BAB 1
PENDAHULUAN

1.1 Latar Belakang Masalah

Sebagai salah satu dari enam Bahasa resmi yang digunakan oleh PBB sejak 18 Desember [1], Bahasa Arab menjadi salah satu bahasa yang memiliki beragam literatur kajian. Literatur berbahasa arab tersebar diberbagai belahan dunia. Bahasa Arab paling banyak digunakan pada buku-buku rujukan ilmu orang islam yang umumnya disebut kitab kuning, dan banyak dipelajari di Indonesia khususnya pesantren yang membahas ilmu Tafsir, Fiqih, dan Hadits. Selain itu, di Indonesia kitab suci Al-Qur'an digunakan oleh umat muslim yang bertuliskan bahasa arab.

Pada dasarnya untuk memahami bahasa arab juga memiliki cakupan ilmu gramatikal yaitu susunan kalimat sebagaimana pada bahasa lainnya. Pada bahasa arab dikenal dengan *nahwu*, yaitu ilmu yang membahas struktur kalimat dalam Bahasa Arab. Selain ilmu *nahwu*, terdapat pula *shorof* untuk menentukan perubahan kata itu sendiri dan *balaghah* untuk membuat bait kata menjadi lebih ringkas dan indah.

Faiz In'amurrohman, melakukan penelitian tentang yang berfokus pada pembelajaran ilmu *nahwu*. Metode yang digunakan adalah FSA untuk mengenali token-token pertanyaan [14]. Namun demikian, penelitian ini lebih berfokus pada pembelajaran nahwu, bukan mengenai bagaimana penerapan ilmu *nahwu* dalam *computer science.* Secara singkat, ilmu *nahwu* digunakan untuk menentukan *syakl* (baris) akhir dari tiap kata dalam kalimat. Al-Qur'an didalamnya sudah memiliki *syakl* pada setiap hurufnya, sedangkan dalam kitab kuning sama sekali tidak terdapat *syakl* (arab gundul).

Syakl atau baris merupakan tanda pada kosakata arab yang menunjukkan perbedaan makna. Abdulrahman Ahmed Khudhur dalam penelitiannya melakukan penerjemahan AlQuran per kata kedalam bahasa inggris. Penelitiannya menunjukkan hasil dimana dari 11 kata terdapat 7 kata (69%) dengan nilai akurasi sedang dan 4 kata diantaranya (31%) [12]. Meskipun penelitiannya berhubungan

tentang makna, penelitian ini tidak banyak membahas tentang struktur kalimat dari bahasa arab, khususnya pada pembahasan *syakl*.

Penelitian lainnya dilakukan oleh mudafiq Ryan dengan melakukan analisis sintaksis untuk menghilangkan suffix dan prefix pada kata Bahasa Arab, sehingga dapat mengenali jenis dari kata tersebut. Penelitian yang menggunakan metode *Light Stemming* ini mampu mendeteksi kata dengan tingkat keberhasilan sebesar 82.22%. Tingkat kegagalan 17.78% karena terjadi bias apakah kata tersebut merupakan kata kerja atau kata benda [13]. Penelitian ini membahas tentang prediksi kategori kata dengan memanfaatkan ciri-ciri kata tersebut. Penelitian ini belum membahas bagaimana gramatikal bahasa arab sehingga dapat menemukan *syakl* akhir kata. Meskipun begitu, penelitian ini menjadi rujukan untuk menyelesaikan salah satu proses dalam *natural language processing*, yaitu *stemming*.

Syakl akhir sebagai bahasan ilmu *nahwu* sangat penting untuk diketahui. Suatu kata dapat berubah kedudukannya dari subjek menjadi objek hanya dengan melihat *syakl* akhirnya saja. Meskipun begitu, mengetahui *syakl* akhir diperlukan pemahaman tentang kaidah atau aturan ilmu *nahwu*. Hal ini menjadi tantangan bagi para pelajar yang menggunakan literatur bahasa arab, karena literatur arab kebanyakan tidak memiliki *syakl*. Terlebih lagi, terdapat literatur arab yang tidak seuai dengan *syakl* yang semestinya. Oleh karenanya, peneliti ingin melakukan penelitian tentang "Impelementasi *Natural Language Processing* dalam penggunaan ilmu *nahwu* pada bahasa arab" dimana sebagai bentuk representasi ilmu *nahwu* dengan memprediksi *syakl* akhir pada setiap kata dalam bahasa arab dengan menggunakan metode *rule based*.

