

ABSTRAK

Nama : Annissa Mutiaputri
Program Studi : Informatika
Judul : Analisis Sentimen Pada Video Ulasan Produk Kecantikan Menggunakan *Random Forest Classifier*
Dosen Pembimbing : Dino Hariatma, M.Kom.

Banyak sekali video ulasan produk kecantikan yang tersebar di *platform* YouTube, namun sangat sedikit yang menganalisis sentiment pada video ulasan tersebut. Penelitian ini melakukan analisis terhadap video ulasan produk kecantikan menggunakan *Random Forest Classifier*. Data dikumpulkan dari *platform* YouTube. Data yang masuk kemudian diseleksi agar tidak terjadi duplikasi data dan berisi hanya ulasan dari satu jenis produk kecantikan yang tidak lebih dari 12 menit. Video dari data yang telah diseleksi akan diunduh dan diubah menjadi bentuk teks melalui dua tahap, yaitu tahap pengubahan video menjadi audio, dan tahap pengubahan audio menjadi teks. Video yang telah menjadi teks kemudian diseleksi kembali agar data yang gagal direkognasi dapat dihapus. Labelisasi dilakukan secara manual dengan label *rating* dari angka 1 hingga angka 5. Data teks yang telah dilabelisasi selanjutnya akan dilakukan *Preprocessing* agar data dapat dimasukkan ke dalam model. Pelatihan data menggunakan *Random Forest Classifier* dengan banyak *tree* sejumlah 5. Hasil akhir analisis adalah klasifikasi *rating* data uji dengan nilai akurasi model sebesar 0.5 atau 50%.

Kata Kunci : Klasifikasi *Rating*, *Random Forest Classifier*, Video Ulasan Produk Kecantikan, YouTube.

ABSTRACT

There are so many review videos of beauty product in YouTube platform, but there are only few people that analyze the sentiment of video reviews. This research is analyzing the beauty product's review videos using Random Forest Classifier. Data collected from YouTube platform. The data that has been collected will be selected to prevent data duplication and to keep only one type of product's reviews that are no longer than 12 minutes. The selected data's video will be downloaded and converted into text by two steps, the first one is to convert the video into audio, and the second one is to convert the audio into text. The converted video will be selected to delete the data that have recognitions failure. The labeling is done manually by adding the 'rate' label from 1 to 5. The labeled data will be preprocessed so it can be processed by the model. The data was trained using Random Forest Classifier with 5 trees. The result of analysis is the rate classification of test data with accuracy point 0.5 or 50%.

Keywords : Rate Classification, Random Forest Classifier, Video Review of Beauty Products, YouTube.