

BAB I

PENDAHULUAN

1.1 Latar Belakang

Pengaruh perkembangan teknologi saat ini menghasilkan berbagai jenis media sosial, antara lain *instagram*, *twitter*, *facebook*, *youtube*, dan banyak lagi, dari beberapa media sosial tersebut *youtube* menjadi salah satu media berbagi informasi video yang terbesar, *youtube* memiliki indikator penilaian dengan *rating like* atau *dislike* (Munthe et al., 2021). Hadirnya *youtube* menjadi inovasi dalam penyebaran informasi, dan kepopuleran *youtube* saat ini dimanfaatkan banyak pihak sebagai wadah untuk *branding* dan publikasi salah satunya seperti *Beauty Vlogger*.

Beauty Vlogger merupakan seseorang yang menghasilkan video berisikan tutorial teknik *make up* serta *review* dan juga saran produk (Mariezka et al., 2018). Orang-orang yang menonton video *Beauty Vlogger* pada salah satu saluran *youtube*, mereka dapat memberikan pendapat mereka terhadap video dari *review* bedak dan *skincare* yang di sampaikan *vlogger* tersebut melalui kolom komentar jika pemilik video tidak menonaktifkan fitur komentar. Namun tidak setiap komentar dari pengguna *youtube* memiliki arti yang sesuai, oleh karena itu analisis sentimen hal yang tepat untuk mengumpulkan pendapat dari masyarakat pengguna *youtube* yang meninggalkan komentar pada saluran *youtube Beauty Vlogger* tersebut.

Analisis sentimen adalah suatu cara untuk menentukan suatu opini atau tanggapan positif negatif dan netral dengan mengklasifikasikan sebuah dokumen teks berdasarkan pendapat dan sudut pandang orang-orang, hal ini biasanya dilakukan dengan memahami, mengelola dan mengekstrak data secara otomatis dari dokumen teks tersebut, untuk memperoleh informasi sentimen terhadap sesuatu kejadian yang sedang diteliti (Sianturi & Hendriani, 2021). Maka dirancanglah sebuah sistem analisis sentimen komentar pada *youtube* dengan hasil yang diperoleh berupa klasifikasi teks kedalam kategori sentimen yang berbeda (positif, atau negatif).

Machine learning yang digunakan untuk tugas klasifikasi dan regresi. *Support Vector Machine* (SVM) bekerja dengan membangun sebuah model yang dapat memisahkan dua kelas menggunakan margin maksimum di antara kelas-kelas tersebut. dengan memanfaatkan teknik *support vector* yang keakuratannya tinggi Banyak penelitian yang memanfaatkan *Support Vector Machine* sebagai algoritma klasifikasi yang menunjukkan kinerjanya cukup baik (Monika & Furqon, 2018).

Penelitian sebelumnya dilakukan oleh Reyhan Achmad Rizal, Imron Sanjaya Girsang, Sidik Apriyadi Prasetyo tentang tingkat rata-rata true detection sebesar 90% dan false detection sebesar 10% pada klasifikasi wajah menggunakan *Support Vector Machine* (SVM) dicapai karena beberapa faktor penting dalam metode SVM yang diimplementasikan. Metode SVM dikenal memiliki kemampuan tinggi dalam menentukan batas antara dua kelas (*hyperplane*) yang memaksimalkan margin antara titik data dari masing-masing kelas. Dalam penelitian ini, proses klasifikasi wajah dilakukan dengan berbagai sudut kemiringan (dari -90° hingga $+90^\circ$), dan SVM mampu mengakomodasi variasi posisi ini secara efektif dengan mempertahankan tingkat akurasi yang tinggi (Rizal et al., 2019).

