

Penerapan Metode Multinomial Naïve Bayes dengan *Feature Selection Particle Swarm Optimization* terhadap Sentimen Ulasan Produk

Serlii Ratmala Crusita¹, Yuliant Sibaroni, S.Si., M.T.²

^{1,2}Fakultas Informatika, Universitas Telkom, Bandung

¹serliiratmalacruscita@students.telkomuniversity.ac.id, ²yuliant@telkomuniversity.ac.id

1. Pendahuluan

Latar Belakang

Produk kecantikan merupakan produk populer yang sering dicari wanita pada saat ini. Produk yang dicari mulai dari perawatan kulit, perawatan rambut maupun *make-up* untuk tampil segar setiap harinya[1]. Untungnya pada era teknologi yang canggih, menjadikan masyarakat di Indonesia lebih modern dalam mengakses sebuah situs (*e-commerce*). Perubahan yang dapat kita lihat adalah pergeseran perilaku masyarakat dalam membeli kebutuhan barang secara *offline shop* menjadi *online shop* atau *E-commerce*[2]. *E-commerce* banyak digunakan karena mudah dan cepat dalam mengakses sebuah informasi yang dapat dilihat dua puluh empat jam sehari tanpa henti. Calon konsumen dapat melihat *review* pada sebuah produk untuk meningkatkan kepercayaan calon konsumen membeli produk yang diinginkan[2]. Namun dalam mengkategorikan ulasan negatif atau ulasan positif dengan dataset yang mencapai ratusan ribu akan memakan waktu yang lama. Untuk mengatasi kategori ulasan negatif atau ulasan positif dapat menggunakan *sentiment analysis*[3]. *Sentiment analysis* merupakan bidang studi yang menganalisis pendapat, sentimen, penilaian, evaluasi, sikap dan emosi seseorang terkait suatu topik, layanan, produk, individu, organisasi atau kegiatan tertentu[4].

Pelabelan salah satu proses yang harus diperhatikan dalam suatu klasifikasi sentimen. Pelabelan merupakan kegiatan memberikan label pada suatu *dataset*[2]. Pada penelitian ini, pelabelan yang dilakukan adalah pelabelan secara manual dengan menggunakan bantuan tenaga ahli bahasa[5]. Terdapat penelitian sebelumnya mengenai analisis sentimen terhadap ulasan produk kecantikan pada situs penyedia *review* produk yang didapat dari *amazon* menggunakan metode Support Vector Machine (SVM)[8]. Pada penelitian tersebut dilakukan perbandingan Support Vector Machine menggunakan seleksi fitur *Particle Swarm Optimization* dengan Support Vector Machine menggunakan seleksi fitur *Genetic Algorithm* pada analisis sentimen[8]. Hasil yang didapat dari penelitian tersebut, penggunaan Support Vector Machine dengan seleksi fitur *Particle Swarm Optimization* memiliki akurasi lebih tinggi yaitu 97%[8]. Penerapan fitur seleksi *Particle Swarm Optimization* (PSO) telah digunakan di berbagai penelitian seperti penelitian tentang data *review* hotel[9], opini publik dengan data kebakaran hutan[10], analisa dengan data *review* restoran[11], data pada sosial media twitter[12], penilaian kepuasan pelanggan makanan kuliner[13], data *furniture*[14] dan data perusahaan penjualan online[15]. Pada penelitian ini dibangun sebuah sistem untuk mengklasifikasikan ulasan (positif/negatif) menggunakan fitur seleksi *Particle Swarm Optimization* (PSO) untuk meningkatkan nilai akurasi analisa sentimen. *Particle Swarm Optimization* (PSO) banyak digunakan untuk memecahkan masalah optimasi serta masalah seleksi fitur karena sifatnya yang sederhana[16].

Penelitian ini menggunakan metode *Multinomial Naïve Bayes* pada sentimen ulasan produk kecantikan. *Multinomial Naïve Bayes* merupakan salah satu metode yang lebih khusus atau spesifik dari metode *Naïve Bayes*[17]. Dengan menggunakan metode *Multinomial Naïve Bayes* diharapkan dapat mencapai performa terbaik[18]. *Multinomial Naïve Bayes* mengasumsikan fitur pada data bersifat *conditionally independent given the class* sehingga setiap kelas memiliki fitur yang merepresentasikan kelasnya masing-masing[19]. Sehingga metode ini dapat dijadikan *classifier* yang tepat dalam permasalahan klasifikasi sentimen data ulasan produk.

Berdasarkan latar belakang yang sudah diuraikan, sebuah sistem dibuat dengan menerapkan Multinomial *Naïve Bayes* terhadap dataset ulasan sebuah produk kecantikan. Fitur seleksi yang digunakan adalah *Particle Swarm Optimization*, dengan menggunakan fitur seleksi *Particle Swarm Optimization* dapat mengetahui bagaimana pengaruh penggunaan *feature selection* pada metode Multinomial *Naïve Bayes*. Penggabungan metode dengan fitur seleksi *Particle Swarm Optimization* juga dilakukan untuk mengetahui bagaimanakah performa metode Multinomial *Naïve Bayes* dengan menggunakan *feature selection Particle Swarm Optimization* (PSO) terhadap sentimen ulasan produk.

Batasan masalah dari penelitian ini adalah dataset yang digunakan adalah data *Beauty* milik Amazon berbahasa inggris yang berjumlah 5000 *record*. Penelitian ini menggunakan teknik pelabelan secara manual dengan angka 1 sebagai positif dan angka 0 sebagai negatif. Pelabelan pada dataset dilakukan secara manual karena pelabelan yang dilakukan manusia lebih tepat sehingga menghasilkan akurasi yang lebih tinggi[6].

Tujuan dari penelitian ini adalah melakukan pengujian untuk mengetahui pengaruh penggunaan *feature selection Particle Swarm Optimization* (PSO) pada metode Multinomial *Naïve Bayes* dan mengetahui hasil performa sistem yang dibangun dengan menggunakan metode Multinomial *Naïve Bayes* dan penggunaan *feature selection Particle Swarm Optimization* (PSO).

Struktur penulisan makalah tugas akhir ini berisi pendahuluan, studi terkait, sistem yang dibangun, evaluasi, dan kesimpulan. Pada bagian pendahuluan, dijelaskan mengenai latar belakang masalah yang diangkat pada tugas akhir ini, batasan masalah dalam pengerjaan sistem, dan tujuan yang ingin dicapai. Pada bagian 2 merupakan studi terkait yang berisi teori dan literatur yang mendukung penyelesaian tugas akhir. Bagian 3 membahas teori tentang rancangan sistem yang dibangun dan dijelaskan proses gambaran umum sistem. Pada bagian 4 berupa evaluasi penelitian yang terdiri dari penjelasan analisis dari pengujian yang telah dilakukan. Bagian 5 berupa kesimpulan dan saran yang berisi penjelasan analisis dari sistem yang dibangun.