

***CLUSTERING TERJEMAHAN AYAT AL-QUR'AN
MENGUNAKAN METODE *LATENT DIRICHLET ALLOCATION*
(LDA) DENGAN VISUALISASI *COLLAPSIBLE RADIAL TREE****



SKRIPSI

**Disusun Sebagai Salah Satu Syarat
Untuk Memperoleh Gelar Sarjana Komputer
pada Departemen Ilmu Komputer/ Informatika**

Disusun oleh :

DIMAS DWI PRASETYA

24010313130071

**DEPARTEMEN ILMU KOMPUTER/ INFORMATIKA
FAKULTAS SAINS DAN MATEMATIKA
UNIVERSITAS DIPONEGORO**

2017

HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Saya yang bertanda tangan di bawah ini :

Nama : Dimas Dwi Prasetya
NIM : 24010313130071
Judul : *Clustering* Terjemahan Ayat Al-Qur'an Menggunakan Metode *Latent Dirichlet Allocation* (LDA) dengan Visualisasi *Collapsible Radial Tree*

Dengan ini saya menyatakan bahwa dalam tugas akhir/ skripsi ini tidak terdapat karya yang pernah diajukan untuk memperoleh gelar kesarjanaan di suatu Perguruan Tinggi, dan sepanjang pengetahuan saya juga tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain, kecuali yang secara tertulis diacu dalam naskah ini dan disebutkan di dalam daftar pustaka.

Semarang, 10 November 2017



Dimas Dwi Prasetya
NIM. 24010313130071

HALAMAN PENGESAHAN

Judul : *Clustering Terjemahan Ayat Al-Qur'an Menggunakan Metode Latent Dirichlet Allocation (LDA) dengan Visualisasi Collapsible Radial Tree*

Nama : Dimas Dwi Prasetya

NIM : 24010313130071

Telah diujikan pada sidang tugas akhir pada tanggal 30 Oktober 2017 dan dinyatakan lulus pada tanggal 30 Oktober 2017.

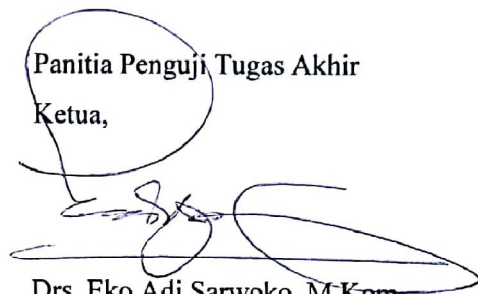
Semarang, 10 November 2017

Mengetahui
Ketua Departemen Ilmu Komputer/ Informatika



Dr. Rano Kusumaningrum, S.Si, M.Kom
NIP. 198104202005012001

Panitia Penguji Tugas Akhir
Ketua,



Drs. Eko Adi Sarwoko, M.Kom
NIP. 196511071992031003

HALAMAN PENGESAHAN

Judul : *Clustering* Terjemahan Ayat Al-Qur'an Menggunakan Metode *Latent Dirichlet Allocation* (LDA) dengan Visualisasi *Collapsible Radial Tree*

Nama : Dimas Dwi Prasetya

NIM : 24010313130071

Telah diujikan pada sidang tugas akhir pada tanggal 30 Oktober 2017.