Fitur utama dari SVM adalah kemampuannya untuk menghasilkan model yang kuat tanpa memerlukan pengetahuan tambahan atau parameter yang kompleks, menjadikannya sangat cocok untuk data berdimensi tinggi seperti pengenalan wajah. *True detection* tinggi yang tercapai merupakan indikasi bahwa model SVM mampu mengenali pola wajah secara konsisten, bahkan dengan perubahan kecil dalam posisi wajah. Sedangkan *false detection* 10% yang relatif rendah menunjukkan bahwa kesalahan dalam klasifikasi tetap terjaga pada tingkat minimal, yang merupakan hasil dari seleksi fitur yang optimal dan margin maksimal yang ditetapkan SVM, sehingga mampu mengurangi kesalahan klasifikasi secara signifikan.

Pada penelitian sebelumnya dengan judul “Analisis dan Penerapan Algoritma *Support Vector Machine* (SVM) dalam Data Mining untuk Menunjang Strategi Promosi” juga dilakukan oleh Agung Handayanto, Khoirya Latifa, Nugroho Dwi Saputro, Rahmat Robi Waliyansyah. Tingkat akurasi sebesar 73,6% dalam penelitian mereka tentang penerimaan mahasiswa baru menggunakan *Support*

Vector Machine (SVM) pada Universitas PGRI Semarang dapat dijelaskan sebagai hasil dari beberapa faktor penting dalam pengolahan data dan konfigurasi model. Dalam penelitian ini, SVM digunakan untuk menganalisis data rekapitulasi penerimaan mahasiswa baru dari tahun 2014 hingga 2018, mencakup 6.371 baris data dengan berbagai parameter seperti jenis kelamin, agama, asal sekolah, jurusan, dan asal wilayah (Handayanto et., 2019).

SVM dipilih untuk memisahkan kelas dengan *hyperplane* yang paling optimal, tetapi kompleksitas data dari berbagai latar belakang calon mahasiswa menimbulkan kesulitan untuk meningkatkan akurasi lebih tinggi. Kombinasi faktor-faktor, seperti keberagaman fitur input yang cukup luas, penggunaan kernel sigmoid, dan proses random sampling dengan banyak perulangan, mempengaruhi hasil. Dengan adanya confusion matrix, diperoleh distribusi yang mencerminkan tingkat kesulitan dalam memprediksi calon mahasiswa yang akan melakukan daftar ulang, yang pada akhirnya menghasilkan akurasi sebesar 73,6%.

Penelitian terkait juga dilakukan oleh Renaldi AL Anshari, Syariful Alam, Moch Hafid T dengan judul “Komparasi *Payment* Digital Untuk Analisis Sentimen Berdasarkan Ulasan di *Google Playstore* Menggunakan Metode *Support Vector Machine*. Hasil penelitian ini menunjukkan efektivitas SVM dengan tingkat akurasi tinggi mencapai 92% pada aplikasi Dana dalam menangkap pola sentimen yang relevan dari data ulasan pengguna. *Precision* sebesar 92% dan *recall* 100% pada aplikasi Dana menunjukkan bahwa model memiliki kemampuan tidak hanya mengidentifikasi ulasan positif secara konsisten, tetapi juga hampir sepenuhnya menghindari kesalahan dalam mendeteksi ulasan negatif sebagai positif. Dalam aplikasi ini, SVM memisahkan kelas dengan *hyperplane* yang baik, menghasilkan keseimbangan yang optimal dalam prediksi berdasarkan sentimen ulasan di *Google Play Store*

Untuk aplikasi OVO, akurasi mencapai 90% dengan *precision* 88% dan *recall* 98%, menunjukkan performa yang mendekati Dana namun dengan sedikit penurunan pada *precision*. Hal ini mungkin karena ulasan di OVO memiliki variasi yang lebih besar atau kompleksitas pola yang mempengaruhi kemampuan SVM dalam mempertahankan *precision* setinggi Dana. Namun, tingginya nilai *recall*

menunjukkan bahwa model cukup baik dalam mengenali ulasan negatif atau positif yang benar secara luas, meskipun precision sedikit menurun (Anshari et al., 2023).

