

1. Pendahuluan

Latar Belakang

Beredar produk kecantikan yang di jual di internet oleh berbagai macam produsen baik luar negeri maupun dalam negeri. Akan tetapi masih diragukan kualitas kosmetik yang dijual oleh tiap produsen, agar mengetahui apakah produk tersebut baik digunakan maka produsen perlu mendapatkan ulasan/review dari konsumen yang memakai produk tersebut. Dengan banyaknya ulasan terhadap produk tersebut, konsumen lain dapat melihat dampak dan kegunaan produk tersebut sesuai kebutuhan konsumen itu sendiri. Terutama di bidang kesehatan, apakah produk yang di pakai memiliki dampak pada kulit, seperti jerawat, bibir kering, kulit terbakar, dan masalah kulit lainnya. Salah satu forum yang menampung review tentang produk kecantikan adalah Female Daily. Di dalam forum tersebut produsen dapat mengetahui apa kekurangan dari produk yang terdapat di dalam produk tersebut dengan menerima review dari para pengguna produk tersebut terutama review di bidang kesehatan.

Untuk itu agar produsen lebih mudah untuk mencari produk yang relevan dengan kesehatan maka dibutuhkan sebuah sistem untuk mengklasifikasikan review produk tersebut termasuk kategori relevan atau tidak relevan terhadap aspek kesehatan.

Pada Tugas Akhir ini digunakan Machine learning pada klasifikasi sentimen menggunakan Random Forest, Support Vector Machine(SVM), dan K-Nearest Neighbour(KNN) untuk mencari akurasi tertinggi dan *F1- score* dari ketiga algoritma tersebut dengan menggunakan feature extraction yaitu chi-square dengan *feature selection* menggunakan *Selected K Best* untuk proses *preprocessing*.

Topik dan Batasannya

Dengan adanya permasalahan seperti yang telah dijelaskan pada latar belakang maka dapat disimpulkan oleh penulis permasalahan yang dapat disimpulkan adalah dengan mengetahui cara untuk mengklasifikasikan data teks berdasarkan data *review Female Daily* yang telah ada, dan juga menganalisis hasil teks dari data teks *review Female Daily* yang telah melalui proses klasifikasi.

Tujuan

Tujuan pada penelitian ini sesuai yang sudah dijelaskan pada latar belakang adalah peneliti dapat menerapkan metode klasifikasi SVM, KNN, *Random Forest* menggunakan *chi-square* dalam analisis sentimen berbasis aspek kesehatan dalam review produk *Female Daily*, dan dapat melakukan analisis dari hasil klasifikasi perbandingan tiga algoritma tersebut dengan menggunakan seleksi fitur TF-IDF

Organisasi Tulisan

Penulisan tugas akhir ini tersusun dalam beberapa bagian, yaitu sebagai berikut:

- Pendahuluan
Menjelaskan latar belakang, rumusan masalah dan tujuan dari topik yang diambil
- Studi Terkait
Berisi penelitian-penelitian sebelumnya yang digunakan sebagai literatur acuan dalam pembuatan tugas akhir ini
- Sistem yang dibangun
Menggambarkan alur berjalannya system yang berawal dari metodologi penelitian untuk mencapai objektif, rancangan analisis hingga metode pengujian yang dilakukan
- Evaluasi
Analisis terhadap sentimen yang digunakan menjelaskan hasil klasifikasi algoritma yang digunakan menggunakan seleksi fitur TF-IDF
- Kesimpulan
Penelitian yang dilakukan serta saran untuk penelitian selanjutnya