

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

IMPLEMENTASI *VECTOR SPACE MODEL* UNTUK MENCARI KALIMAT YANG MENGANDUNG IDE POKOK TEKS ARTIKEL BERBAHASA INDONESIA

MERISKA ZAMURI

11251204834

Tanggal Sidang : 18 Agustus 2017

Periode Wisuda : November 2017

Jurusan Teknik Informatika

Fakultas Sains dan Teknologi

Universitas Islam Negeri Sultan Syarif Kasim Riau

ABSTRAK

Ide pokok adalah gagasan utama yang lebih penting dari yang lain yang diwujudkan dalam sebuah kalimat. Masalah yang sering dihadapi oleh pembaca saat membaca artikel panjang sehingga pembaca membutuhkan cukup banyak waktu untuk memahami isi artikel. Selain itu, pembaca juga sulit untuk menentukan ide pokok pada artikel yang berada diawal (deduktif), ditengah (ineratif), dan diakhir (induktif). Untuk mengatasi masalah tersebut maka perlu dibangun sebuah sistem yang dapat mencari kalimat yang mengandung ide pokok sehingga pembaca dapat memahami inti artikel dari kalimat ide pokok tersebut. Penelitian ini menggunakan dua metode yaitu metode *vector space model* dan metode TF-IDF. Perancangan sistem pencarian kalimat ide pokok ini diawali dengan lima tahap *text preprocessing* yaitu pemecahan kalimat, *case folding*, *tokenizing*, *filtering*, dan *stemming*. Algoritma *stemming* yang digunakan adalah algoritma *Porter Stemmer*. Proses selanjutnya menghitung bobot tf-idf dan bobot *vector space model* untuk mengukur kemiripan antara suatu dokumen dengan suatu kueri. Dalam penelitian ini judul artikel yang diinputkan sebagai kueri. Berdasarkan hasil pengujian terhadap 100 artikel, diperoleh akurasi 70%. Pengujian dilakukan dengan tiga variasi yaitu artikel ide pokok diawal (deduktif) diperoleh akurasi sebesar 78,57%, artikel ide pokok ditengah (ineratif) diperoleh akurasi sebesar 40% dan artikel ide pokok diakhir (induktif) diperoleh akurasi sebesar 60%.

Kata Kunci: Algoritma *Porter Stemmer*, Artikel, Kueri, *Text Preprocessing*, *TF-IDF*.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

IMPLEMENTATION OF VECTOR SPACE MODEL TO SEEK THE SENTENCE CONTAINING MAIN IDEAS ON INDONESIAN ARTICLES TEXT

MERISKA ZAMURI

11251204834

Final Exam Date : August 18, 2017

Graduation Ceremony Period : November 2017

Informatics Engineering Department

Faculty of Science and Technology

State Islamic University of Sultan Syarif Kasim Riau

ABSTRACT

The main idea is a central idea which is more important than the others embodied in a sentence. The problem that readers often encounter when reading long articles so that readers need enough time to understand the contents of the article. In addition, the reader is also difficult to determine the main idea in the article that begins (deductive), in the middle (ineratif), and end (inductive). To overcome these problems it is necessary to build a system that can search for sentences that contain the main idea so that the reader can understand the essence of the article from the sentence of the main idea. This research uses two methods namely vector space model and TF-IDF method. The design of the sentence search system of this basic idea begins with five stages of preprocessing text ie sentence splitting, case folding, tokenizing, filtering, and stemming. Stemming algorithm used is Porter Stemmer algorithm. The next process calculates the weight of the tf-idf and the weight of the vector space model to measure the similarity between a document and a query. In this study the title of the article entered as a query. Based on the test results of 100 articles, obtained 70% accuracy. The test is done with three variations, the main idea article at the beginning (deductive) obtained by 78.57% accuracy, the main idea article in the middle (ineratif) obtained by the accuracy of 40% and the main idea article ends (inductive) obtained 60% accuracy.

Keywords: *Porter Stemmer Algorithm, Article, Query, Text Preprocessing, TF-IDF.*