BAB I
PENDAHULUAN

1.1 Latar Belakang Masalah

Pusat Penelitian dan Pengembangan Sumber Daya Air adalah salah satu instansi yang bergerak dibidang penelitian dan pengembangan Sumber Daya Air. Di Pusat Penelitian dan Pengembangan Sumber Daya Air hasil analisis dari penelitian akan didokumentasikan dan disimpan diperpustakaan Pusat Penelitian dan Pengembangan Sumber Daya Air serta diberikan kata kunci untuk indeks pencarian.

Suatu kata kunci merupakan sebuah kata atau konsep dengan keistimewaan, yang berarti kata apapun yang digunakan sebagai kunci dan kode atau digunakan untuk menghubungkan kekata lain atau informasi lain [1]**.** Kata kunci yang paling baik adalah kata kunci yang terdiri dari rangkaian kata benda atau kata sifat [2]. Kata kunci suatu dokumen mewakili isi dari dokumen tersebut. Dengan kata kunci, kita dapat mengetahui inti dari dokumen itu. Kata kunci juga digunakan dalam pencarian suatu dokumen dengan mesin pencari. Oleh karena itu, menentukan kata kunci sangatlah penting. Kita dapat memahami isi dari dokumen sebelum membaca keseluruhan dari dokumen tersebut dengan kata kunci. Apabila kata kunci yang dipilih tidak tepat, maka seseorang dapat salah dalam memahami isi dari dokumen sebelum ia membaca dari keseluruhan dokumen dan juga seorang informan atau peneliti yang ingin melakukan pencarian terhadap dokumen yang dibutuhkannya tidak terarahkan dengan baik sesuai dengan data yang diinginkan.

Berdasarkan observasi di Perpustakaan Pusat Penelitian dan Pengembangan Sumber Daya Air, hasil penelitian tahun-tahun sebelumnnya masih terdapat abstrak yang belum diberikan kata kunci. Saat ini di perpustakaan dan peneliti Pusat Penelitian dan Pengembangan Sumber Daya Air untuk menentukan kata kunci suatu dokumen masih dilakukan dengan membaca ulang isi dari abstrak suatu dokumen. Hal tersebut membuat proses pembentukan kata kunci

membutuhkan waktu yang lama dan jika dokumen tersebut banyak maka akan membutuhkan sumber daya manusia yang cukup banyak pula.

Oleh karena itu, dibutuhkan penggunaan kata kunci yang tepat pada dokumen tersebut, sehingga kata kunci yang akan digunakan tepat pemilihannya sesuai dengan isi dokumen. Selain itu juga dapat mempercepat waktu dalam pemberian kata kunci setiap dokumen.

Salah satu algoritma yang dapat digunakan dalam penentuan kata kunci adalah Algoritma *Baum Welch.* Algoritma *Baum Welch* merupakan algoritma yang digunakan pada proses learning dengan mengoptimalkan nilai parameter Model Markov Tersembunyi atau *Hidden Markov Method* (HMM). Dalam menentukan rangkaian yang belum diketahui, algoritma *Baum Welch* mengoptimalkan nilai parameter yang diberikan. Dinamai Algoritma *Baum Welch* berasal dari nama Leonard E. Baum dan Lloyd R. Welch [3]. Algoritma *Baum Welch* dikenal juga dengan Algoritma Forward-Backward [4].

HMM telah banyak digunakan karena model yang kaya akan struktur matematika dan mengandung teori dasar yang dapat diterapkan dalam berbagai bidang [5]. Sudah banyak penelitian dilakukan menggunakan Algoritma *Baum Welch* dibidang teknik elektro, ilmu komputer, statistik komputasi dan bioinformatika [3], terutama diaplikasikan di bidang speech recognition dan bioinformatics [6].

Setiap algoritma memiliki performansi yang berbeda-beda. Nilai performansi suatu algortima dilihat dari kemangkusannya. Suatu algoritma dikatakan bagus apabila algoritma tersebut mangkus atau efisien. Kemangkusan algoritma dapat dilihat dari beberapa faktor. Salah satunya adalah menghitung jumlah waktu yang dibutuhkan untuk menjalankannya dengan menghitung nilai kompleksitas waktu asimptotik menggunakan notasi Big-O [7] [8].

Berdasarkan permasalahan di atas maka penelitian ini dilakukan untuk Analisis Performansi Algoritma *Baum Welch* Dalam Menentukan Kata Kunci Suatu Dokumen serta mendapatkan teknik pencarian kata kunci suatu dokumen.

1.2 Perumusan Masalah

Berdasarkan latar belakang masalah maka dapat dirumuskan bagaimana mendapatkan teknik pencarian kata kunci suatu dokumen dengan menganalisis performansi Algoritma *Baum Welch*.

