IMPLEMENTASI *WEB SCRAPING* UNTUK ULASAN FILM PADA *TWITTER* MENGGUNAKAN ASOSIASI TEKS

(Studi Kasus: Film KKN di Desa Penari)

INTISARI

Media sosial *twitter* merupakan salah satu media sosial yang sudah sangat akrab dengan masyarakat Indonesia. Pengguna *twitter* biasanya akan memberikan informasi atau komentar terbaru mengenai hal yang sedang dibicarakan. Salah satu komentar yang dapat disampaikan oleh pengguna *twitter* adalah mengenai ulasan film. Data ulasan tentu saja sangat bermanfaat untuk menggali informasi mengenai film yang tayang. Banyaknya ulasan yang masuk pada komentar di *twitter* tentu sangat menyulitkan apabila disalin secara manual. Teknik yang digunakan untuk mengambil komentar pada *twitter* adalah dengan menggunakan *Web Scraping*. Penelitian ini menggunakan komentar yang berisi mengenai ulasan Film KKN di Desa Penari pada tanggal 30 April 2022. Tanggal tersebut merupakan tanggal film tayang perdana. Komentar yang diperoleh tidak semuanya berisi ulasan, sehingga perlu diseleksi lagi. Hasil seleksi dari 866 komentar diperoleh 116 ulasan positif dan 83 ulasan negatif. Setelah itu data harus melewati tahap *text preprocessing* yang terdiri dari *cleansing data*, *case folding*, *spelling normalization*, *filtering*, dan *tokenizing*. Setelah melalui tahapan *text preprocessing*, ulasan tersebut kemudian dianalisis untuk mendapatkan informasi yang penting dengan menggunakan asosiasi teks. Hasil asosiasi teks untuk ulasan positif diperoleh informasi bahwa penonton memberikan penilaian terhadap tokoh, akting, dan sinematografi yang bagus, kemudian film yang ditayangkan juga sesuai dengan *thread* yang ada pada *twitter*, dan sinematografi juga keren. Sedangkan untuk ulasan negatif penonton memberikan penilaian bahwa Film KKN di Desa Penari film yang biasa, serta hantu yang ditayangkan juga kurang.

Kata Kunci: Ulasan, *Text Preprocessing*, Film KKN di Desa Penari

WEB SCRAPING IMPLEMENTATION FOR FILM REVIEWS ON TWITTER USING TEXT ASSOCIATION

(Case Study: KKN di Desa Penari Movie)

ABSTRACT

Twitter social media is one of the social media that is very familiar with the people of Indonesia. Twitter users will usually provide the latest information or comments on the matter being discussed. One of the comments that can be submitted by Twitter users is regarding film reviews. Review data is of course very useful for digging up information about films that are showing. The large number of comments that enter comments on Twitter is certainly very difficult if copied manually. The technique used to retrieve comments on Twitter is to use Web Scraping. This study uses comments containing KKN di Desa Penari movie reviews on April 30, 2022. That date is the premiere date of the film. The comments obtained do not all contain reviews, so they need to be selected again. The results of the selection of 866 comments obtained 116 positive reviews and 83 negative reviews. After that, the data must pass through the text preprocessing stage which consists of data cleansing, case folding, spelling normalization, filtering, and tokenizing. After going through the text preprocessing stages, the reviews are then analyzed to obtain important information using text associations. The results of the text association for positive reviews obtained information that the audience gave an assessment of the characters, acting, and good cinematography, then the films shown were also in accordance with the threads on Twitter, and the cinematography was also cool. As for the negative reviews, the audience gave an assessment that the KKN Film in the Dancer Village was an ordinary film, as well as the ghosts that were shown were also lacking.

Keywords: Review, Text Preprocessing, KKN di Desa Penari Movie