BAB 1
PENDAHULUAN

1.1 Latar Belakang Masalah

Ambiguitas merupakan hal yang bermakna dua atau mempunyai lebih dari dua makna. Sebagian besar orang salah mengartikan apa yang dibaca dan yang didengarnya. Hal ini dapat terjadi baik dalam ujaran lisan maupun tulisan. Penaafsiran lebih dari satu makna seringkali dapat menimbulkan keraguan dan kebingungan dalam mengambil keputusan tentang makna yang dimaksud.

Dalam bidang kelimuan *computer science*, khususnya *Natural Language Processing* (NLP), ambiguitas pada suatu kata dapat diselesaikan dengan pendekatan *Word Sense Disambiguation*. Dimana suatu keambiguan diproses, dan diberikan makna kata terhadap keambiguan tersebut, dengan tujuan menghilangkan keambiguan itu sendiri [1].

Penelitian atas *word sense disambiguation* sendiri sudah banyak dilakukan serta menghasilkan banyak *annotated corpus* yang dapat dipakai sebagai referensi dan data latih untuk penelitian selanjutnya, namun sangatlah disayangkan *annotated corpus word sense disambiguation* dalam bahasa Indonesia masih sangatlah minim.

Sebuah *annotated corpus word sense disambiguation*, yang layak digunakan sebagai data latih penelitan, haruslah merupakan hasil anotasi manual, yang ditandai dan telah melewati proses validasi oleh ahli bahasa. Proses tersebut seringkali menyulitkan ahli bahasa yang belum tentu memiliki tingkat literasi pada bidang komputer yang cukup tinggi untuk melakukan atau anotasi terhadap corpus, atau membaca sebuah *annotated corpus* untuk keperluan proses validasi.

Berdasarkan uraian dari permasalahan di atas, maka pada penelitian ini akan dibangun sebuah aplikasi anotasi, yang secara fungsional dapat membantu

memenuhi kebutuhan pengguna dalam membangun sebuah annotated corpus *word sense disambiguation* pada teks bahasa Indonesia.

1.2 Perumusan Masalah

Berdasarkan latar belakang penulisan, maka dapat disimpulkan bahwa, yang menjadi rumusan dari permasalahan dari penelitian ini adalah, sulitnya pembangunan *annotated corpus word sense disambiguation* dalam bahasa Indonesia.

1.3 Maksud dan Tujuan

Adapun maksud dari penelitian ini adalah pembangunan sebuah aplikasi yang dapat membantu keperluan anotasi dan validasi terhadap *corpus word sense disambiguation*. Sedangkan tujuan yang ingin dicapai dari penelitian ini adalah, dengan adanya aplikasi yang dibangun, proses anotasi dan validasi pada *annotated corpus word sense disambiguation* lebih dipermudah.

1.4 Batasan Masalah

Agar permasalahan yang ditinjau tidak meluas dari maskud dan tujuan awal yang ingin dicapai, maka terdapat batasan-batasan masalah sebagai berikut.

Batasan Data Masukan :

1. Data masukan berupa file dokumen dalam format .txt.
2. Data masukan artikel berita dalam Bahasa Indonesia.
3. Artikel yang digunakan dalam data masukan berasal dari portal berita online (sindonews, vice)

Batasan Proses :

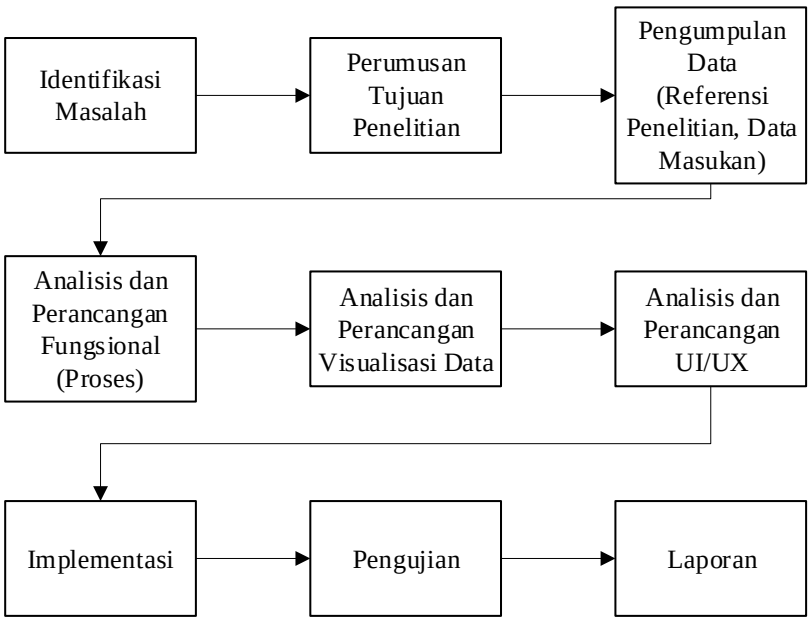
1. Aplikasi yang dibangun dibagi menjadi 2 proses, yaitu anotasi, dan validasi.
2. Teks, dan penandaan ambigu menggunakan dbms mysql sebagai media penampung data.

Batasan Data Keluaran :

Data keluaran dari penelitian ini berupa *annotated corpus* hasil anotasi, dan *input* berita yang telah divalidasi.