Semarang, 10 November 2017

Dosen Pembimbing



Dr. Retno Kusumaningrum, S.Si, M.Kom
NIP. 198104202005012001

ABSTRAK

Al-Qur'an adalah kitab suci yang digunakan oleh umat Islam sebagai pedoman hidup, oleh karena itu menjadi suatu kewajiban bagi umat Islam untuk mempelajari Al-Qur'an. Dalam mempelajari Al-Qur'an, seringkali hanya dilakukan dengan memahami setiap ayat pada surat tertentu tanpa memperhatikan ayat-ayat pada surat lainnya. Padahal belum tentu suatu masalah yang dikemukakan di dalamnya berurutan pada satu surat saja. Sebab letak sebagian ayat dan surat dalam Al-Qur'an terpencar satu dengan lainnya, tidak dalam satu bahasan. Hal tersebut menyebabkan diperlukannya penelitian mengenai pengelompokan ayat Al-Qur'an berdasarkan topik, agar lebih memudahkan dalam mempelajari Al-Qur'an. Pada penelitian ini, diusulkan menggunakan metode *Latent Dirichlet Allocation* (LDA) untuk melakukan pengelompokan (*clustering*) terhadap terjemahan ayat Al-Qur'an dalam bahasa Indonesia. LDA akan melakukan *clustering* terhadap ayat-ayat Al-Qur'an berdasarkan tingkat kemiripan kata-kata yang ada pada ayat tersebut kedalam sejumlah topik. LDA akan menghasilkan dua buah keluaran yaitu, *topic proportion* dan probabilitas kata-topik. Agar dapat melihat lebih jelas hasil *clustering* terjemahan ayat Al-Qur'an, dibutuhkan suatu visualisasi. Salah satu visualisasi yang dapat menampilkan hasil *clustering* adalah dalam bentuk *Collapsible Radial Tree*. Tujuan dari penelitian ini adalah untuk menghasilkan *Collapsible Radial Tree* yang menampilkan hasil *clustering* dari terjemahan ayat Al-Qur'an dengan menggunakan kombinasi parameter LDA yang terbaik. Kombinasi parameter terbaik dalam kasus ini adalah dengan menggunakan nilai *alpha* 0.1, *beta* 0.1, dan *threshold* 0.001, sedangkan jumlah topik ditentukan berdasarkan kata unik di dalam Al-Qur'an, sehingga mendapatkan nilai sebesar 12. Kombinasi parameter tersebut menghasilkan nilai *perplexity* sebesar 9,79. Hasil pengujian *usability* pada visualisasi memperoleh nilai untuk indikator *learnability* 92 %, *efficiency* 65.75 %, *memorability* 70 %, *errors* 5 %, dan *satisfaction* 80 %.

Kata kunci : *Latent Dirichlet Allocation, clustering, collapsible radial tree, pengujian usability*

ABSTRACT

Al-Qur'an is a holy book used by Muslims as a guide of life, therefore it becomes an obligation for Muslims to study Al-Qur'an. In studying Al-Qur'an, it is often only done by understanding each verse in a particular surah without regard to the verses in other surah. Though not necessarily a problem put forward in it sequentially on one letter only. The other reason is the location of some verses and surah in Al-Qur'an scattered one with another. This causes the need of research on grouping of verses of Al-Qur'an based on the subject, in order to make it easier for us in studying Al-Qur'an. In this research, we use the *Latent Dirichlet Allocation* (LDA) method to cluster the translation of the Al-Qur'an in Indonesian language. LDA will clustering the verses of the Al-Qur'an according to the degree of resemblance to the words in the verse. LDA will produce two outputs, there are topic proportion and probability of word-topics. In order to see more clear result of clustering translations of the verses of Al-Qur'an, it need a visualization to display the result. One visualization that can display the clustering results is a *Collapsible Radial Tree*. The purpose of this research is to produce *Collapsible Radial Tree* which shows the clustering result from Al-Qur'an verse translation by using the best LDA combination parameter. The best combination of parameters in this case is to use *alpha* value 0.1, *beta* 0.1, and *threshold* 0.001, while the number of topics is determined based on unique words in the Al-Qur'an, so get a value of 12. The combination of parameters yields a *perplexity* value of 9.79. The *usability testing* results on visualization for *learnability* is 92 %, *efficiency* is 65.75 %, *memorability* is 70 %, *errors* is 5 %, dan *satisfaction* is 80 %.

Keywords : *Latent Dirichlet Allocation, clustering, collapsible radial tree, usability testing*

KATA PENGANTAR

Segala puji syukur bagi Allah SWT atas karunia-Nya yang diberikan kepada penulis sehingga penulis dapat menyelesaikan tugas akhir ini. Tugas akhir yang berjudul “*Clustering Terjemahan Ayat Al-Qur’an Menggunakan Metode Latent Dirichlet Allocation (LDA) dengan visualisasi Collapsible Radial Tree*”. Tugas akhir ini disusun sebagai salah satu syarat untuk memperoleh gelar sarjana strata satu pada Departemen Ilmu Komputer/Informatika Fakultas Sains dan Matematika Universitas Diponegoro Semarang.

