhadap kebijakan penggunaan aplikasi My Pertamina. Dalam penelitian ini, peneliti akan melakukan perbandingan dari hasil analisis 2 algoritma, yaitu *Support Vector Machine* (SVM) dan *Naïve Bayes Classifier*. *Support Vector Machine* (SVM) adalah teknik yang relatif baru untuk membuat prediksi, baik dalam konteks klasifikasi atau regresi yang sangat populer belakangan ini. Penggunaan algoritma ini dalam menyelesaikan masalah sehari-hari sudah meluas, baik di kalangan

akademisi maupun praktisi. Terbukti dalam banyak implementasi, SVM secara konsisten mencapai solusi yang sama untuk setiap proses, memberikan hasil yang lebih baik dari ANN secara keseluruhan proses [21]. Algoritma *Naive Bayes* dikenal sebagai *Teorema Bayes* karena memprediksi arah kejadian di masa depan berdasarkan apa yang terjadi di masa lalu. Salah satu hal yang membuat *Naive Bayes Classifier* ini terkenal adalah asumsi yang sangat kuat (naif) pada masing-masing kondisi atau kejadian. Dibandingkan dengan model pengklasifikasi lainnya, *Naive Bayes Classifier* berperforma sangat baik [24].

Terdapat beberapa penelitian menggunakan algoritma *Support Vector Machine* (SVM) dan atau *Naive Bayes Classifier* yang sudah dilakukan. Pertama, Maulana Rizky Hidayat (2022) telah melakukan penelitian dengan judul “*Perbandingan Akurasi dan Waktu Proses Algoritma Naive Bayes dengan Algoritma Support Vector Machine dalam Analisis Sentimen Twitter Aplikasi Peduli Lindungi*”. Peneliti menggunakan dataset sebanyak 3588 dataset kemudian dengan pengujian menggunakan teknik *split 8020* serta *K-Fold Cross*. Hasil dari penelitian ini adalah tingkat akurasi dari kedua algoritma yang digunakan sama sehingga keduanya efektif untuk digunakan dalam proses analisis sentimen. Selain berdasarkan tingkat akurasi, peneliti juga membandingkan berdasarkan waktu proses pelatihan (*training*) dengan hasil yang menunjukkan bahwa algoritma *Naive Bayes Classifier* jauh lebih cepat jika dibandingkan dengan algoritma SVM karena proses cara kerja dari kedua algoritma tersebut berbeda [3].

Kedua, Deni Gunawan, dkk. (2020) telah melakukan penelitian dengan judul “*Komparasi Algoritma Support Vector Machine Dan Naive Bayes Dengan Algoritma Genetika Pada Analisis Sentimen Calon Gubernur Jabar 2018-2023*”. Pada penelitian ini, menggunakan data sebanyak 2643 data *tweet* untuk keempat kandidat calon gubernur Jabar 2018-2023 yang dikelompokkan menjadi 2 tabel yaitu positif dan negatif. Hasil dari penelitian yang sudah dilakukan ini adalah rata-rata akurasi pada algoritma *Support Vector Machine* (SVM) adalah 92,61% dengan AUC 0.950

sedangkan pada algoritma *Naive Bayes* menghasilkan rata-rata akurasi 93,29% dengan AUC 0,525, Algoritma *Support Vector Machine* berbasis *Genetic Algorithm* menghasilkan rata-rata akurasi 93,03% dengan AUC 0,869 dan Algoritma *Naive Bayes* berbasis *Genetic Algorithm* menghasilkan rata-rata akurasi 92,85% dengan AUC 0,543. Dengan demikian, dapat disimpulkan bahwa hasil komparasi dari penelitian ini adalah algoritma *Support Vector Machine* (SVM) memiliki performa lebih baik jika dibandingkan dengan algoritma *Naive Bayes* [4].

Ketiga, Dimas Pramana Indratama (2022) melakukan penelitian dengan judul "*Analisis Sentimen Terhadap Ulasan Pengguna Aplikasi Peduli Lingkungan Menggunakan Algoritma Support Vector Machine*". Pada penelitian ini, peneliti melakukan penelitian terhadap data hasil crawling dari cuitan (*tweets*) di aplikasi *twitter* menggunakan *API twitter*. Pengambilan data dilakukan pada data tweet yang ditulis oleh pengguna dalam rentang waktu bulan Februari hingga Mei tahun 2022. Hasil dari penelitian ini berupa persentase untuk masing - masing kelas sentimen dari dataset yang diuji. Adapun hasil untuk kelas sentimen positif sebesar 21,43% dan untuk kelas sentimen negatif sebesar 78,57%. Selain itu, untuk hasil rata-rata akurasi dengan pengujian menggunakan *10-Fold Cross Validation* didapatkan sebesar 85%. Berdasarkan hasil persentase dari penelitian yang sudah dilakukan, menunjukkan bahwa algoritma *Support Vector Machine* dapat melakukan analisis sentimen pada data *twitter* berbahasa Indonesia dengan baik [5].

