

BAB I

PENDAHULUAN

1.1. Latar Belakang

Di era digital sekarang kesadaran masyarakat terhadap perawatan kulit (*skincare*) semakin meningkat terutama pada kalangan perempuan (Ścieszko et al., 2021). Hal ini didorong oleh kemudahan akses informasi yang membuat masyarakat lebih cepat teredukasi tentang pentingnya menjaga kesehatan kulit. Media sosial seperti *TikTok* telah menjadi *platform* utama bagi *influencer* kecantikan untuk memberikan rekomendasi produk *skincare* yang dapat membantu mengatasi berbagai permasalahan kulit. Salah satu *influencer* populer di Indonesia, Tasya Farasya sering membahas produk dan kandungan *skincare* yang bermanfaat untuk berbagai jenis kulit (Mirza, 2024). Salah satu alasan mengapa Tasya Farasya memiliki pengaruh besar karena Tasya selalu memberikan ulasan jujur setelah mencoba produk *skincare* secara pribadi selama kurang lebih dua minggu (PUTRI, 2020). Pendekatan ini membuat banyak orang lebih percaya dan merasa yakin terhadap produk yang telah mendapatkan persetujuan dari Tasya. Selain itu, ada juga Syafik Amry atau yang dikenal dengan *username* sebagai Dosen *Skincare* pada aplikasi *TikTok*. Sebagai seorang *influencer* sekaligus ahli kecantikan, Syafik kerap membagikan pengetahuan mendalam tentang kandungan *skincare* dan cara memilih produk yang sesuai dengan jenis kulit. Hal ini membantu *audiens* membuat keputusan yang lebih cerdas dalam merawat kulit.

Namun meskipun informasi ini sangat melimpah, banyak perempuan masih kesulitan dalam memahami kandungan *skincare* serta bagaimana cara menggabungkannya secara tepat sesuai dengan jenis kulit dan permasalahan kulit mereka. Hal ini sering kali menyebabkan penggunaan produk yang tidak sesuai sehingga dapat memperburuk kondisi kulit atau memberikan hasil yang tidak optimal. Salah satu penyebab utama dari masalah ini adalah rendahnya literasi terhadap kandungan *skincare* di kalangan pengguna. Kandungan seperti niacinamide, hyaluronic acid, dan retinol sering disebut dalam ulasan produk. Tetapi tidak semua pengguna mengetahui manfaat, risiko, atau cara penggunaan

kandungan yang tepat untuk jenis kulit dan permasalahan kulit yang dihadapi. Selain itu, kombinasi bahan tertentu juga memerlukan pemahaman khusus untuk memastikan keamanan dan efektivitasnya. Hal ini diperkuat oleh hasil penelitian Pratiwi et al (2023) yang dilakukan pada remaja putri di SMPN 1 Awangpone. Sejumlah 60 dari total 88 responden mengaku tidak memiliki pengetahuan yang memadai tentang kandungan dan penggunaan produk *skincare*. Temuan serupa juga ditemukan dalam studi Hilmi et al (2022) di salah satu universitas di Karawang, Jawa Barat, yang menunjukkan bahwa 102 dari 392 responden memiliki tingkat pengetahuan yang kurang baik tentang *skincare*. Hasil dari kedua penelitian ini menunjukkan bahwa pemahaman mengenai kandungan *skincare* masih menjadi tantangan bagi banyak pengguna. Sebagian besar orang merasa perlu mencari informasi tambahan agar dapat menggunakan produk dengan tepat dan aman.

Fitur kolom komentar pada aplikasi *TikTok* kerap digunakan sebagai wadah konsultasi bagi perempuan yang masih belum paham dengan kandungan *skincare*. Dengan jumlah komentar ulasan pada *TikTok* yang besar dan tidak terstruktur, diperlukan cara untuk menganalisis informasi ini secara efektif. *Natural Language Processing* (NLP) menawarkan solusi inovatif untuk menganalisis komentar-komentar di *TikTok* secara otomatis. Dengan teknik NLP, teks yang tidak terstruktur dapat diubah menjadi data yang terorganisir (Ladas et al., 2023). Sehingga dapat mengekstraksi informasi penting seperti kandungan *skincare*, jenis kulit, dan *impact* penggunaan. Selain itu, teknik seperti *Named Entity Recognition* (NER) juga dapat digunakan untuk mengenali dan mengelompokkan entitas tersebut sehingga mempermudah analisis dan penyusunan rekomendasi *skincare* yang lebih relevan (Guo et al., 2020). NLP dan NER tidak hanya mempercepat proses analisis tetapi juga memberikan akurasi tinggi dalam memahami kebutuhan pengguna berdasarkan data teks.

Penelitian ini bertujuan untuk membuat korpus dan mengembangkan model NER untuk mengekstraksi informasi penting dari komentar *TikTok* mengenai kandungan *skincare*, jenis kulit, permasalahan kulit, tips, waktu perawatan, *impact*, perawatan, dan jenis produk dengan cara yang sistematis. Proses ini dilakukan

dengan mengumpulkan data komentar, membentuk korpus menggunakan NLP dan mengembangkan model NER untuk mengenali informasi yang relevan. Validitas hasil penelitian ini dijamin melalui proses verifikasi oleh *expert skincare* serta pengujian performa model menggunakan metrik evaluasi yang sesuai untuk NER, yaitu *precision*, *recall*, dan *F1-score*. Pendekatan ini memastikan hasil yang akurat dan dapat diandalkan, sehingga memberikan kontribusi penting dalam memahami kebutuhan pengguna dan tren perawatan kulit yang berkembang.