Dalam penyusunan laporan tugas akhir ini tentulah telah banyak mendapat bantuan dan dukungan dari berbagai pihak. Untuk itu, pada kesempatan ini penulis mengucapkan rasa hormat dan terima kasih kepada:

1. Dr. Retno Kusumaningrum, S.Si, M.Kom, selaku Ketua Departemen Ilmu Komputer / Informatika FSM Universitas Diponegoro Semarang dan selaku dosen Pembimbing yang telah meluangkan waktu dan berkenan memberikan bimbingan, arahan, masukan, serta motivasi yang sangat berharga dan fokus akan tujuan bagi penulis.
2. Helmie Arif Wibawa, S.Si, M.Cs, selaku Koordinator Tugas Akhir Departemen Ilmu Komputer/Informatika FSM Universitas Diponegoro Semarang
3. Orangtua, keluarga, teman dekat dan sahabat yang telah mendukung, membantu dan memberikan semangat kepada penulis dalam menyelesaikan tugas akhir ini.
4. Semua pihak yang tidak dapat disebutkan satu persatu yang telah membantu kelancaran penelitian ini, semoga Tuhan yang memberikan balasan yang lebih baik.

Penulis menyadari bahwa dalam laporan ini masih banyak kekurangan baik dari segi materi ataupun dalam penyajiannya karena keterbatasan kemampuan dan pengetahuan penulis. Oleh karena itu, kritik dan saran sangat penulis harapkan. Semoga laporan ini dapat bermanfaat bagi pembaca pada umumnya dan penulis pada khususnya.

Semarang, 13 Oktober 2017

Penulis

DAFTAR ISI

HALAMAN PERNYATAAN KEASLIAN SKRIPSI.....	ii
HALAMAN PENGESAHAN	iii
HALAMAN PENGESAHAN	iv
ABSTRAK	v
ABSTRACT	vi
KATA PENGANTAR.....	vii
DAFTAR ISI	viii
DAFTAR GAMBAR.....	xi
DAFTAR TABEL	xiii
DAFTAR LAMPIRAN	xv
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	3
1.3 Tujuan dan Manfaat.....	4
1.4 Ruang Lingkup.....	4
1.5 Sistematika Penulisan.....	4
BAB II TINJAUAN PUSTAKA	6
2.1 <i>Clustering</i> Terjemahan Ayat Al-Qur'an.....	6
2.2 <i>Latent Dirichlet Allocation</i>	7
2.3 <i>Radial / Architecture Tree</i>	11
2.4 <i>Preprocessing</i>	12
2.4.1 Tokenisasi	12
2.4.2 <i>Stopword</i> Removal.....	12
2.4.3 <i>Stemming</i>	12
2.5 <i>Perplexity</i>	21
2.6 <i>Usability Testing</i>	22

2.7 Pengembangan Perangkat Lunak	26
BAB III METODOLOGI PENELITIAN	28
3.1 <i>Preprocessing</i>	30
3.1.1 Tokenisasi	31
3.1.2 <i>Stopword Removal</i>	32
3.1.3 <i>Stemming</i>	34
3.2 <i>LDA Collapsed Gibbs Sampling</i>	42
3.3 Evaluasi Kinerja Model LDA	44
3.4 Visualisasi <i>Collapsible Radial Tree</i>	46
3.5 Evaluasi Visualisasi <i>Collapsible Radial Tree</i>	50
3.6 Analisa dan Desain Sistem	52
3.6.1 Analisis Sistem	52
3.6.1.1 Deskripsi Sistem	52
3.6.1.2 Kebutuhan Fungsional Sistem	52
3.6.1.3 Kebutuhan Non-Fungsional Sistem	53
3.6.2 Perancangan Sistem	53
3.6.2.1 Pemodelan Data	53
3.6.2.2 Pemodelan Fungsi	54
3.6.2.3 Perancangan Antarmuka	56
BAB IV HASIL EKSPERIMEN DAN ANALISA	60
4.1 Hasil Pengembangan Sistem	60
4.1.1 Lingkungan Implementasi	60
4.1.2 Implementasi Antarmuka	61
4.2 Skenario Pengujian Sistem	65
4.2.1 Pengujian Fungsional Sistem	65
4.2.2 Pengujian Kinerja Sistem	65
4.2.2.1 Data	65

4.2.2.2	Eksperimen	65
4.3	Hasil dan Analisis Sistem.....	66
4.3.1	Pengujian Fungsional Sistem.....	66
4.3.2	Hasil Eksperimen 1	67
4.3.3	Hasil Eksperimen 2.....	72
BAB V	PENUTUP	79
5.1	Kesimpulan	79
5.2	Saran	79
DAFTAR	PUSTAKA.....	80

