

1. Pendahuluan

Dalam bagian ini terdiri dari latar belakang, perumusan masalah, batasan masalah, tujuan, metodologi penyelesaian masalah, dan organisasi tulisan.

1.1 Latar Belakang

Semakin luas penyebaran ajaran Islam di dunia ini, maka akan semakin banyak orang-orang yang mempelajari pedoman dari ajaran tersebut, yaitu salah satunya ialah Al-Qur'an Al-Karim. Al-Qur'an merupakan kitab suci umat Islam yang merupakan kumpulan firman-firman Allah yang diturunkan kepada Nabi Muhammad Sallahu 'Alaihi Wasallam. Tujuan utama diturunkan Al-Qur'an adalah untuk menjadi pedoman manusia dalam menata kehidupan supaya memperoleh kebahagiaan di dunia dan di akhirat [1].

Karena Al-Qur'an diturunkan kepada Rasulullah menggunakan bahasa Arab (QS. Yusuf [12] : 2), maka dibuatlah Al-Qur'an dalam berbagai versi terjemahan. Walaupun begitu, banyak orang yang bukan berasal dari Arab bahkan ingin mempelajari Al-Qur'an dari bahasa asalnya guna mengetahui makna yang terkandung didalam bahasa tersebut. Namun untuk orang-orang yang baru mempelajari bahasa Arab harus paham arti dari setiap kata yang ada pada Al-Qur'an dan hal itu membutuhkan ilmu *al-sharf*, yaitu ilmu yang mempelajari seluk-beluk bentuk kata dalam bahasa Arab [2].

Dengan adanya perkembangan teknologi, dapat dibuat klasifikasi ayat-ayat pada Al-Quran dengan *text classification*. *Text classification* adalah sub bidang dari *text mining* yang bertujuan untuk mengklasifikasikan kumpulan teks dari beberapa dokumen ke dalam kategori-kategori tertentu. Sebelum teks tersebut dapat diklasifikasi maka perlu dilakukan *pre-processing* terlebih dahulu, yang meliputi : *tokenization*, *filtering*, *stemming*, dan *weighting* [3]. Dalam hal ini, penulis secara khusus mencoba meneliti proses *stemming* karena *stemming* merupakan langkah penting dalam pemrosesan kata dalam bahasa Arab, terutama untuk Al-Quran. Disamping itu juga dibutuhkan algoritma yang tepat untuk mendapatkan akurasi hasil *stemming* yang baik guna mempermudah orang-orang dalam mempelajari Al-Quran.

Algoritma Khoja Stemmer adalah salah satu *stemmer* bahasa Arab yang terkenal dan banyak digunakan. Cara kerja dari algoritma Khoja Stemmer ialah dengan menghilangkan akhiran terpanjang dan awalan terpanjang dari sebuah kata. Kemudian mencocokkan kata yang tersisa dengan pola verbal dan kata benda (*noun*), untuk mendapatkan akar dari kata [4]. Oleh karena itu, peneliti membangun aplikasi *stemming* pada Al-Quran menggunakan algoritma Shereen Khoja Stemmer untuk mengetahui apakah algoritma tersebut cocok digunakan untuk melakukan *stemming* pada Al-Quran sehingga mempermudah pelajar dalam memahami morfologi atau *sharf* pada Al-Quran.

1.2 Perumusan Masalah

Dari latar belakang yang telah disampaikan diatas, dapat dirumuskan masalah yaitu bagaimana tahapan implementasi Khoja Stemmer dalam melakukan *stemming* pada Al-Quran, bagaimana hasil Khoja Stemmer dalam melakukan *stemming* pada Al-Qur'an, dan bagaimana akurasi Khoja Stemmer dalam melakukan *stemming* pada Al-Qur'an.

1.3 Batasan Masalah

Adapun batasan masalah pada tugas akhir ini yaitu dataset Al-Quran berupa 30 juz digunakan untuk menentukan akurasi sistem secara keseluruhan, sedangkan analisisnya hanya mengambil beberapa ayat awal pada dataset dan pendekatan yang digunakan ialah berupa *root-based stemmers*.

