

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.

PENERAPAN METODE *WINNOWERING FINGERPRINT* DAN *MULTINOMIAL LOGISTIC REGRESSION* PADA PENGELOMPOKAN DOKUMEN ABSTRAK TUGAS AKHIR

MUHAMMAD ADI PUTRA
11051103394

Tanggal Sidang: 28 Juli 2017

Periode Wisuda: Agustus 2017

Jurusan Teknik Informatika

Fakultas Sains dan Teknologi

Universitas Islam Negeri Sultan Syarif Kasim Riau

ABSTRAK

Keanekaragaman dan jumlah dokumen abstrak tugas akhir pada saat ini terus bertambah dan tidak terorganisir dengan baik menyebabkan pencarian informasi semakin sulit, solusi untuk mengatasinya adalah dengan membangun suatu sistem klasifikasi dokumen. Penelitian ini menerapkan metode *winnowering* untuk pemilihan fitur yaitu *fingerprint*, *tf-idf* untuk pembobotan dan *multinomial logistic regression* untuk pengelompokan. Pengelompokan dokumen dengan menggunakan *winnowering fingerprint* dan *multinomial logistic regression* mempunyai 8 bidang keahlian dengan menggunakan 1050 dokumen abstrak dengan 90% data latih dan 10% data uji. Pengujian menghasilkan akurasi 84,7% ($k\text{-gram}=8$, bilangan prima=2 dan jumlah $window=8$), Sedangkan pengujian menggunakan data yang seimbang, yaitu 400 data latih (masing-masing kelas memiliki 50 dokumen) menghasilkan akurasi 72,5% ($k\text{-gram}=8$, bilangan prima=2 dan jumlah $window=8$).

Kata Kunci: *Fingerprint*, *Multinomial logistic regression*, Pengelompokan Dokumen Abstrak Tugas akhir, *Text Mining*, *TF-IDF*, *Winnowering*.

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.

IMPLEMENTATION OF WINNOWER FINGERPRINT METHOD AND NAIVE BAYES FOR THESIS ABSTRACT DOCUMENTS CLASSIFICATION

MUHAMMAD ADI PUTRA
11051103394

Date of Final Exam : July 28th, 2017

Graduation Ceremony Period : August 2017

Informatics Engineering Department

Faculty of Sciences and Technology

State Islamic University of Sultan Syarif Kasim Riau

ABSTRACT

The diversity as well as the number of abstract documents of thesis at this time continues to grow and are not well organized resulting in increasingly difficult information retrieval, the solution to overcome this is to build a documents classification system . This research applies winnowing fingerprint method for feature selection, tf-idf for weighting and multinomial logistic regression for classification. Documents classification with winnowing fingerprint and multinomial logistic regression have 8 expertise areas using 1050 of abstract documents with 90% of training data and 10% of test data. The results get accuracy 84.7% (k-gram=8, prime number=2 dan number of windows=8). While using balanced data that is, 400 of training data (each class have 50 documents) giving accuracy 72,5% (k-gram=8, prime number=2 dan number of windows=8).

Keyword : *Fingerprint, Multinomial Logistic Regression, Text Mining, TF-IDF, Thesis Abstract Documents Classification, Winnowing.*