DAFTAR GAMBAR

Gambar 2.1 Graphical Model PLSA dan LDA	7
Gambar 2.2 Model LDA generatif dan inferensi (Kusumaningrum, et al., 2014)	8
Gambar 2.3 LDA sebagai generatif (Liu, 2013).....	9
Gambar 2.4 LDA sebagai inferensi (Kusumaningrum, 2014).....	9
Gambar 2.5 Architecture Tree	11
Gambar 2.6 Waterfall Model (Sommerville, 2011)	26
Gambar 3.1 Gambaran Umum Penelitian	28
Gambar 3.2 Flowchart Preprocessing.....	30
Gambar 3.3 Flowchart Tokenisasi.....	31
Gambar 3.4 Flowchart stopword removal	33
Gambar 3.5 Flowchart Stemming Sastrawi (Bashri, 2017).....	34
Gambar 3.6 Flowchart subproses stemmingSastrawi (Bashri, 2017).....	35
Gambar 3.7 Flowchart subproses stemmingPlural (Bashri, 2017).....	36
Gambar 3.8 Flowchart subproses stemmingSingular (Bashri, 2017).....	37
Gambar 3.9 Flowchart loopPengembalianAkhiran (Bashri, 2017).....	39
Gambar 3.10 Flowchart proses LDA <i>Collapsed Gibbs Sampling</i>	43
Gambar 3.11 Flowchart pembentukan file json	47
Gambar 3.12 Struktur file json	48
Gambar 3.13 Rancangan Collapsible Radial Tree.....	48
Gambar 3.14 Rancangan tooltip tree	49
Gambar 3.15 Rancangan pop-up tree	49
Gambar 3.16 Entity Relationship Diagram (ERD).....	53
Gambar 3.17 Data Context Diagram (DCD).....	54
Gambar 3.18 Data Flow Diagram Level 1	55
Gambar 3.19 Antarmuka Halaman Beranda.....	56
Gambar 3.20 Antarmuka halaman data	56
Gambar 3.21 Antarmuka Halaman Preprocessing	57
Gambar 3.22 Antarmuka halaman LDA.....	58
Gambar 3.23 Antarmuka halaman visualisasi Collapsible Radial Tree	59
Gambar 3.24 Antarmuka Halaman visualisasi Collapsible Radial Tree detail	59

Gambar 4.1 Implementasi Antarmuka Beranda	61
Gambar 4.2 Implementasi Antarmuka Data	61
Gambar 4.3 Implementasi Antarmuka Preprocessing	62
Gambar 4.4 Implementasi Antarmuka LDA	63
Gambar 4.5 Implementasi Antarmuka Visualisasi Tree.....	64
Gambar 4.6 Implementasi Antarmuka Visualisasi Tree Detail.....	64
Gambar 4.7 Grafik nilai perplexity untuk seluruh kombinasi parameter	68
Gambar 4.8 Contoh grafik probabilitas kata dari parameter K-03	69
Gambar 4.9 Grafik grafik probabilitas kata dari parameter K-22	69
Gambar 4.10 Visualisasi Collapsible Radial Tree untuk kombinasi K-03 level topik.....	70
Gambar 4.11 Visualisasi Collapsible Radial Tree untuk kombinasi K-03 level surat	70
Gambar 4.12 Visualisasi Collapsible Radial Tree untuk kombinasi K-03 level ayat	71
Gambar 4.13 Pop-up salah satu node ayat untuk kombinasi K-03.....	71
Gambar 4.14 Hasil Post-test Kuesioner.....	74
Gambar 4.15 Persentase indikator usability	78

