

1. Pendahuluan

Latar Belakang

Indonesia merupakan salah satu negara dengan populasi muslim terbesar di dunia [1]. Pemeluk agama Islam memiliki kitab yang diwajibkan untuk dibaca setiap harinya, yaitu Al-Qur'an. Pentingnya membaca Al-Qur'an tidak hanya untuk menambah amal dan pahala bagi seorang individu, melainkan untuk dijadikan pedoman hidup sehari-hari bagi umat Islam. Perkembangan agama islam semakin luas, telah terbukti pada jumlah umat muslim yang bertambah sebanyak 2,9 persen pertahun [6]. Maka dari itu, semakin banyak insan yang memilih Al-Qur'an sebagai pedoman hidupnya. Dikarenakan Al-Qur'an memiliki banyak surat dan ayat di dalamnya, alangkah lebih baik jika proses pembelajaran Al-Qur'an dibantu dengan alat otomatis yang memanfaatkan NLP.

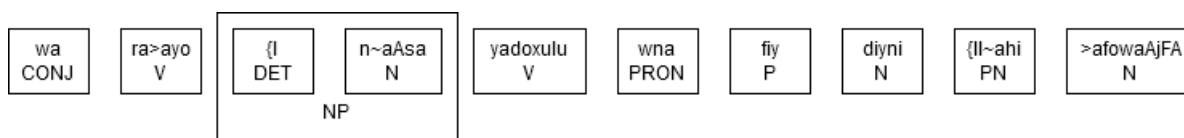
Bahasa Arab menimbulkan banyak tantangan untuk *Natural Language Processing (NLP)*, karena secara morfologis Bahasa Arab merupakan bahasa yang rumit [4]. Bahasa Arab kurang memiliki sumber daya dan alat beranotasi [8]. Pada penelitian kali ini akan mencoba melakukan *noun phrase chunking* otomatis terhadap ayat-ayat yang terdapat di dalam Al-Qur'an. Sistem *chunking* bisa menggunakan *rule-based* atau bisa juga menggunakan *machine learning* [2]. Pada penelitian ini, penulis menggunakan pendekatan *rule-based* karena pendekatan tersebut mudah beradaptasi dengan pola *chunking* yang berbeda [2]. Tujuan dari *noun phrase chunking* ini sendiri untuk mengetahui dan memahami struktur dari sebuah ayat dalam Al-Qur'an dan diharapkan dapat mempermudah orang awam untuk mengerti dan mempelajari Al-Qur'an secara keseluruhan.

Topik dan Batasannya

Pada penelitian kali ini akan dilakukan *rule-based noun phrase chunking* otomatis terhadap ayat-ayat pada Al-Qur'an yang akan menghasilkan potongan-potongan ayat berupa *noun phrase* entitas 'orang' beserta dengan aturan tata bahasa per kata dari ayat tersebut.

Noun phrase chunking yang akan dilakukan dalam penelitian kali ini dibatasi dengan hanya mencari kumpulan kata yang membentuk *noun phrase* dalam suatu ayat serta karakter yang diambil merupakan bahasa latinnya. Dan pada penelitian ini hanya dilakukan sampai percobaan *noun phrase chunking* saja, serta *noun phrase chunking* yang akan dicari adalah *noun* entitas 'orang'. Ada juga *data test* dan *data train* yang hanya mengambil juz 30 pada Al-Qur'an.

Gambar 1 merupakan contoh dari *noun phrase chunking* terhadap entitas 'person'.



Gambar 1. Contoh *noun phrase chunking*.

Gambar 1 merupakan potongan dari surat An-Nasr ayat kedua. Terdapat potongan kata '{l n aAsa' yang berarti 'seorang manusia' diberi label sebagai satu *chunk noun phrase* entitas 'orang'. Bisa dilihat juga salah satu contoh dari *rule-based chunking* yang penulis buat adalah gabungan dari DET dan N sebagai sebuah *chunk noun phrase* entitas 'orang'.

Tujuan

Tujuan dari penelitian ini adalah untuk mencari tahu seberapa akurat metode *rule-based chunking* dapat diterapkan pada Bahasa Arab khususnya dalam Al-Qur'an.

Organisasi Tulisan

Selanjutnya akan dijelaskan apa itu *chunking*, *noun phrase chunking*, *flowchart* yang menunjukkan sistem yang dibangun serta data apa saja yang akan dipakai, lalu diakhiri dengan evaluasi serta analisis dari hasil penerapan sistem yang telah dibangun.