1.5 Metode Penelitian

Metode penelitian merupakan tahapan yang telah ditentukan dalam melakukan sebuah penelitian yang digunakan sebagai pedoman dalam proses penelitian agar penelitian yang dilakukan dapat berjalan secara baik dan sistematis. Metode penelitian yang digunakan pada penelitian ini adalah metode deskriptif [2]. Metode deskriptif digunakan karena pada penelitian ini karakteristik subjek atau objek penelitian digambarkan secara sistematis. Metode penelitian dibagi ke dalam dua jenis, yaitu metode pengumpulan data dan metode pengembangan perangkat lunak. Adapun alur penelitian yang dapat dilihat pada Gambar 1.1.



Gambar 1.1 Alur Penelitian

1.5.1 Metode Pengumpulan Data

Teknik pengumpulan data yang dilakukan adalah dengan menggunakan metodologi penelitian. Metodologi penelitian merupakan tahap-tahap penelitian

yang harus diterapkan sebelum melakukan pemecahan masalah. Dalam menganalisa digunakan metode–metode sebagai berikut :

1. Crawling

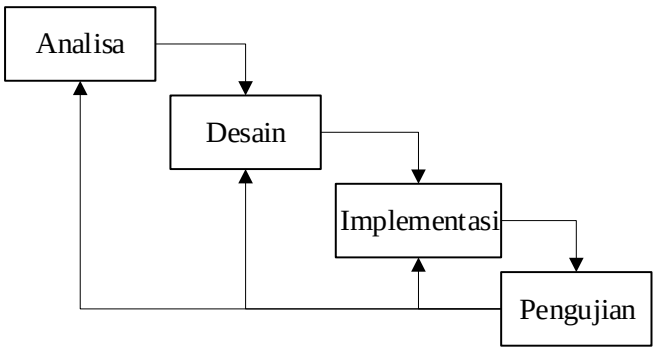
Melakukan pengamatan dan pengumpulan data obyek penelitian. Pengumpulan data dilakukan dengan melakukan *web-crawling* pada portal portal berita dalam Bahasa Indonesia.

2. Penelitian Perpustakaan (Library Research)

Metode yang dipakai untuk mengumpulkan data yang sifatnya teoritis dengan membaca buku-buku atau literatur yang ada hubungannya dengan laporan tugas akhir.

1.5.2 Metode Pembangunan Perangkat Lunak

Proses pembangunan perangkat lunak menggunakan model *waterfall*. Menurut Roger S. Pressman, model yang terkadang disebut siklus classic life atau model *waterfall*, model linear sequential menyarankan pendekatan sistematis dan sekuensial untuk pengembangan perangkat lunak yang dimulai pada tingkat system dan berlanjut melalui analisis, perancangan, pengkodean, pengujian, dan dukungan [3]. Berikut ini merupakan metode *waterfall* menurut Pressman yang dapat dilihat pada Gambar 1.2.



Gambar 1.2 Metode *Waterfall*

Adapun tahapan dari metode *waterfall* di atas menurut Roger S. Pressman adalah sebagai berikut: [3]

1. Analisa

   Analisis merupakan tahap pengumpulan dan pengelompokan masalah serta kebutuhan perangkat lunak yang akan dibangun.

2. Desain

   Desain atau perancangan perangkat lunak adalah proses yang berfokus pada empat atribut program yang berbeda, yaitu: struktur data, arsitektur sistem perangkat lunak, perancangan antarmuka, dan detail prosedural.

3. Implementasi

   Merupakan tahap implementasi dimana aplikasi mulai dibangun. Pada tahap ini akan dilakukan pembangunan aplikasi sesuai dengan tahap analisis dan desain yang sebelumnya telah dilakukan.

4. Pengujian

   Ketika aplikasi selesai dibangun, maka tahap selanjutnya adalah tahap pengujian. Proses pengujian dilakukan untuk memastikan bahwa aplikasi telah sesuai dengan tujuan awal.

1.6 Sistematika Penulisan

Sebagai pedoman penulisan agar penulisan skripsi ini dapat terarah dan tersusun sesuai dengan yang penulis harapkan, maka akan disusun sistematika penulisan sebagai berikut.

BAB 1 PENDAHULUAN

Bab ini berisi uraian latar belakang masalah, identifikasi masalah, maksud dan tujuan, batasan masalah, metodologi penelitian, tahap pengumpulan data, model pembangunan perangkat lunak dan sistematika penulisan.

BAB 2 TINJAUAN PUSTAKA

Bab ini akan membahas berbagai konsep dasar dan teori-teori pendukung yang berhubungan dengan pembangunan aplikasi.

BAB 3 ANALISIS DAN PERANCANGAN SISTEM

Bab ini akan membahas tentang deskripsi sistem, analisis kebutuhan dalam pembangunan aplikasi serta perancangan sistem yang dikembangkan pada aplikasi anotasi *word sense disambiguation* pada teks bahasa Indonesia.

BAB 4 IMPLEMENTASI DAN PENGUJIAN SISTEM

Bab ini berisi hasil implementasi analisis dari BAB 3 dan perancangan aplikasi yang dilakukan, serta hasil pengujian aplikasi untuk mengetahui apakah aplikasi yang dibangun sudah memenuhi kebutuhan.

BAB 5 KESIMPULAN DAN SARAN

Bab ini menguraikan tentang hasil yang didapat dari penelitian yang telah dilakukan, serta masukan untuk pengembangan lebih lanjut agar penelitian yang akan datang dapat memberikan hasil yang lebih baik.