DAFTAR TABEL

Tabel 2.1	State of The Art	6
Tabel 2.2	Kombinasi awalan-akhiran yang tidak diperbolehkan	14
Tabel 2.3	Aturan pemenggalan awalan algoritma <i>Stemming</i> Nazief dan Adriani.....	15
Tabel 2.4	Modifikasi dan penambahan aturan pemenggalan awalan oleh Algoritma <i>Stemming Confix Stripping</i>	17
Tabel 2.5	Daftar aturan <i>Rule Precedence</i>	18
Tabel 2.6	Modifikasi aturan pemenggalan awalan oleh Algoritma <i>Stemming Enhanced Confix Stripping</i> (Tahitoe & Purwitasari, 2010).....	19
Tabel 2.7	Modifikasi aturan pemenggalan awalan dan penambahan aturan pemenggalan sisipan oleh algoritma <i>Stemming modified Enhanced Confix Stripping</i>	21
Tabel 2.8	Penambahan dan modifikasi aturan pemenggalan awalan <i>stemmer</i> Sastrawi..	21
Tabel 2.9	Skor Jawaban Responden	24
Tabel 2.10	Rekapitulasi hasil kuesioner	24
Tabel 2.11	Keterangan Persentase tanggapan responden	25
Tabel 2.12	Contoh rekapitulasi hasil kuesioner.....	25
Tabel 3.1	Proses Tokenisasi	32
Tabel 3.2	Daftar <i>Stopword</i>	32
Tabel 3.3	Proses Penghilangan <i>Stopword</i>	33
Tabel 3.4	Proses <i>Stemming</i>	40
Tabel 3.5	Bentuk	41
Tabel 3.6	<i>Vocabulary</i>	41
Tabel 3.7	Contoh Probabilitas kata-topik	43
Tabel 3.8	Contoh <i>Topic Proportion</i>	44
Tabel 3.9	Aspek penilaian indikator <i>usability</i>	51
Tabel 3.10	Kebutuhan Fungsional Sistem.....	52
Tabel 3.11	Kebutuhan Non Fungsional Sistem.....	53
Tabel 4.1	Rencana Pengujian Fungsional Sistem.....	66
Tabel 4.2	Daftar Kombinasi Parameter	67
Tabel 4.3	Responden pengujian <i>Usability</i>	72
Tabel 4.4	Hasil Post-Test Kuesioner	73
Tabel 4.5	Penyelesaian pengerjaan Task.....	76

Tabel 4.6 Waktu eksekusi pengerjaan task..... 76

DAFTAR LAMPIRAN

Lampiran 1. Perhitungan LDA <i>Collapsed Gibbs Sampling</i>	85
Lampiran 2. Lembar Pengujian <i>Usability</i>	98
Lampiran 3. Deskripsi dan Hasil Pengujian Fungsional Sistem	101

BAB I

PENDAHULUAN

Bab ini membahas latar belakang masalah, rumusan masalah, tujuan dan manfaat, serta ruang lingkup penelitian dan sistematika penulisan tugas akhir mengenai *Clustering* Terjemahan Ayat Al-Qur'an Menggunakan Metode *Latent Dirichlet Allocation* dengan Visualisasi *Collapsible Radial Tree*.

1.1 Latar Belakang

Al-Qur'an adalah kitab suci yang digunakan oleh umat Islam sebagai pedoman hidup, sehingga menjadi suatu kewajiban bagi umat Islam untuk mempelajari Al-Qur'an. Dalam mempelajari Al-Qur'an, seringkali hanya dilakukan dengan memahami setiap ayat pada surat tertentu tanpa memperhatikan ayat-ayat pada surat lainnya, padahal belum tentu suatu masalah yang dikemukakan di dalamnya berurutan pada satu surat saja. Sebab yang lain adalah letak sebagian ayat dan surat dalam Al-Qur'an terpencar satu dengan yang lainnya, tidak dalam satu bahasan (Choiruddin, 2005). Mengetahui adanya hubungan antara ayat-ayat dan surat-surat dapat membantu untuk memahami dengan tepat ayat-ayat dan surat-surat yang bersangkutan (Yusuf, 2012). Hal tersebut menyebabkan diperlukannya penelitian mengenai pengelompokan ayat Al-Qur'an berdasarkan topik, agar lebih memudahkan dalam mempelajari Al-Qur'an.

Salah satu cara untuk melakukan pengelompokan (*clustering*) ayat di Al-Qur'an adalah berdasarkan topiknya, yang dapat dipecahkan dengan menggunakan *topic modeling*. *Topic Modeling* adalah kumpulan algoritma yang digunakan untuk menemukan struktur tersembunyi dari tema yang terdapat dalam setiap dokumen (Blei, 2012). *Topic modeling* memiliki beberapa model, yaitu *linear topic modeling* dan *probabilistic topic modeling* (Pleplé, 2013). *Probabilistic topic modeling* merupakan pengembangan dari metode *linear topic modeling*, yaitu *Latent Semantic Indexing* (LSI) (Liu, 2013). LSI mengasumsikan bahwa kata-kata dalam suatu dokumen akan mempengaruhi topik yang dapat digunakan sebagai rujukan ke dokumen itu sendiri (Deerwester, et al., 1990). Sedangkan *probabilistic topic modeling* mendefinisikan topik sebagai distribusi berdasarkan kata-kata dan

mendefinisikan dokumen sebagai campuran topik (Blei, 2012). Oleh karena itu, pada penelitian ini akan dilakukan *clustering* mengenai terjemahan ayat Al-Qur'an dengan menggunakan *probabilistic topic modeling* sebagai salah satu bentuk *topic modeling*. *Probabilistic topic modeling* merupakan sebuah algoritma yang berfungsi untuk menemukan topik utama dalam suatu dokumen dan memberi label pada dokumen tersebut (Blei, 2012).