1.4 Tujuan

Tugas akhir ini memiliki beberapa tujuan, yaitu mengetahui tahapan implementasi Khoja Stemmer dalam melakukan *stemming* pada Al-Quran, mengetahui proses Khoja Stemmer dalam melakukan *stemming* pada Al-Quran, mengetahui akurasi Khoja Stemmer dalam melakukan *stemming* pada Al-Quran, menganalisa hasil Khoja Stemmer dalam melakukan *stemming* pada Al-Quran, dan mengetahui apakah Khoja Stemmer cocok digunakan untuk membangun aplikasi *stemming* pada Al-Quran.

1.5 Metodologi Penyelesaian Masalah

Metode penyelesaian masalah dalam tugas akhir ini adalah sebagai berikut:

1. Kajian Pustaka
Pada tahap ini dilakukan pencarian dan pemahaman sumber yang diperlukan untuk pengumpulan data dan perancangan sistem yang akan dibuat. Sumber didapat dengan mencari ilmu, wawasan dari buku, jurnal, paper, perpustakaan mengenai konsep Shereen Khoja Stemmer, *stemming* pada Al-Qur'an, dan *stemming* pada bahasa arab dalam menunjang pengerjaan tugas akhir.
2. Pengumpulan Data
Pada tahap ini dilakukan pengumpulan data Al-Quran dan kamus akar kata yang dijadikan sebagai dataset..
3. Perancangan dan Implementasi
Pada tahap ini dilakukan perancangan dan pembuatan aplikasi *stemming* pada Al-Quran yang menerapkan algoritma Shereen Khoja Stemmer.
4. Pengujian dan Analisis
Pada tahap ini dilakukan pengujian dan analisis terhadap hasil *stem* pada aplikasi *stemming* Al-Qur'an yang menerapkan algoritma Shereen Khoja Stemmer serta menganalisa akurasi sistem yang dihasilkan.
5. Penyusunan Laporan Tugas Akhir
Pada tahap ini dilakukan penyusunan laporan tugas akhir sebagai dokumentasi berdasarkan penelitian yang telah dilakukan serta mengambil kesimpulan dari hasil penelitian ini.

1.6 Organisasi Tulisan

Pada tugas akhir ini memiliki susunan sebagai berikut:

1. Pendahuluan, yang terdiri dari latar belakang, perumusan masalah, batasan masalah, tujuan, metodologi penyelesaian masalah, dan organisasi tulisan.
2. Studi terkait, yaitu teori / studi / literatur yang mendukung (terkait erat) dengan topik Tugas Akhir. Bagian ini terdiri dari pengertian *stemming*, *stemming* pada bahasa arab; yang terdiri dari morfologi pada bahasa arab dan pendekatan pada *stemming* bahasa arab, dan *stemming* pada Al-Quran.
3. Sistem yang Dibangun, yaitu rancangan sistem yang dibuat dan tahapan implementasi yang dilakukan guna membangun aplikasi *stemming* pada Al-Quran menggunakan algoritma Shereen Khoja Stemmer. Bagian ini terdiri dari pendekatan atau algoritma yang digunakan, dataset yang digunakan, metrik evaluasi dan *baseline*, serta gambaran umum sistem yang dibangun; yang terdiri dari membaca dataset input, tokenisasi, mencari *stem / root*, dan akurasi sistem.
4. Evaluasi, yang berisi hasil pengujian dan analisis, serta performansi pada aplikasi *stemming* Al-Quran yang menggunakan algoritma Shereen Khoja Stemmer. Bagian ini terdiri dari pengujian sistem; yang mencakup tujuan pengujian, skenario data, dan skenario pengujian, spesifikasi kebutuhan sistem; yang mencakup spesifikasi perangkat keras dan spesifikasi perangkat lunak, dan hasil pengujian; yang mencakup pengujian *stemming*, *stemmed words*, *words not stemmed*, *stopwords*, dan akurasi sistem.
5. Kesimpulan, yang berisi kesimpulan hasil penelitian pada tugas akhir ini dan saran guna mengembangkan aplikasi *stemming* Al-Quran menggunakan algoritma Shereen Khoja Stemmer menjadi lebih baik lagi kedepannya.
6. Daftar pustaka yang berisi referensi yang digunakan dalam mengerjakan tugas akhir.
7. Lampiran yang berisi dataset yang digunakan serta rincian pengujian yang dilakukan.