1.2 Rumusan Masalah

Berdasarkan latar belakang sebelumnya, dapat dirumuskan masalah dalam penelitian ini, yaitu bagaimana melakukan prediksi *syakl* akhir kata, dan seberapa menerapkan dengan metode *rule based* yang merujuk pada kaidah ilmu *nahwu*.

1.3 Maksud dan Tujuan

Pada bagian ini dijelaskan maksud dan tujuan dilakukannya penelitian berjudul Implementasi *Natural Language Processing* dalam Penggunaan Ilmu *Nahwu* Pada Bahasa Arab ini.

1.3.1 Maksud

Berdasarkan permasalahan pada identifikasi masalah, maka maksud dari penulisan tugas akhir ini adalah memprediksi *syakl* (baris) akhir setiap kata, serta menentukan keterangan kedudukan kata sesuai grammatikal yang berlaku dengan menggunakan metode *rule based*.

1.3.2 Tujuan

Tujuan utama dari penelitian ini secara umum untuk mengimplementasikan *natural language processing* dalam penggunaan ilmu *nahwu*, Adapun secara khusus tujuan penelitian ini adalah:

1. Memprediksi *syakl* (baris *harokat*) akhir setiap kata, serta menentukan keterangan kedudukan kata sesuai grammatikal yang berlaku dengan menggunakan metode *rule based*.
2. Mengetahui akurasi hasil prediksi *syakl* (baris *harokat*) akhir dengan menggunakan metode *rule based*.

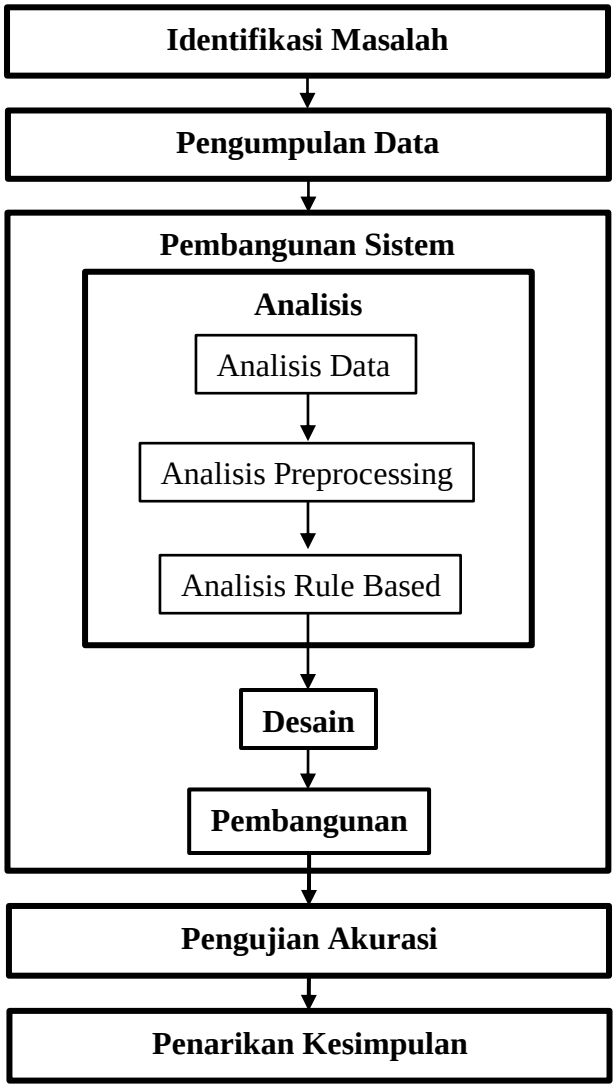
1.4 Batasan Masalah

Pada penelitian ini terdapat batasan masalah yang tidak menjadi bagian utama dalam penyelesaian masalah berdasarkan identifikasi masalahnya. Adapun Batasan masalah tersebut diantaranya

