

INTISARI

KLASIFIKASI RINGKASAN BERITA KRIMINAL MENGGUNAKAN ALGORITMA *ITERATIVE DICHOTOMISER TREE*

Oleh

Herfia Rhomadhona
14/371992/PPA/04627

Teknologi internet menjadi media penyebaran berita kepada khayalak ramai dengan berbagai kategori. Salah satu kategori yang diberitakan di media online adalah berita kriminal. Semakin banyak kriminalitas yang terjadi semakin meningkat jumlah dan volume berita pada sebuah media online. Oleh karena itu dibutuhkan sebuah metode agar bisa meringkas dimensi dokumen tersebut sehingga dapat meningkatkan kinerja akurasi dari proses klasifikasi dokumen berita.

Penelitian ini melakukan pemodelan sistem klasifikasi ringkasan berita dengan menerapkan metode *Iterative Dichotomiser Tree* (ID3). Ringkasan bertujuan untuk meringkas dimensi dokumen. Pembangunan ringkasan otomatis menggunakan metode *Maximum Marginal Relevance* (MMR). Metode ini dikombinasikan dengan metode *cosine similarity* untuk mencari kesamaan judul dengan kalimat dan kalimat dengan kalimat lainnya. Ringkasan berita tersebut digunakan untuk data *input* pada proses klasifikasi. Tahapan pra pemrosesan teks berupa *tokenization*, *stemming* dan *stopword* dan pembobotan dengan TF-IDF untuk *feature selection*. Fitur tersebut digunakan sebagai atribut yang dihitung nilai entropy dan gain untuk pembentukan *Decision Tree*.

Pengujian klasifikasi dilakukan dengan menggunakan data uji sebanyak 200 data dengan persentase ringkasan 10%, 20%, 30%, 50% dan 100%. Nilai akurasi yang dihasilkan setiap persentase ringkasan tidak memiliki perbedaan yang signifikan. Sehingga klasifikasi berita kriminal menggunakan ringkasan (*summary*) dapat mempercepat kinerja klasifikasi dan nilai akurasi yang dihasilkan tetap terjaga.

Kata kunci : *Summary*, MMR, Klasifikasi, *Decision Tree*, TF-IDF.

ABSTRACT

CLASSIFICATION OF CRIMINAL NEWS SUMMARY USING ITERATIVE DICHOTOMISER TREE ALGORITHM

By

**Herfia Rhomadhona
14/371992/PPA/04627**

Internet technology to propagate news to the mob crowded with various categories. One category that is reported in the online media is crime news. The more crime that occurs is increasing the number and volume of news on an online media. Therefore we need a method that can reduce the dimensions of the document so that it can improve the performance of the process of document classification accuracy of the news.

This research news summary classification system modeling by applying the method Iterative Dichotomiser Tree. Summary aims to reduce the dimensions of the document. Development of automatic summary using the Maximum Marginal Relevance (MMR). This method is combined with the method of cosine similarity to seek common ground with the title phrase and sentence with other sentences. The news summaries are used for data input in the classification process. Phases preprocessing text is tokenization, stemming and stopword and TF-IDF weighting for feature selection. The feature is used as an attribute calculated entropy value and the gain for the establishment Decision Tree.

Test on the classification performed using 200 test data, with the percentage of summaries of 10%, 20%, 30%, 50% and 100%. The resulting accuracy value of each percentage of the summary does not have significant differences. So that the classification of crime news using summaries can accelerate the classification performace and value of the resulting accuracy is maintained.

Keywords : Summary, MMR, Classification, Decision Tree, TF-IDF