

BAB 1

PENDAHULUAN

1.1 Latar Belakang

Meskipun wanita mendambakan penampilan menarik dan wajah yang enak dipandang, kecantikan sejati tidak hanya berasal dari pakaian dan aksesoris. Kunci utama terletak pada kesehatan kulit[1]. Kecanggihan teknologi dan kemudahan akses internet mendorong pergeseran tren belanja kosmetik. Masyarakat kini lebih memilih berbelanja kosmetik secara online. [2]. Kejelian konsumen sangatlah penting dalam memilih produk kosmetik di pasaran yang beragam, karena tidak semua produk memiliki kualitas yang terjamin dan aman untuk digunakan [3]. Salah satu produk yang menarik perhatian di media sosial adalah *Skintific*. *Skintific* dikenal dengan produk perawatan kulit yang inovatif dan efektif, serta banyak digunakan oleh konsumen dari berbagai kalangan.

saat ini, konsumen sering melakukan riset kosmetik membeli produk. Mereka akan membandingkan harga dan berdiskusikan dengan konsumen lain tentang pendapat mereka, atau membaca ulasan produk[4]. Informasi dari ulasan produk (5,06) lebih dipercaya dan mampu menarik minat konsumen dibandingkan informasi marketing (4,36) yang bersifat subjektif [5]. Membaca semua review memang ideal, tapi waktunya bisa lama. Memilih hanya beberapa review memang praktis, tapi risikonya adalah penilaian yang tidak objektif. Analisis sentimen memungkinkan proses review yang efisien dan objektif, sehingga menghemat waktu dan menghasilkan penilaian yang lebih akurat [6]

Analisis sentimen adalah proses otomatis untuk memahami dan menganalisis pendapat atau opini dalam teks yang tidak terstruktur, seperti review, dengan mengekstrak, mengolah, dan memahaminya[7]. Analisis sentimen adalah proses otomatis untuk memahami dan menganalisis pendapat atau opini dalam teks yang tidak terstruktur, seperti review, dengan mengekstrak, mengolah, dan memahaminya[8]. Klasifikasi analisis sentimen dapat dilakukan dengan algoritma klasifikasi seperti *Decision Tree*, *Naive Bayes*, *Bayes Networks*, *Support Vector Machine (SVM)*, *Random Forest*, dan lain-lain[9].

Machine learning adalah kecerdasan buatan yang memungkinkan komputer belajar dan meningkatkan performanya sendiri tanpa diprogram secara eksplisit, berdasarkan data yang diberikan [10]. Pendekatan machine learning, khususnya algoritma Random Forest, menawarkan solusi efektif untuk menganalisis data sentimen yang berlimpah dari media sosial. *Random Forest*, sebuah metode klasifikasi yang efektif, digunakan untuk mengidentifikasi sentimen (positif, negatif, atau netral) dalam komentar pelanggan secara otomatis [11]. Breiman memperkenalkan algoritma *Random Forest* yang memiliki beberapa keunggulan, yaitu menghasilkan error yang relatif rendah, memiliki performa yang baik dalam klasifikasi, mampu menangani data training berdimensi besar secara efisien, dan merupakan metode yang efektif untuk mengestimasi dataset yang mengandung missing data [12].

Penelitian ini penting dilakukan mengingat semakin tingginya ketergantungan masyarakat pada ulasan dan opini di media sosial dalam pengambilan keputusan pembelian. Dengan memahami sentimen publik, perusahaan dapat merumuskan strategi pemasaran yang lebih tepat sasaran, meningkatkan kepuasan pelanggan, dan pada akhirnya meningkatkan penjualan produk kecantikan mereka. Dengan demikian, judul penelitian ini berjudul “STUDI SENTIMEN PUBLIK TERHADAP PRODUK KECANTIKAN DENGAN PENDEKATAN MACHINE LEARNING : ANALISIS RANDOM FOREST PADA DATA MEDIA SOSIAL”.

1.2 Rumusan Masalah

Penelitian ini berfokus pada bagaimana memperoleh dan mengolah data media sosial yang mencerminkan pendapat dan persepsi pengguna tentang produk kecantikan, bagaimana membangun model *Random Forest* yang optimal untuk mengklasifikasikan sentimen pengguna terhadap produk kecantikan dan bagaimana mengukur dan mengevaluasi kinerja model *Random Forest* dalam menganalisis sentimen publik terhadap produk kecantikan.