1. Bahasa arab yang digunakan telah mengikuti kaidah ilmu *nahwu.*
2. Pengolahan kata tidak meliputi penggunaan *tasydid*, *tanwin*, serta *illat* bertemunya dua sukun.
3. Pemberian label pada kata bergantung pada keberadaan kata dalam kamus.
4. Keterangan kata yang *mabniy* (tidak berubah *syakl*) tidak disertai *mahal* (kedudukannya dalam kalimat).
5. Pengujian teks dilakukan dengan mengambil teks dari kitab Bahasa Arab, dan membandingkan hasil prediksi dengan *syakl* secara manual.

1.5 Metodologi Penelitian

Peneliti menggunakan metode kualitatif pada penelitian berjudul "Implementasi *Natural Language Processing* dalam penggunaan ilmu *nahwu* pada Bahasa Arab" ini. Metode ini menggambarkan secara sistematis fakta dan karakteristik objek dan subjek yang diteliti secara tepat [2]. Alur penelitian dapat dilihat pada Gambar berikut:



Gambar 0.1 Alur Penelitian

1.5.1 Identifikasi Masalah

Identifikasi masalah dilakukan untuk mengkaji tentang beberapa penelitian berkaitan dengan *natural language processing* Bahasa Arab yang telah dilakukan

sebelumnya. Hasil dari beberapa penelitian tersebut diidentifikasi sehingga dapat ditemukan metode yang cukup efektif untuk melakukan implementasi *natural language processing* pada Bahasa Arab.

1.5.2 Pengumpulan Data

Penelitian ini membutuhkan data dan informasi sebagai landasan dalam menemukenali solusi penanganan masalah yang telah diidentifikasi. Adapun metode pengumpulan data pada penelitian ini dilakukan dengan cara sebagai berikut :

1. Studi Literatur

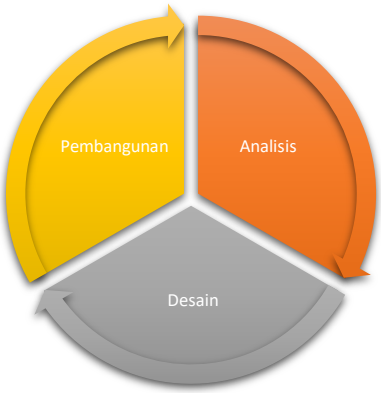
   Pengumpulan data dilakukan dengan cara mempelajari, meneliti, dan menelaah berbagai literatur dari perpustakaan yang bersumber dari buku-buku, jurnal ilmiah, situs internet, dan bacaan lainnya yang berkaitan dengan penelitian yang dilakukan.

2. Wawancara

   Wawancara akan dilakukan kepada professional didalam bidangnya, dalam hal ini adalah pengajar Bahasa arab. Tujuannya adalah untuk menyesuaikan bahasa manusia (dalam hal ini Bahasa arab) kedalam bahasa yang dapat dibaca oleh komputer.

1.5.3 Pembangunan Sistem

Pembangunan sistem ini akan dilakukan dengan metode pengembangan sistem yang cukup familiar bagi pengembang aplikasi yaitu pendekatan *Agile Development*. Model ini membagi proyek kedalam beberapa bagian pengembangan dalam interval waktu tertentu. *Agile* baik digunakan dalam pengembangan apikasi yang ingin dilakukan dalam waktu yang relatif cepat. Berikut merupakan gambaran metode *Agile Development*:



Gambar 0.2 Metode Agile

1. Analisis

   Pada tahapan ini dilakukan pencarian dan penentuan pokok permasalahan yang terjadi berdasarkan latar belakang masalah dalam penelitian ini. Proses *analysis* merupakan perencanaan pemecahan permasalahan dan langkah-langkah yang dilakukan untuk pengumpulan data-data yang diperlukan. Tahapan ini meliputi Analisis data, preprocessing, dan translasi