Terdapat beberapa metode dari *probabilistic topic modeling* diantaranya, yaitu *Probabilistic Latent Semantic Analysis* (PLSA) dan *Latent Dirichlet Allocation* (LDA). LDA merupakan pengembangan dari PLSA (Blei, et al., 2003) yang lebih stabil untuk mengolah data dalam jumlah besar (Liu, 2013). Metode LDA mengasumsikan bahwa pada satu dokumen terdapat lebih dari satu topik, yang masing-masing merupakan distribusi melalui kosakata. Metode LDA dapat diterapkan sebagai proses generatif maupun proses inferensi. LDA sebagai proses generatif akan menghasilkan sebuah korpus dari probabilitas kata-topik dan *topic proportion* yang telah diketahui sebelumnya. Sedangkan LDA sebagai proses inferensi akan menghasilkan probabilitas kata-topik dan *topic proportion* dari korpus yang digunakan sebagai masukan.

Pada penelitian ini akan diterapkan LDA dengan proses inferensi. Metode LDA dengan proses inferensi, akan menghasilkan *topic proportion* dan probabilitas kata-topik berupa *matrix*, oleh karena itu perlu dilakukan visualisasi agar lebih mudah membaca hasil dari analisis *topic modeling* menggunakan LDA ini. Beberapa penelitian LDA memvisualisasikan hasilnya dalam bentuk grafik, seperti pada penelitian *Fuzzy-Gibbs Latent Dirichlet Allocation Model for Feature Extraction on Indonesian Documents* yang dilakukan oleh Prihatini dan kawan-kawannya pada tahun 2017 dengan menggunakan grafik batang, lalu pada penelitian yang dilakukan oleh Zulhanif pada tahun 2016 yang memvisualisasikan hasil LDA menggunakan grafik batang dan *wordcloud*, lalu pada analisis dan visualisasi suara pelanggan pada pusat pelayanan pelanggan dengan pemodelan topik menggunakan LDA studi kasus: PT. Petrokimia Gresik oleh Agustina pada tahun 2017, yang menggunakan visualisasi grafik batang dan garis yang menampilkan topik dan probabilitas yang didapatkan dari LDA. Pada ketiga penelitian tersebut, hasil dari LDA divisualisasikan dengan grafik yang digunakan untuk menunjukkan probabilitas dari setiap kata, sedangkan pada penelitian ini ingin menunjukkan hasil *clustering* dari

sejumlah dokumen yang diberikan. Sehingga kurang cocok apabila menggunakan visualisasi dalam bentuk grafik.

Pada penelitian ini menggunakan data berupa terjemahan ayat Al-Qur'an, sehingga dokumen yang akan di proses adalah sebanyak jumlah ayat dalam Al-Qur'an, yaitu 6,236 dokumen. Karena menggunakan data dalam jumlah besar, diberikan usulan untuk menampilkan visualisasi dari hasil proses *clustering* berbasis LDA dalam bentuk *radial tree*. Visualisasi dengan bentuk *tree* akan menampilkan hasil *clustering* dengan lebih jelas (Beham, et al., 2014). Usulan tersebut mereferensi *Architecture Tree* dari Marmelab yang menampilkan sejumlah topik yang disusun dalam bentuk *radial* atau melingkar.