1.3 Tujuan Penelitian

Penelitian ini bertujuan untuk Menganalisis dan mengidentifikasi sentimen publik secara efektif terhadap produk kecantikan, khususnya produk *Skintific* berdasarkan data yang dikumpulkan dari berbagai *platform* media sosial, Mengukur kinerja model *Random Forest* dalam menganalisis sentimen publik terhadap produk kecantikan *Skintific*, termasuk akurasi klasifikasi dan kehandalan prediksi dan Membangun model klasifikasi menggunakan pendekatan *machine learning*, khususnya analisis *Random Forest*, untuk mengklasifikasikan sentimen pengguna baik positif, negatif, maupun netral yang terkait dengan produk kecantikan *Skintific*.

1.4 Manfaat Penelitian

Beberapa manfaat yang diharapkan dari penelitian ini adalah:

1. Dengan memahami sentimen publik terhadap produk kecantikan, perusahaan dapat meningkatkan strategi pemasaran mereka.
2. Penelitian ini memperkaya literatur di bidang analisis sentimen dan penerapan *machine learning*, khususnya algoritma *Random Forest*, dalam konteks produk kecantikan.
3. Penelitian ini menunjukkan aplikasi praktis dari algoritma *Random Forest* dalam analisis sentimen, yang dapat mendorong penggunaan lebih luas dari teknik *machine learning* di berbagai industri.

1.5 Batasan Masalah

Ruang lingkup penelitian ini ialah sebagai berikut:

1. Penelitian ini akan terbatas pada data yang tersedia dari *platform* media sosial tertentu saja. Keterbatasan ini dapat mempengaruhi representasi sentimen publik secara keseluruhan, terutama jika data yang dikumpulkan tidak mencakup berbagai macam sumber dan kelompok demografis.
2. Meskipun analisis *Random Forest* dapat memprediksi sentimen dengan tingkat akurasi yang tinggi, tetapi masih mungkin terdapat kesalahan dalam klasifikasi, terutama dalam kasus sentimen netral atau ambigu.
3. Penelitian ini hanya berfokus pada produk skincare dari *Skintific*. Produk kecantikan lainnya dari merek yang berbeda tidak akan dianalisis.

1.6 Keterbaruan

1. Penelitian yang dilakukan oleh Green Arther Sandag (2020) tentang Pemanfaatan Algoritma Random Forest untuk Memprediksi Rating Aplikasi di App Store: Sebuah studi menggunakan algoritma Random Forest untuk memprediksi rating aplikasi di App Store. Hasilnya menunjukkan akurasi sebesar 86.27%, recall 84.68%, presisi 84.64%, dan nilai RMSE 0.313 [13].
2. Penelitian yang dilakukan oleh Ibnu Afdhal, Rahmad Kurniawan, Iwan Iskandar, Roni Salambue, Elvia Budianita, dan Fadhilah Syafria (2020) tentang Analisis Sentimen Komentar YouTube tentang Islamofobia Menggunakan Algoritma Random Forest. Analisis menemukan 631 komentar positif dan 369 komentar negatif/islamofobia. Algoritma random forest mencapai akurasi 79% dalam eksperimen[14].

3. Penelitian yang dilakukan oleh Evita Fitri, Yuri Yuliani, Susy Rosyida dan Windu Gata tentang membandingkan kinerja tiga algoritma klasifikasi, yaitu Naive Bayes, Random Forest, dan Support Vector Machine, dalam menganalisis sentimen pengguna terhadap aplikasi Ruangguru. Hasil menunjukkan bahwa algoritma Random Forest menghasilkan akurasi tertinggi, dengan peningkatan 7.16% dibandingkan penelitian sebelumnya [15].
4. Penelitian yang dilakukan oleh Oktaviami Manullang Cahyo Prianto, dan Nisa Hanum Harani penggunaan analisis sentimen untuk memprediksi hasil pemilihan presiden, dengan membandingkan dua metode: Lexicon Based dan Random Forest. Hasil menunjukkan bahwa Random Forest menghasilkan performa yang lebih baik, terutama dalam mengidentifikasi sentimen negatif terhadap kandidat[16].