

Pengembangan dan Analisis Sistem Pendeteksi Clickbait Thumbnail pada Portal Video Menggunakan Face Recognition dan Analisis Sentimen = Development and Analysis of Video Portal's Clickbait and Thumbnail Detection System using Face Recognition and Sentiment Analysis

Agastya Vitadhani, author

Deskripsi Lengkap: <https://lib.ui.ac.id/detail?id=20516147&lokasi=lokal>

Abstrak

Kombinasi thumbnail dan judul pada portal berbagi video sangat berguna untuk memberikan gambaran mengenai isi video. Tetapi, terdapat juga thumbnail yang bersifat clickbait yang lebih bertujuan memancing masyarakat untuk menonton video. Bahkan terdapat clickbait thumbnail yang dibuat untuk menyesatkan masyarakat dengan agenda untuk membangun opini tertentu. Untuk membatasi variasi pada dataset permasalahan dipersempit pada clickbait thumbnail yang mengandung ujaran kebencian terhadap dai. Thumbnail yang dipilih sebagai dataset terdiri dari foto satu orang atau lebih dengan disertai teks narasi. Dari thumbnail diperoleh informasi berupa identitas individu dengan menggunakan face recognition dan teks narasi dengan menggunakan optical character recognition. Untuk mengolah teks narasi dari thumbnail dan judul video digunakan teknik analisis sentimen. Dibangun model machine learning dengan data yang diperoleh dari identitas, teks narasi dan judul video dengan menggunakan model SVM. Evaluasi terhadap model machine learning dilakukan dengan validasi silang 5-fold. Evaluasi menghasilkan akurasi 0.968, sensitivitas 0.968, nilai presisi 0.9698 dan F1-Score 0.967.

.....The combination of thumbnails and titles on video sharing portals is effective for providing an overview of the content of the video. However, there are also clickbait thumbnails that are more intended to lure people into watching the videos. There are even clickbait thumbnails created to mislead the public with an agenda to build a particular opinion. To limit the variation in the dataset the problem was narrowed down to clickbait thumbnails containing hate speech against dai. Thumbnails selected as datasets consist of photos of one or more people accompanied by narrative text. From the thumbnails, information is obtained in the form of individual identities using face recognition and narrative texts using optical character recognition. To process narrative text from thumbnails and video titles, sentiment analysis techniques are used. A machine learning model was built with data obtained from identity, narration text, and video titles using the SVM model. Evaluation of the machine learning model is carried out with 5-fold cross-validation. The evaluation resulted in an accuracy of 0.968, a sensitivity of 0.968, a precision value of 0.9698, and an F1-Score of 0.967