Berdasarkan penjelasan dan penelitian yang sudah dilakukan, kedua algoritma baik *Support Vector Machine* (SVM) maupun *Naive Bayes Classifier* memiliki hasil tingkat akurasi yang berbeda - beda dalam setiap kasus yang dianalisis. Selain itu, kedua algoritma tersebut juga memiliki kelebihan dan kekurangan masing-masing dalam sistem kerja analisis sentimen. Oleh karena itu, penelitian ini bertujuan mengetahui perbandingan sistem kerja algoritma *Support Vector Machine* (SVM) dan

*Naïve Bayes Classifier* dalam analisis sentimen pengguna media sosial *twitter* terhadap penerapan aplikasi MyPertamina.

### 1.2. Rumusan Masalah

Berdasarkan latar belakang yang telah dipaparkan, maka diperoleh rumusan masalah untuk penelitian ini, sebagai berikut :

1. Bagaimana implementasi algoritma *Support Vector Machine* (SVM) dan *Naive-Bayes Classifier* dalam analisis sentimen pengguna *twitter* terhadap penerapan aplikasi My Pertamina?
2. Bagaimana hasil analisis sentimen dari penerapan aplikasi My Pertamina pada media sosial *twitter*?
3. Apa perbedaan dalam implementasi algoritma *Support Vector Machine* (SVM) dan *Naive-Bayes Classifier* dalam analisis sentimen pengguna *twitter* terhadap penerapan aplikasi MyPertamina?

### 1.3. Batasan Penelitian

Batasan masalah dari penelitian ini adalah:

1. Algoritma yang digunakan dalam penelitian ini adalah *Support Vector Machine* (SVM) dan *Naive-Bayes Classifier* dengan menerapkan *tools* yang sudah terdapat dalam *library python*.
2. Data yang digunakan dalam penelitian ini adalah hasil dari *crawling* data media sosial *twitter* dengan kata kunci "Aplikasi My Pertamina"
3. Pengambilan data dilakukan pada cuitan (*tweet*) yang ditulis dalam rentang waktu mulai tanggal 15 Juni 2022 hingga 25 Juli 2022.
4. Jumlah dataset yang digunakan dalam penelitian ini adalah 2500 data
5. Proses pengambilan (*crawling*) data dari media sosial *twitter* menggunakan *library python sncrape* agar proses *crawling* lebih mudah.
6. Dataset hasil *crawling* akan melalui proses *pre-processing* kemudian akan diklasifikasikan ke dalam 2 kelas sentimen yaitu sentimen positif dan sentimen negatif

#### 1.4. Tujuan Penelitian

Maksud dan tujuan yang ingin dicapai dari penelitian ini adalah:

1. Mengetahui perbandingan algoritma *Support Vector Machine* (SVM) dan *Naive-Bayes Classifier* dalam analisis sentimen pengguna *twitter* terhadap penerapan aplikasi My Pertamina
2. Mengetahui hasil analisis sentimen pengguna *twitter* terhadap penerapan aplikasi My Pertamina dengan algoritma *Support Vector Machine* (SVM) dan *Naive-Bayes Classifier*
3. Mengetahui reaksi masyarakat terhadap penerapan kebijakan baru dalam penggunaan aplikasi My Pertamina.

#### 1.5. Manfaat Penelitian

Manfaat yang dapat diambil dari penelitian ini adalah:

1. Mengetahui perbandingan dari dua algoritma yang digunakan yaitu *Support Vector Machine* (SVM) dan *Naive-Bayes Classifier* dalam analisis sentimen pengguna *twitter* terhadap penerapan aplikasi My Pertamina
2. Hasil analisis sentimen dari topik yang diteliti dapat digunakan sebagai acuan pengambilan keputusan terhadap kebijakan penggunaan aplikasi My Pertamina.
3. Sebagai referensi untuk penelitian selanjutnya baik dari segi algoritma yang digunakan maupun topik yang diteliti.

#### 1.6. Sistematika Penelitian

Rincian sistematika penulisan dari penelitian ini adalah 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, dasar-dasar teori yang digunakan dalam penelitian, dan ringkasan dari hasil penelitian sejenis yang sudah dilakukan sebelumnya sebagai referensi.

BAB III METODE PENELITIAN, didalamnya terdapat objek yang akan menjadi sasaran pada penelitian, langkah-langkah / alur prosedur penelitian, dan alat serta bahan yang akan digunakan dalam penelitian.

BAB IV HASIL DAN PEMBAHASAN, bab ini berisi proses penelitian yang dilakukan, implementasi algoritma pada dataset yang digunakan sebagai objek/sasaran penelitian dan pembahasan tentang hasil dari penelitian yang sudah dilakukan.

BAB V PENUTUP, bab ini berisi kesimpulan yang dapat peneliti rangkum selama proses penelitian serta saran untuk penelitian selanjutnya berdasarkan masalah dan kekurangan dari penelitian ini.

