



### **Hak cipta dan penggunaan kembali:**

Lisensi ini mengizinkan setiap orang untuk menggubah, memperbaiki, dan membuat ciptaan turunan bukan untuk kepentingan komersial, selama anda mencantumkan nama penulis dan melisensikan ciptaan turunan dengan syarat yang serupa dengan ciptaan asli.

### **Copyright and reuse:**

This license lets you remix, tweak, and build upon work non-commercially, as long as you credit the origin creator and license it on your new creations under the identical terms.

# BAB I

## PENDAHULUAN

### 1.1. Latar Belakang

Akhir-akhir ini teknologi komputer berkembang dengan pesat. Hal tersebut dibuktikan dengan semakin canggihnya perangkat keras seperti prosesor, memori, dan sebagainya. Perkembangan-perkembangan pada perangkat keras tersebut membuat kebutuhan para pengguna komputer dalam hal perangkat lunak pun bertambah. (Adamss, 2011) Untuk itulah para pengembang terus mengembangkan perangkat lunak terutama yang berguna bagi kebutuhan manusia.

Semakin berkembangnya teknologi, dokumen-dokumen dalam format digital mulai banyak ditemukan melalui internet, seperti dokumen-dokumen skripsi, jurnal ilmiah, dan apapun yang berisikan teks. Melihat semakin banyaknya dokumen digital, diperlukan suatu metode pencarian untuk mencari informasi yang sesuai dengan yang diharapkan dan relevan. Saat ini pencarian kata-kata terhadap sekumpulan dokumen dilakukan secara manual dengan cara mencari satu kata per satu kata dari suatu dokumen. Hal ini membuat proses pencarian menjadi pekerjaan yang sangat menyita waktu. Berdasarkan permasalahan tersebut solusi dari pencarian kata-kata untuk beberapa dokumen digital dapat dilakukan dengan membangun sebuah aplikasi pencarian multi dokumen, yang didalamnya terdapat proses *stemming*.

## 1.2. Identifikasi Masalah

Dari latar belakang dapat disimpulkan bahwa permasalahan yang ada adalah kesulitan pengguna dalam mencari kata-kata dalam beberapa dokumen digital.

## 1.3. Batasan Masalah

Batasan – batasan masalah yang digunakan adalah sebagai berikut

1. Bahasa yang dapat diproses adalah bahasa Inggris.
2. Dokumen yang dapat dibaca adalah dokumen yang berformat PDF, DOCX, TXT.
3. Konten yang dapat diambil berupa teks.
4. Untuk menyimpan data hasil stem, dokumen disimpan didalam file TXT.
5. Algoritma yang digunakan dalam penelitian ini adalah Algoritma Porter.
6. Library PDFBox digunakan untuk proses mendapatkan teks dari pdf.
7. Menggunakan library TextExtractor untuk proses mendapatkan teks dari docx.
8. Aplikasi yang akan dibangun dalam penelitian ini dikembangkan dengan bahasa pemrograman Visual C# .NET versi 3.5 SP1.

#### **1.4. Tujuan Penelitian**

Tujuan dari penelitian ini adalah

Mengimplementasikan algoritma Porter untuk pencarian kata dan kalimat yang ada dalam beberapa dokumen *digital*, berupa aplikasi yang dapat mencari kalimat dari sekumpulan dokumen digital

#### **1.5. Manfaat Penelitian**

Penelitian ini diharapkan akan memberikan manfaat kepada *user* terkait dengan

1. Membantu pencarian di beberapa dokumen digital seperti format PDF, DOCX, dan TXT.
2. Implementasi algoritma porter pada proses pencarian dalam beberapa dokumen digital.

#### **1.6. Metode Penelitian**

##### **1. Studi Literatur**

Dalam metode ini dilakukan pencarian dan pembelajaran proses literatur yang berhubungan dengan rumusan masalah, manfaat penelitian, dan tujuan penelitian.

##### **2. Analisa**

Pada metode penelitian ini dilakukanlah suatu analisa terhadap suatu masalah yang timbul terkait dengan permasalahan yang ada.

##### **3. Rancangan**

Pada metode penelitian rancangan dilakukan beberapa perancangan untuk pembuatan aplikasi.

#### 4. Pembuatan Aplikasi C#

Pada metode ini aplikasi akan dibuat dengan menggunakan bahasa pemrograman C#.

#### 5. Uji Coba dan Evaluasi

Dalam metode uji coba dan evaluasi dilakukanlah percobaan-percobaan pada aplikasi yang sudah dibuat. Jika ditemukan kesalahan, maka akan dilakukan perbaikan pada tahap ini juga. Metode penelitian ini juga akan membahas tentang hasil evaluasi dari uji coba – uji coba yang dilakukan.

### 1.7. Sistematika Penulisan

Sistematika penulisan laporan skripsi ini dijelaskan sebagai berikut

#### Bab I Pendahuluan

Bab ini berisikan tentang penulisan penelitian seperti latar belakang, rumusan masalah, batasan masalah, tujuan penelitian, manfaat penulisan, metode penelitian, dan sistematika penulisan.

#### Bab II Landasan Teori

Dalam Bab ini akan dijelaskan teori-teori yang berhubungan dengan penelitian ini seperti *Stemming* dan *Porter Stemmer*.

#### Bab III Analisis dan Perancangan

Bab ini menerangkan tentang analisa permasalahan, penjelasan solusi, dan perancangan dari solusi tersebut.

#### Bab IV Implementasi dan Evaluasi

Bab ini akan membahas pengimplementasian rancangan yang telah dijelaskan pada bab III. Selain itu bab ini juga berisi tentang hasil uji coba dan evaluasi dari aplikasi yang sudah berhasil dibangun

#### Bab V Simpulan dan Saran

Bab ini berisi simpulan dan saran dari hasil penelitian yang sudah berhasil dilakukan dan dievaluasi.



UMN