1.2. Rumusan Masalah

Berdasarkan latar belakang yang telah dijelaskan, terdapat beberapa permasalahan yang dapat dirumuskan yaitu belum optimalnya pemanfaatan data komentar pada aplikasi *TikTok* untuk menganalisis kebutuhan pengguna terkait kandungan *skincare*, jenis kulit, permasalahan kulit, tips, *impact*, waktu perawatan, perawatan dan jenis produk. Komentar-komentar tersebut memiliki potensi besar untuk menyediakan wawasan mendalam. Namun tanpa pengolahan data yang terstruktur melalui pembuatan korpus dan pengembangan model NER, informasi penting di dalamnya tetap tidak terorganisir dan sulit digunakan. Oleh karena itu, diperlukan pendekatan sistematis menggunakan teknik NLP dan NER untuk mengekstraksi informasi ini secara efektif, guna memberikan manfaat yang lebih besar dalam mendukung kebutuhan pengguna dan inovasi di industri kecantikan.

1.3 Pertanyaan Penelitian

Berdasarkan rumusan masalah di atas, maka pertanyaan penelitiannya adalah:

1. Bagaimana membuat korpus dari komentar aplikasi *TikTok* yang terstruktur terkait kandungan *skincare*, jenis kulit, permasalahan kulit, tips, *impact*, jenis produk, waktu perawatan, dan perawatan?
2. Bagaimana pengembangan model NER dapat membantu mengenali entitas penting seperti kandungan *skincare*, jenis kulit, permasalahan kulit, tips, *impact*, jenis produk, waktu perawatan, dan perawatan secara akurat?

1.4 Tujuan Penelitian

Adapun tujuan dari penelitian ini adalah:

1. Membuat korpus dari data komentar aplikasi *TikTok* yang terstruktur dan relevan dengan topik *skincare*.
2. Mengembangkan model NER untuk mengenali entitas penting, seperti kandungan *skincare*, jenis kulit, permasalahan kulit, tips, *impact*, dan jenis produk, waktu perawatan, dan perawatan.

1.5 Batasan Masalah

1. Penelitian ini hanya menggunakan data berupa komentar pada video aplikasi *TikTok* yang berhubungan dengan ulasan atau *review* produk *skincare* dan konten yang membahas tentang tips pemilihan kandungan.
2. Konten yang dianalisis dalam penelitian ini dibatasi pada komentar dari video *TikTok* yang berbahasa Indonesia.
3. Fokus penelitian ini hanya pada konten yang berkaitan dengan produk *skincare* untuk perempuan.
4. Entitas yang dianalisis dalam penelitian ini dibatasi pada kandungan *skincare*, jenis kulit, permasalahan kulit, tips, *impact*, waktu perawatan, perawatan dan jenis produk.
5. Entitas tips dan *impact* difokuskan hanya pada konteks pemakaian produk *skincare*.
6. Proses NER hanya sampai pada tahap testing model.

1.6 Manfaat Penelitian

Hasil dari penelitian ini diharapkan akan memberikan manfaat antara lain:

1. Manfaat Teoritis

Penelitian ini memberikan kontribusi pada pengembangan ilmu NLP dan NER, khususnya dalam menganalisis data teks dari komentar media sosial di bidang *skincare*. Selain itu, penelitian ini memperkaya literatur tentang pemanfaatan teknologi NLP dalam memahami kebutuhan pengguna dan tren perawatan kulit, serta menjadi referensi untuk penelitian lanjutan di bidang serupa.

2. Manfaat Praktis

Penelitian ini membantu pengguna memahami kandungan *skincare* yang sesuai dengan jenis dan permasalahan kulit mereka. Bagi industri kecantikan, hasil penelitian ini dapat menjadi dasar untuk menciptakan produk yang relevan dan inovatif. Selain itu, penelitian ini memudahkan pengembangan sistem rekomendasi dan memberikan wawasan bagi ahli kecantikan serta dermatologi dalam memberikan rekomendasi yang lebih tepat.

1.7 Sistematika Penulisan

Proposal ini disusun dengan sistematika bagian, terdiri dari 4 (empat) bab yaitu:

1. Bab I Pendahuluan

Bab I berisi penjelasan permasalahan mengenai kesulitan perempuan dalam memahami kandungan *skincare* dan penggabungan bahan aktif yang sesuai dengan jenis kulit berdasarkan komentar *TikTok*. Terdapat beberapa poin yaitu latar belakang masalah, perumusan masalah, pertanyaan penelitian, batasan masalah, tujuan penelitian, dan manfaat penelitian.

2. Bab II Tinjauan Pustaka

Bab II berisi tinjauan pustaka tentang penelitian-penelitian terdahulu yang relevan tentang NLP dan NER.

3. Bab III Landasan Teori

Bab III berisi tentang teori dasar mengenai NLP dan NER. Sumber referensi berasal dari literasi ilmiah seperti jurnal, buku, karya ilmiah.

4. Bab IV Metodologi Penelitian

Bab IV berisi tentang alur penelitian dan kebutuhan yang diperlukan dari pengumpulan data, pembuatan korpus menggunakan metode NLP, dan pengembangan model NER.

5. Bab V Hasil dan Pembahasan

Bab V berisi tentang penjelasan hasil penelitian dan pembahasan terkait penelitian yang telah dilakukan pada pengumpulan data, pembuatan korpus menggunakan metode NLP, dan pengembangan model NER.

6. Bab VI Kesimpulan dan Saran

Bab VI berisi tentang kesimpulan dari penelitian yang sudah dilakukan pada penelitian ini dan berisi saran untuk pengembangan penelitian selanjutnya.