1.3 Maksud dan Tujuan

Berdasarkan permasalahan yang diteliti, maka maksud dari penulisan tugas akhir ini adalah untuk melakukan analisis performansi Algoritma *Baum Welch* dalam menentukan kata kunci suatu dokumen.

Sedangkan tujuan yang akan dicapai dalam penelitian ini adalah :

1. Untuk mengetahui nilai performansi Algoritma *Baum Welch* dalam menentukan kata kunci suatu dokumen di lingkungan perpustakaan Pusat Penelitian dan Pengembangan Sumber Daya Air dengan menghitung kompleksitas waktu asimptotik Algoritma *Baum Welch* menggunakan notasi Big-O.
2. Untuk mendapatkan teknik pencarian kata kunci suatu dokumen di lingkungan perpustakaan Pusat Penelitian dan Pengembangan Sumber Daya Air, sehingga dapat mempercepat waktu dalam penentuan kata kunci dokumen.

1.4 Batasan Masalah

Batasan masalah dari analisis performansi Algortima *Baum Welch* dalam menentukan kata kunci suatu dokumen adalah sebagai berikut:

1. Data masukan pada simulator berupa abstrak yang yang ada di lingkungan perpustakaan Pusat Penelitian dan Pengembangan Sumber Daya Air yang telah diberikan kata kunci serta jumlah kata kunci dan kata kunci abstrak tersebut akan dibandingkan dengan kata kunci dan jumlah kata kunci yang dihasilkan oleh simulator.
2. Algoritma yang diterapkan adalah Algoritma *Baum Welch* dalam metode Model Markov Tersembunyi atau *Hidden Markov Metode* (HMM).
3. Kajian tata bahasa tidak dilakukan ke dalam sistem diasumsikan sudah benar secara tata bahasa.
4. Aplikasi berjalan pada sistem operasi *windows* berbasis desktop.

5. Aplikasi yang dibangun dalam pembuatan simulasi menggunakan Delphi 2009.
6. Pendekatan Pembangunan Perangkat Lunak ini menggunakan metode tersruktur.
7. Menganalisis performansi Algoritma *Baum Welch* dengan menghitung nilai kompleksitas waktu Asimptotik Algoritma *Baum Welch* menggunakan notasi Big-O.
8. Membandingkan hasil kata kunci baru dan jumlah kata kunci baru dengan kata kunci lama dan jumlah kata kunci lama dengan asumsi bahwa kata kunci lama telah benar, sehingga untuk membandingkan mengacu pada kata kunci lama.
9. Membandingkan kesesuaian kata kunci baru dengan kriteria kata kunci dan kesesuaian dengan isi abstrak.
10. Keluaran yang dihasilkan pada simulator adalah kumpulan kata kunci, jumlah kata kunci, nilai kompleksitas waktu asimptotik Big-O dan nilai parmeter yang mempengaruhinya.
11. Kata kunci yang dihasilkan oleh sistem adalah beberapa kata kunci yang terdiri dari gabungan 2 kata.

1.5 Metodologi Penelitian

Metodologi penelitian yang digunakan adalah metode deskriptif, yaitu metode yang menggambarkan fakta-fakta dan informasi dalam situasi atau kejadian sekarang secara sistematis, faktual dan akurat. Metode penelitian ini memiliki dua metode yaitu metode pengumpulan data dan metode simulasi perangkat lunak.

1.5.1. Metode Pengumpulan Data

Metode pengumpulan data yang digunakan dalam penelitian ini adalah sebagai berikut :

a. Studi Literatur

Metode pengumpulan data dilakukan dengan cara mengumpulkan literatur, jurnal, *paper*, bacaan-bacaan dan website yang ada kaitannya dengan judul penelitian.

b. Observasi

Metode pengumpulan data dengan mengadakan penelitian dan peninjauan langsung terhadap permasalahan yang diambil.

c. Interview

Metode pengumpulan data dilakukan dengan mengadakan tanya jawab secara langsung yang ada kaitannya dengan topik yang diambil [9].

1.5.2. Metode Simulasi Perangkat Lunak

Metode simulasi perangkat lunak menggunakan paradigma perangkat lunak secara *waterfall*, yang meliputi beberapa proses diantaranya:

a. System Engineering

Perangkat lunak selalu menjadi bagian dari sistem yang lebih besar, sehingga dimulai dengan menganalisis seluruh sistem memperoleh persyaratan suatu sistem tersebut secara keseluruhan. Kemudian memutuskan fungsi apa yang akan dialokasikan untuk komponen-komponen perangkat lunak. Output dari keuntungan yang muncul dari tahap ini adalah sebagai spesifikasi sistem persyaratan untuk sistem perangkat lunak dalam konteks yang lebih luas dari komponen sistem berakhir.

b. Analysis

Tahan ini berkonsentrasi pada perangkat lunak, memberikan detail yang lebih dan berfokus pada proses pengumpulan persyaratan yang dimulai pada tahap sebelumnya. Untuk memperoleh pemahaman wilayah informasi yang menyeluruh. Setelah memperhatikan pada apa yang perlu dilakukan.

