

## INTISARI

### IMPLEMENTASI MODIFIKASI ENHANCED CONFIX STRIPPING STEMMER UNTUK BAHASA JAWA KUNO (KAWI)

Oleh

Komang Ary Tebuana  
14/363864/PA/15866

Pemrosesan digital dalam mencari informasi pada dokumen teks berbahasa Kawi (Jawa Kuno) memerlukan *filtering* dengan menggunakan teknik dalam sistem temu balik informasi yaitu *Stemming*. Penggunaan *stemming* bertujuan untuk mencari kata dasar yang ada pada sebuah *term* atau kata berimbuhan dengan mengekstraksi kata dari imbuhan yang melekat pada kata tersebut.

*Stemmer* Bahasa Kawi yang dibangun dalam penelitian ini menggunakan algoritma *Enhanced Confix Stripping*. Algoritma berbasis kamus tersebut diuji pula performanya apabila tanpa pengecekan kamus dan dengan beberapa metode pengecualian aturan pemotongan imbuhan. Selain itu, dibandingkan pula dengan *stemmer* yang dibangun dengan algoritma Porter. Obyek penelitian yang digunakan adalah kata-kata dalam Bahasa Kawi pada kitab Sarasamuscaya yang berjumlah 860 kata. Data uji diproses dengan menerapkan aturan pembentukan kata berimbuhan (*wangun*) dalam Bahasa Kawi tanpa aturan sandhi pada masing-masing algoritma *stemming*.

Keluaran dari *stemmer* yaitu berupa kata dasar hasil pemrosesan pemotongan imbuhan. Dari penelitian ini, didapatkan bahwa waktu proses dan akurasi bergantung pada metode yang digunakan. *Enhanced Confix Stripping* memperoleh nilai akurasi yang tertinggi, sedangkan *Enhanced Confix Stripping* tanpa kamus dengan waktu proses yang paling lambat dan terjadi penurunan akurasi. Dengan dilakukan pengecualian pemotongan pada imbuhan tertentu mampu mendapatkan peningkatan performa pada algoritma *ECS* tanpa kamus dan Porter untuk Bahasa Kawi.

**Kata Kunci:** *Stemming, Bahasa Kawi, Bahasa Jawa Kuno, Enhanced Confix Stripping, Porter.*

## ABSTRACT

### IMPLEMENTATION OF MODIFICATION ENHANCED CONFIX STRIPPING STEMMER FOR OLD JAVANESE LANGUAGE (KAWI)

By

Komang Ary Tebuana  
14/363864/PA/15866

Digital processing in searching information on Kawi's (Old Javanese) language document required filtering by using technique in information retrieval that is *Stemming*. Stemming used to find the root word that exists in a term or word with affixes by extracting words from affixes attached to the word.

The Kawi language *stemmer* built in this study uses *Enhanced Confix Stripping* algorithm. This dictionary-based algorithm also tested for performance if without dictionary checking and with some exception method of affix cutting rule. In addition, this *stemmer* compared to stemmer build with Porter algorithm. The research object is words with Kawi language in *Sarasamuscaya* book which amounted to 860 words. The test data is processed by applying the formation (*wangun*) rules on each *stemming* algorithm.

The output of *stemmer* is the root word that the result of affix cutting process. From this research, it is found that process time and accuracy depend on the method used. *Enhanced Confix Stripping* receives the highest accuracy value, while *Enhanced Confix Stripping* without dictionary with the slowest processing time and decreased accuracy. With the exception of cutting at certain affixes it is able to get a performance improvement on the *stemming* algorithm without the dictionary for Kawi Language.

**Keyword:** *Stemming, Kawi Language, Old Javanese Language, Enhanced Confix Stripping, Porter.*