

Klasifikasi Sentimen pada Komentar Youtube Indonesian Idol Menggunakan Metode Naïve Bayes Classifier dan Recurrent Neural Network

Fauziah Giyanti¹, Moch. Arif Bijaksana², Hasmawati³

^{1,2,3}Fakultas Informatika, Universitas Telkom, Bandung

¹faugiyys@students.telkomuniversity.ac.id, ²arifbijaksana@telkomuniversity.ac.id,

³hasmawati@telkomuniversity.ac.id

Abstrak

Semakin berkembangnya teknologi informasi dan komunikasi, media sosial seperti *Youtube* memiliki banyak peminat dan digunakan untuk berkarya seperti membuat konten *Youtube*, *vlog* dan *podcast*. *Youtube* adalah salah satu media informasi dan bisa menjadi sumber data yang sangat besar. Ada beberapa fitur pada *Youtube* yang dapat memberikan informasi yaitu video, gambar, dan komentar pada video yang diunggah. Komentar pada *Youtube* ini akan digunakan untuk mengetahui opini dari penonton *Youtube* terhadap suatu konten. Data berupa teks yang akan digunakan pada Tugas Akhir ini adalah komentar penonton *Youtube Indonesian Idol*. Dengan banyaknya komentar pada video yang diunggah, sulit untuk membedakan komentar yang termasuk positif dan negatif. Hal ini dapat berpengaruh untuk penilaian penampilan peserta dan menjadi penentu hasil *voting* pada *Indonesian Idol*. Cara untuk mengkategorikan komentar tersebut dengan melakukan klasifikasi sentimen. Pada penelitian ini metode klasifikasi yang digunakan yaitu *Naïve Bayes Classifier* dan *Recurrent Neural Network*. *Naïve Bayes* adalah metode klasifikasi data berdasarkan faktor-faktor probabilitas. *Recurrent Neural Network* (RNN) mempunyai kelebihan model komputasi yaitu memproses informasi berurutan yang panjang. Tahap yang dilakukan sebelum klasifikasi yaitu mengumpulkan data dari komentar *Youtube Indonesian Idol* sebagai dataset, lalu *preprocessing*, *labelling data* dan selanjutnya tahap klasifikasi. Proses klasifikasi dengan pembagian data latih 90% dan 10% data uji memberikan hasil akurasi tertinggi pada metode *Naïve Bayes Classifier* sebesar 91.30% dengan *f1-score* untuk kelas positif 95%, negatif 13% dan neutral 0%. Sedangkan pada metode *Recurrent Neural Network* yaitu 91.30% dengan *F1-Score* kelas positif sebesar 94%, negatif 0% dan neutral 0%.

Kata Kunci: *klasifikasi, sentimen, Youtube, Naïve Bayes Classifier, Recurrent Neural Network*

Abstract

With the development of information and communication technology, social media such as YouTube has many enthusiasts and is used for works such as creating YouTube content, vlogs and podcasts. Youtube is one of the information media and can be a very large source of data. There are several features on Youtube that can provide information, namely videos, images, and comments on uploaded videos. Comments on YouTube will be used to find out opinions from YouTube viewers on content. The data in the form of text that will be used in this Final Project are comments from Indonesian Idol Youtube viewers. With so many comments on uploaded videos, it is difficult to distinguish between positive and negative comments. This can affect the evaluation of participant performances and determine the results of voting on Indonesian Idol. The way to categorize these comments is by classifying sentiments. In this study the classification method used is the Naïve Bayes Classifier and Recurrent Neural Network. Naive Bayes is a data classification method based on probability factors. Recurrent Neural Network (RNN) has the advantage of a computational model that processes long sequential information. The stage that was carried out before the classification was collecting data from Indonesian Idol Youtube comments as a dataset, then preprocessing, labeling the data and then the classification stage. The classification process by dividing 90% training data and 10% test data gives the highest accuracy results in the Naïve Bayes Classifier method of 91.30% with f1-scores for 95% positive class, 13% negative and 0% neutral. Whereas in the Recurrent Neural Network method, it is 91.30% with a positive class F1-Score of 94%, 0% negative and 0% neutral.

Keyword: *klasifikasi, sentimen, Youtube, Naïve Bayes Classifier, Recurrent Neural Network*
