

## 1. Pendahuluan

Pada bagian ini akan di jelaskan beberapa hal penting terkait tentang latar belakang, topik dan batasan, tujuan dan rencana kegiatan yang akan di kerjakan pada tugas akhir ini.

### Latar Belakang

Al-Qur'an adalah salah satu kitab suci atau buku umat Islam yang menjadi petunjuk setiap manusia. Sekitar 24% daripada populasi masyarakat dunia yang merupakan umat Islam [1] membaca dan mempelajari tentang Al-Qur'an yang di dalamnya terdapat banyak pengetahuan atau informasi yang tersembunyi.

Al-Qur'an memiliki 30 juz, 114 surat dan lebih dari 6000 ayat yang terkandung di dalamnya [2]. Dari isi Al-Qur'an yang begitu banyak, membuat orang awam kesulitan untuk menemukan dan mempelajari tentang keterkaitan antar kata yang ada di dalam Al-Qur'an . Permasalahan tersebut dapat diatasi dengan adanya sebuah kamus, ensiklopedia atau tesaurus kosa kata Al-Qur'an dan sebagai salah satu kelengkapannya yaitu tiap entri kata memiliki keterkaitan dengan kata lain. T ini membahas tentang keterkaitan antar kata dalam Al-Qur'an dan untuk lebih lanjut diharapkan dapat membantu dalam mencari kemiripan antar ayat Al-Qur'an . Seperti contoh berikut ini: kata مَغْرُوفٌ (perbuatan baik) yang memiliki keterkaitan dengan kata عَفَا (memaafkan) karena dalam Al-Qur'an kedua kata memiliki keterkaitan yaitu salah satu perbuatan baik adalah memaafkan.

Keterkaitan dan kesamaan semantik berkaitan dengan bidang linguistik khususnya pada *Natural Language Processing* (NLP) yang belakangan ini menjadi topik yang menarik dan banyak diteliti. Kesamaan semantik dan keterkaitan kata memiliki peran penting dalam beberapa task dari NLP dan beberapa bidang terkait seperti *text classification*, *document clustering*, *text summarization*, dan lain sebagainya [3] . Pada penelitian sebelumnya yang serupa hanya dilakukan analisis pada pencarian kata-kata yang mengandung atau menginterpretasikan manusia [4], kesamaan semantik antar teks terjemahan Inggrisnya [5]. Pada penelitian tersebut tidak terdapat penjelasan lebih mengenai keterkaitan antar kata dalam Al-Qur'an.

Melengkapi penelitian sebelumnya, maka tugas akhir ini menganalisis ke level keterkaitan antar kata dalam Al-Qur'an . Pembangunan sistem ini membutuhkan beberapa data seperti lema yang ada di dalam Al-Qur'an . Terdapat berbagai metode yang dapat digunakan untuk menganalisis keterkaitan antar kata. Salah satunya adalah menggunakan pendekatan kesamaan distribusional berbasis vektor, karena setiap kata memiliki ketergantungan terhadap jarak atau sudut yang dibangun antara vektor kata tersebut untuk dapat menghasilkan evaluasi kualitas representasi kata yang tinggi. Faktor tersebut yang menjadi alasan pengambilan basis vektor sebagai metode untuk mencari atau menghitung keterkaitan antar kata dalam Al-Qur'an . Keterkaitan antar kata tersebut dapat diketahui dari perhitungan nilai perbedaan sudut vektor kata-kata tersebut dengan menggunakan *cosine-similarity*. Sistem yang dibangun diharapkan dapat menghasilkan performansi yang baik berdasarkan nilai korelasi yang dihitung. Nilai korelasi yang dimaksud adalah nilai korelasi yang didapat dari perhitungan *precision* setelah dikoreksi oleh ahli bahasa.

### Topik dan Batasannya

Berdasarkan dari latar belakang yang telah dijelaskan, topik-topik yang di angkat dalam tugas akhir ini sebagai berikut:

1. Merepresentasikan kata dalam Al-Qur'an menggunakan word2vec CBOW
2. Melakukan perhitungan keterkaitan menggunakan cosine-similarity
3. Melakukan perhitungan performansi dari sistem yang telah dibangun menggunakan precision.

Adapun batasan dari permasalahan yang ada pada tugas akhir ini sebagai berikut:

1. Analisis kata-kata yang dilakukan menggunakan basis vektor .
2. Tidak terdapat jenis keterkaitan.
3. Mengabaikan kata yang sama namun memiliki makna yang berbeda.

**Tujuan**

Berikut adalah tujuan yang ingin dicapai pada tugas akhir ini.

1. Mengimplementasikan pendekatan kesamaan distribusional dalam menganalisis keterkaitan antar kata dalam Al-Qur'an;
2. Mengetahui akurasi yang didapat dari pendekatan kesamaan distribusional dalam menganalisis keterkaitan antar kata dalam Al-Qur'an.

**Organisasi Tulisan**

Susunan dari tulisan ini adalah sebagai berikut : Bagian 2 membahas tentang teori, studi dan literatur terkait dengan tulisan ini. Bagian 3 akan dijelaskan sistem yang akan dibangun. Bagian 4 akan terdapat hasil dan evaluasi dari sistem yang di bangun. Dan Bagian 5 akan dijelaskan kesimpulan.