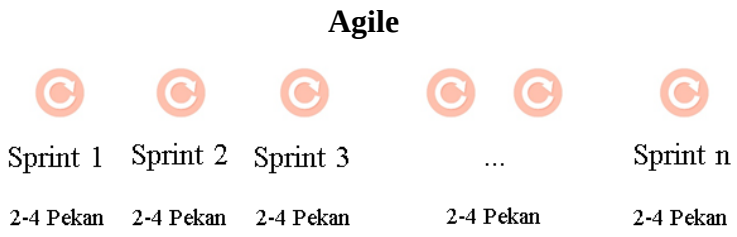
2. Desain

   Proses desain adalah proses pembuatan desain sistem dengan menggunakan diagram-diagram pemodelan sistem seperti *Use Case Diagram*, Class Diagram, Activity Diagram, dan yang lainnya, serta desain *user interface* sistem.

3. Pembangunan

   Pembangunan merupakan proses pengembangan software dengan melakukan *codin*g yaitu pembuatan sistem program dengan menggunakan bahasa pemrograman Python.

Setiap kebutuhan dikerjakan secara bergantian dalam sprint tertentu dengan interval waktu tertentu yang ditentukan. Sebagaimana diilustrasikan berdasarkan pernyataan Matt LeMay menjadi sebagai berikut [3] :



Gambar 1.3 Agile Sprint

1.5.4 Pengujian Akurasi

Tahapan pengujian akurasi dilakukan untuk memeriksa akurasi dari hasil implementasi aplikasi, sehingga dengannya peneliti dapat menemukan kesalahan dan ketidaksesuaian penggunaan metode rule based ini untuk memprediksi *syakl* pada Bahasa Arab. Pengujian dilakukan dengan meneggunakan sejumlah teks yang diambil dari berbagai buku teks, dalam hal ini berupa kitab Bahasa Arab.

1.5.5 Penarikan Kesimpulan

Pada tahap ini peneliti menarik kesimpulan berdasarkan hasil pengujian terhadap perangkat lunak yang telah dilakukan. Setelah itu, selanjutnya peneliti juga akan memberikan saran untuk penelitian selanjutnya berdasarkan hasil evaluasi.

1.6 Sistematika Penulisan

Skripsi ini disusun untuk memberikan gambaran umum tentang masalah yang akan dipecahkan. Skripsi ini akan dijadikan langkah awal untuk menyusun laporan penelitian yang lebih mendalam. Sistematika pada penyusunan laporan penelitian nantinya adalah sebagai berikut :

BAB 1 PENDAHULUAN

Pada BAB I membahas tentang latar belakang, identifikasi masalah, batasan masalah atau ruang lingkup penelitian, maksud dan tujuan penelitian, metodologi penelitian serta sistematika penulisan.

BAB 2 TINJAUAN PUSTAKA

Pada BAB II dijelaskan teori-teori yang mendukung dalam penelitian yang bersumber dari jurnal-jurnal pendukung serta buku-buku literatur. Teori yang dicantumkan dalam laporan hanya teori-teori mendukung yang terkait dengan penelitian.

BAB 3 ANALISIS DAN PERANCANGAN

Pada BAB 3 akan membahas tentang sejarah singkat studi kasus, analisis sistem, analisis data masukan, analisis preprocessing, tokenisasi, hingga analisis pengguna. Pembuatan analisis sistem dapat didukung dengan diagram alir untuk memudahkan pemahaman alur proses, serta pembahasan tentang perancangan sistem yang akan dibuat.

BAB 4 IMPLEMENTASI DAN PENGUJIAN

Pada BAB IV membahas implementasi antarmuka sistem, implementasi sistem, pengujian sistem untuk mengetahui tingkat akurasi .

BAB 5 KESIMPULAN DAN SARAN

Pada BAB V membahas hasil akhir berupa kesimpulan serta saran. Kesimpulan, didalamnya membahas keterangan ringkas dari hasil penelitian yang akan menjawab pokok permasalahan. Saran yang dimaksud pada laporan adalah saran terhadap pembaca untuk yang ingin mengembangkan topik penelitian ini, dengan menunjukkan kekurangan dari sistem yang telah dibuat. Harapannya agar pihak lain tersebut dapat menyempurnakan sistem sehingga bisa menjadi lebih bermanfaat bagi orang banyak.