Pemilihan metode *Support Vector Machine* (SVM) untuk penelitian ini dapat didasarkan pada keunggulan SVM dalam menangani data klasifikasi, terutama pada kasus yang melibatkan identifikasi pola kompleks dan pemilahan sentimen yang cukup sulit diidentifikasi dalam data berdimensi tinggi, seperti pada ulasan pengguna di platform digital. Kemampuan SVM dalam klasifikasi sentimen SVM terbukti efektif untuk memisahkan kelas sentimen yang kompleks pada data berbasis teks, yang mirip dengan tujuan penelitian ini dalam mengklasifikasikan sentimen dalam ulasan atau data lainnya. SVM terbukti dapat bekerja efektif pada data dengan variasi posisi (seperti variasi wajah dan data multivariabel). Ini menunjukkan bahwa dalam konteks penelitian ini, SVM mampu menangani variabilitas dalam ekspresi sentimen atau respons pengguna terhadap layanan yang berbeda, sehingga relevan digunakan. SVM tidak hanya memberikan akurasi yang kuat tetapi juga fleksibilitas dalam berbagai aplikasi yang memiliki pola data kompleks, menjadikannya pilihan yang relevan untuk penelitian ini dalam pengklasifikasian data atau analisis sentimen yang memerlukan tingkat presisi tinggi.

Penelitian ini bertujuan untuk memahami tanggapan masyarakat pengguna *youtube* terhadap saluran *Beauty Vlogger* berbahasa indonesia, serta mengetahui akurasi metode dari *Support Vector Machine* dalam melakukan analisis sentimen. Berdasarkan latar belakang tersebut penulis dapat mengangkat judul tugas akhir, “Analisis Sentimen Komentar Pada Saluran *Youtube Beauty Vlogger* Berbahasa Indonesia Menggunakan Metode *Support Vector Machine*”.

1.2 Rumusan Masalah

Dapat dilihat pada latar belakang di atas, maka dapat disimpulkan beberapa permasalahan yaitu:

1. Bagaimana penerapan *Support Vector Machine* untuk analisis sentimen komentar produk bedak dan *skincare* pada saluran *youtube Beauty Vlogger* berbahasa indonesia.

2. Seberapa baik performa algoritma *Support Vector Machine* dalam pengklasifikasian analisis sentimen *youtube* produk bedak dan *skincare* pada saluran *youtube Beauty Vlogger* berbahasa indonesia.

1.3 Tujuan Penelitian

Berikut adalah tujuan dari penelitian yang akan dilakukan:

1. Merancang sistem untuk menganalisis sentimen komentar produk bedak dan *skincare* pada saluran *youtube Beauty Vlogger* berbahasa indonesia menggunakan metode *Support Vector Machine*
2. Mengetahui bagaimana penerapan algoritma *Support Vector Machine* dalam menganalisis sentimen komentar produk bedak dan *skincare* pada saluran *youtube Beauty Vlogger* berbahasa indonesia.

1.4 Manfaat Penelitian

Manfaat yang diharapkan dari penelitian tugas akhir ini adalah sebagai berikut:

1. Dapat membangun sebuah sistem untuk menganalisis sentimen produk bedak dan *skincare* pada saluran *youtube Beauty Vlogger* berbahasa indonesia menggunakan metode *Support Vector Machine*
2. Dapat mengimplementasikan metode *Support Vector Machine* untuk analisis sentimen produk bedak dan *skincare* pada saluran *youtube Beauty Vlogger* berbahasa indonesia.

1.5 Batasan Masalah

Batasan masalah pada penelitian yang akan dilakukan yaitu:

1. Metode yang digunakan adalah *Support Vector Machine*.
2. Penelitian ini hanya dilakukan pada akun *youtube Beauty Vlogger*.
3. Jumlah data yang diambil sebanyak 1500 komentar hanya pada akun *youtube Beauty Vlogger*.
4. Komentar yang dianalisis hanya komentar berjenis teks.
5. Hasil dari sistem ini adalah dapat mendeteksi sebuah komentar yang hanya berfokus pada positif negatif dan netral.