

BAB I

PENDAHULUAN

1.1. Latar Belakang

Berkembangnya teknologi di era globalisasi ini mendukung persebaran tren kecantikan yang menjadi primadona setiap wanita. Obsesi seorang wanita untuk cantik, menawan, hingga dapat menarik perhatian orang lain menjadi kebanggaan tersendiri sebagai seorang wanita. Hal tersebut menjadi potensi untuk dapat berkembangnya suatu *brand* kecantikan. Daya beli konsumen khususnya di Indonesia sangatlah tinggi sehingga persaingan antar *brand* kecantikan semakin ketat, hal tersebut dapat dilihat dari banyaknya produk kecantikan yang muncul sehingga menyebabkan kenaikan penjualan produk kecantikan sebesar 7% dari tahun 2020 hingga 2021 [1][2][3]. Perkembangan teknologi, telah menciptakan sebuah model baru dalam memperdagangkan suatu produk kepada konsumen. Perkembangan teknologi telah menciptakan internet yang dapat merubah persepsi berdagang yang hanya bisa menjual barang pada konsumen dengan keterbatasan jarak dan waktu menjadi tidak terbatas hingga dapat tersebar luas kepada konsumen lokal hingga universal [4]. Semakin banyaknya produk yang terdapat di pasar menyebabkan konsumen mengalami kesulitan dalam menentukan produk yang sesuai dengan kebutuhannya.

Kesulitan yang dialami konsumen dapat diselesaikan dengan melihat *review* dari konsumen yang telah menggunakan produk tersebut. *Review* yang diberikan mengandung ulasan terhadap produk yang sedang dibahas. Hal yang dibahas pada *review* umumnya akan membicarakan tentang kualitas, aroma, efek yang ditimbulkan, harga, dan kemasan. *Review* yang diberikan akan dijadikan dasar untuk konsumen dalam memilih produk kecantikan. Namun aspek yang ditampilkan seperti jumlah *review*, rating, dan rekomendasi belum mengidentifikasi kejelasan terhadap suatu produk.

Analisis sentimen menjadi solusi yang dapat dilakukan untuk melakukan klasifikasi komentar konsumen terhadap entitas yang sedang dibicarakan [5][6].

Analisis sentimen merupakan sebuah metode yang biasa digunakan untuk memahami, mengekstraksi, serta mengolah suatu data tekstual secara otomatis untuk mendapatkan nilai sentimen dalam sebuah kalimat maupun paragraf[7]. Nilai yang didapatkan setelah melakukan analisis sentimen dikelompokkan menjadi 2 yaitu positif dan negatif. Suatu kalimat dikatakan memiliki sentimen positif apabila kalimat atau paragraf berisikan tersebut berisikan opini yang positif seperti senang, kebanggaan, persetujuan. Sedangkan Suatu kalimat dikatakan memiliki sentimen negatif dinyatakan apabila suatu kalimat atau paragraf berisikan opini yang negatif seperti marah, kecewa, bentuk ketidak setujuan dan sebagainya [8].

Pengambilan data pada penelitian kali ini menggunakan metode *Text Mining* atau pengumpulan data yang merupakan proses ekstraksi informasi dari sumber data yang tidak terstruktur [9]. Salah satu penerapan text mining yang sering dijumpai yaitu *Web Scraping*. Pada penelitian ini, digunakan metode *web scraping* dengan mengambil data dari *web femaledaily* yang merupakan sebuah platform untuk membantu konsumen yang mengalami kesulitan dalam menentukan pilihannya terhadap suatu produk kecantikan berdasarkan pengalaman serta ilmu yang dibagikan oleh pengguna produk kecantikan lain. Teknik yang digunakan dalam mendapatkan data yang dibutuhkan yaitu *Web Scraping* merupakan suatu proses pengambilan data teks semi-terstruktur dari WWW (*World Wide Web*) kemudian disimpan pada file sistem untuk dijadikan data yang siap untuk dianalisis. Pengambilan data pada WWW berupa halaman web dalam bahasa HTML maupun XHTML [10][11][12].

Pada proses analisis sentimen, dibutuhkan suatu metode untuk dapat melatih sistem mengetahui sentimen dari setiap kalimat yang diberikan. Pada penelitian kali ini metode yang digunakan yaitu menggunakan deep learning. Penggunaan Deep learning menjadi populer dikarenakan dapat dengan baik mengolah data dengan jumlah yang banyak dan kompleks dengan baik, waktu pemrosesan yang lebih cepat, dan akurasi yang lebih baik daripada metode machine learning yang regular. Deep Learning memiliki beberapa jenis yang mana di setiap jenisnya memiliki keunggulan serta kekurangannya masing-masing. Metode deep learning

yang akan digunakan kali ini yaitu *Recurrent Neural Network* (RNN) sebagai salah satu arsitektur yang populer dalam *Natural Language Processing* (NLP) dan *Convolutional neural network* (CNN) sebagai algoritma yang dianggap efektif dalam melakukan klasifikasi dalam jumlah besar karena memiliki lapisan yang dapat mempelajari fitur dari data yang dimiliki [13][14]. Setelah melalui proses klasifikasi, kemudian akan dilakukan proses evaluasi dan validasi metode dengan menggunakan *confusion matrix* untuk mendapatkan data benar salahnya sistem dalam melakukan klasifikasi. Hasil akurasi yang didapatkan dari kedua metode akan dibandingkan untuk menyimpulkan algoritma terbaik yang dapat digunakan dalam melakukan analisis sentimen menggunakan dataset berupa *review* konsumen pada website *Femaledaily*.

1.2. Rumusan Masalah

Berdasarkan latar belakang diatas, rumusan masalah yang diperoleh adalah sebagai berikut.

1. Bagaimana tahapan untuk mengklasifikasikan *review* pengguna produk kecantikan terhadap sebuah produk kecantikan dengan menggunakan metode algoritma RNN dan CNN.
2. Bagaimana perbandingan performa antara algoritma RNN dengan CNN dalam melakukan analisis sentimen terhadap *review* pengguna produk kecantikan.

1.3. Batasan Masalah

Agar lebih terfokus dalam menyelesaikan permasalahannya, batasan masalah pada penelitian ini adalah sebagai berikut.

1. Dataset yang digunakan merupakan *review* konsumen produk kecantikan pada website *femaledaily* dari awal produk diterbitkan hingga tanggal 22 Juni 2022.
2. Sentimen yang digunakan pada proses klasifikasi yaitu positif dan negatif.

1.4. Tujuan Penelitian

Berdasarkan pernyataan sebelumnya tujuan yang ingin dicapai pada penelitian ini adalah sebagai berikut.

1. Membangun model algoritma *RNN* dan *CNN* untuk mengklasifikasikan *review* konsumen produk kecantikan terhadap produk yang digunakan pada website *femaledaily*.
2. Membandingkan kinerja metode algoritma *RNN* dan *CNN* dalam melakukan klasifikasi teks *review* konsumen terhadap produk kecantikan yang digunakan.

1.5. Manfaat Penelitian

1. Mengetahui kinerja algoritma *RNN* dan *CNN* untuk mengklasifikasikan *review* konsumen produk kecantikan terhadap produk yang digunakan pada website *femaledaily*.
2. Memberikan informasi yang dapat digunakan sebagai sumber penelitian selanjutnya.