Masukan untuk tahap ini adalah sistem spesifikasi persyaratan yang diproduksi oleh keadaan sistem rekayasa. Hasil dari tahap ini adalah spesifikasi persyaratan perangkat lunak yang berwujud, cukup tepat dan lengkap digunkan sebagai dasar untuk pemecahan rancangan tersebut.

c. Design

Proses desain berkaitan dengan banyak aspek dari sistem perangkat lunak. Untuk memperoleh gambaran abstrak perangkat lunak secara menyeluruh untuk dilaksanakan.

d. Coding

Tahap penerjemahan data atau pemecahan masalah adalah tugas yang lebih sederhana. Tujuannya adalah untuk mengubah gambaran abstrak yang dihasilkan oleh tahap perencanaan ke dalam sesuatu yang lebih berwujud dan nyata. Mengubah kode dalam bahasa tertentu yang dapat dikompilasi atau ditafsirkan untuk menghasilkan modul yang dapat dilaksanakan.

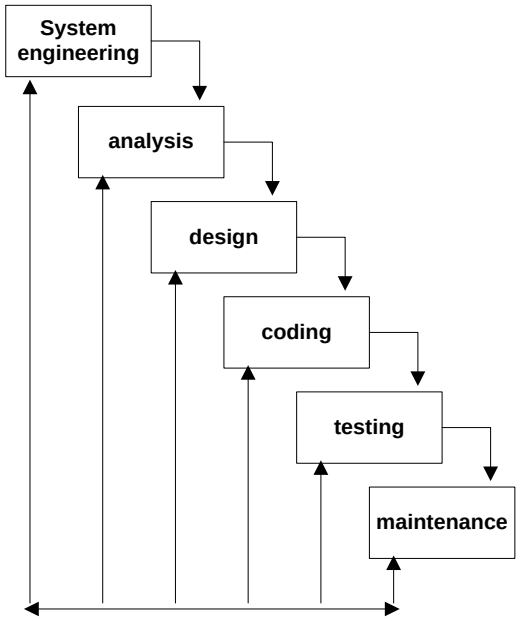
e. Testing

Memastikan bahwa dari sudut pandang dalam, semua pernyataan dalam kode telah diuji, dan dari sudut pandang luar, semua persyaratan telah dipenuhi. Hasil dari tahap ini adalah suatu verifikasi dan validasi produk perangkat lunak tersebut.

f. Maintenance

Tahap ini tidak terlalu banyak dalam pembangunan tetapi lebih merupakan gagasan suatu lingkup untuk mencakup sisa penggerak sebuah sistem perangkat lunak setelah disampaikan kepada siapa pun yang ditugaskan hal tersebut. [10]

Berikut adalah gambar *Waterfall Model*.



Gambar I.1 *Waterfall model* [10]

1.6 Sistematika Penulisan

Sistematika penulisan dalam penelitian ini disusun untuk memberikan gambaran umum tentang penelitian yang dijalankan. Sistematika penulisan tugas akhir ini adalah sebagai berikut :

BAB I PENDAHULUAN

Bab I menguraikan tentang latar belakang permasalahan, mencoba merumuskan inti permasalahan yang dihadapi, menentukan tujuan dan kegunaan penelitian, yang kemudian diikuti dengan pembatasan masalah, asumsi, serta sistematika penulisan.

BAB II LANDASAN TEORI

Bab II membahas berbagai konsep dasar dan teorI.teori yang berkaitan dengan topik penelitian yang dilakukan dan hal-hal yang berguna dalam proses analisis permasalahan serta tinjauan terhadap penelitian-penelitian serupa yang telah pernah dilakukan sebelumnya termasuk sintesisnya.

BAB III ANALISIS DAN PERANCANGAN SISTEM

Bab III menganalisis masalah dari model penelitian untuk memperlihatkan keterkaitan parameter yang diteliti serta model matematis untuk analisisnya.

BAB IV IMPLEMENTASI DAN PENGUJIAN SISTEM

Bab IV mengimplementsi hasil analisis dan pengujian terhadap sistem yang telah dirancangan dan dibangun.

BAB V KESIMPULAN DAN SARAN

Bab V berisi kesimpulan dan saran yang sudah diperoleh dari hasil penulisan tugas akhir.