Visualisasi dengan *Architecture Tree* masih memiliki beberapa permasalahan, yaitu data yang digunakan masih sangatlah besar untuk ditampilkan sekaligus dalam satu halaman, karena pada *Architecture Tree* node tidak akan tertutup secara otomatis apabila membuka node yang lain, sehingga pengguna harus memilih kembali node yang sudah terbuka untuk menutupnya. Oleh karena itu diberikan modifikasi pada bagian *node tree* yang mereferensi dari visualisasi yang dilakukan oleh Dachsel dan Ebert yang berupa *collapsible node*. Sehingga apabila terdapat *node* yang terbuka, dan akan membuka *node* lain pada level yang sama, maka *node* yang sudah terbuka akan tertutup secara otomatis. Hal tersebut dilakukan untuk menghindari menampilkan informasi yang berlebihan (Dachsel & Ebert, 2001). Sesuai fungsi dan bentuk dari modifikasi *Architecture Tree* ini, selanjutnya akan disebut *Collapsible Radial Tree*. Modifikasi ini diharapkan lebih menghemat ruang untuk menampilkan hasil *clustering* berbasis LDA.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah dijelaskan, maka dapat dirumuskan permasalahan yaitu bagaimana melakukan *clustering* terjemahan ayat Al-Qur'an menggunakan metode LDA dan bagaimana memvisualisasikan hasil *clustering* berbasis LDA dengan menggunakan usulan visualisasi *Collapsible Radial Tree*.

1.3 Tujuan dan Manfaat

Tujuan dari penelitian tugas akhir ini antara lain :

1. Menentukan model *clustering* terbaik dengan menggunakan metode LDA untuk dijadikan sebagai data pada visualisasi *Collapsible Radial Tree*.
2. Mengukur seberapa mudah penggunaan visualisasi *Collapsible Radial Tree* bagi pengguna dengan melakukan pengujian *usability*.

Adapun manfaat yang diharapkan dari penelitian tugas akhir ini adalah sistem yang dibuat dan dikembangkan dapat memberikan kontribusi terhadap penelitian mengenai *clustering* terjemahan ayat Al-Qur'an agar dapat memudahkan umat Islam untuk mempelajari Al-Qur'an.

1.4 Ruang Lingkup

Ruang lingkup yang akan dibahas dalam penelitian ini adalah:

1. *Input* dari sistem ini berupa data berupa *text* dari terjemahan ayat Al-Qur'an dalam bahasa Indonesia dari halaman website <http://www.qurandatabase.org/>.
2. Kinerja dari *clustering* menggunakan LDA hanya dilihat dari nilai *perplexity*.
3. *Output* dari sistem ini berupa hasil *clustering* terjemahan ayat Al-Qur'an berdasarkan topik yang divisualisasikan dalam bentuk *Collapsible Radial Tree*.
4. Pengujian *usability* dilakukan untuk menilai visualisasi *Collapsible Radial Tree* berdasarkan indikator *usability* menurut Nielsen.
5. Pengembangan sistem ini menggunakan bahasa pemrograman Python, HTML, CSS, PHP, Javascript dan Sistem manajemen basis data MySQL.

1.5 Sistematika Penulisan

Sistematika penulisan yang digunakan dalam tugas akhir ini terbagi dalam beberapa pokok bahasan, yaitu:

BAB I PENDAHULUAN

Bab ini memberikan gambaran mengenai latar belakang masalah, rumusan masalah, tujuan dan manfaat, ruang lingkup serta sistematika penulisan tugas akhir mengenai *Clustering* Terjemahan Ayat Al-Qur'an menggunakan LDA dengan Visualisasi *Collapsible Radial Tree*.

BAB II TINJAUAN PUSTAKA

Bab ini memberikan kajian pustaka yang berhubungan dengan tema tugas akhir sebagai landasan untuk perumusan dan analisis permasalahan pada tugas akhir. Bab ini menyajikan *clustering* terjemahan ayat Al-Qur'an, perkembangan penelitian mengenai *clustering* terjemahan ayat Al-Qur'an, tahap *preprocessing*, penjelasan mengenai LDA dengan visualisasi . *Collapsible Radial Tree*.

BAB III METODOLOGI PENELITIAN

Bab ini menyajikan garis besar penyelesaian masalah tugas akhir yang diawali dengan penyajian garis besar penyelesaian masalah dalam bentuk blok proses. Garis besar penyelesaian masalah diawali dengan melakukan *preprocessing* untuk pembentukan *bag of words*, proses LDA, proses evaluasi dengan visualisasi *Collapsible Radial Tree*.

BAB IV HASIL EKSPERIMEN DAN ANALISA

Bab ini menguraikan hasil skenario eksperimen dan analisa pada penelitian yang dimulai dari teknis pengumpulan data, penjelasan mengenai semua skenario eksperimen dan analisa dari setiap hasil eksperimen yang telah dilakukan.

BAB V PENUTUP

Bab ini menjabarkan kesimpulan dari uraian yang telah diulas pada bab-bab sebelumnya dan saran untuk pengembangan penelitian lebih lanjut.