

**Pengelompokan Tematik Al-Qur'an Menggunakan Metode LSA  
Dengan Pembobotan *Raw Count* dan *IDF-Smooth***

**SKRIPSI**

Diajukan Untuk Memenuhi Salah Satu Syarat  
Memperoleh Gelar Sarjana Teknik Pada Fakultas Teknik  
Universitas Islam Riau



**OLEH:**  
**HAPPY FIRSTY AMELIA**  
**143510806**

**PROGRAM STUDI TEKNIK INFORMATIKA  
FAKULTAS TEKNIK  
UNIVERSITAS ISLAM RIAU  
PEKANBARU  
2020**

## LAMPIRAN PERNYATAAN BEBAS PLAGIARISME

Saya yang bertanda tangan di bawah ini :

Nama : Happy Firsty Amelia  
Tempat/Tgl Lahir : Padang, 05 Januari 1995  
Alamat : Bedeng Sikuran, RT 002/RW 002, Kel. Bedeng  
Sikuran, Kec. Inuman, Kab. Kuantan Singingi

Adalah Mahasiswi Universitas Islam Riau yang terdaftar pada :

Falkultas : Teknik  
Jurusan : Teknik Informatika  
Program Studi : Teknik Informatika  
Jenjang Pendidikan : Strata Satu (S1)

Dengan ini menyatakan sesungguhnya bahwa skripsi yang telah saya tulis adalah benar dan asli hasil dari penelitian yang telah saya lakukan dengan judul **“Pengelompokan Tematik Al-Qur’an Menggunakan Metode LSA dengan Pembobotan Raw Count dan IDF-Smooth”**.

Apabila dikemudian hari ada yang merasa dirugikan atau menuntut karena penelitian ini menggunakan sebagian hasil tulisan atau karya orang lain tanpa mencantumkan nama penulis yang bersangkutan, atau terbukti karya ilmiah ini bukan karya saya sendiri atau **plagiat** hasil karya orang lain, maka saya bersedia menerima sanksi sesuai dengan peraturan perundang-undangan yang berlaku.

Demikian surat pernyataan ini saya buat dengan sesungguhnya untuk dapat digunakan sebagaimana mestinya.

Pekanbaru, 08 Mei 2020

Yang membuat pernyataan,



HAPPY FIRSTY AMELIA

## HALAMAN PERSEMBAHAN

*Assalamualaikum Wr.Wb.*

Alhamdulillahirobbil'alamin, Puji syukur penulis ucapkan kehadiran Allah yang Maha Esa, karena atas rahmat dan hidayah-Nya. Maka saya dapat menyelesaikan Tugas Akhir yang berjudul **"Pengelompokan Tematik Al-Qur'an Menggunakan Metode LSA Dengan Pembobotan Raw Count dan IDF-Smooth"**

Adapun Tugas Akhir ini dibuat untuk diajukan sebagai salah satu syarat guna memperoleh gelar sarjana Strata Satu (S-1) pada Program Studi Teknik Informatika Fakultas Teknik Universitas Islam Riau. Dalam kesempatan ini penulis juga ingin mengucapkan terimakasih yang sebesar-besarnya kepada :

- ❖ Akmar Efendi S.Kom., M.Kom selaku Dosen Penasehat Akademik yang senantiasa selalu memberikan nasehat dan motivasi kepada penulis.
- ❖ Dr. Arbi Haza Nasution, M.IT., B.IT selaku Dosen Pembimbing Skripsi yang selalu senantiasa memberikan Wejangan dan Bimbingan sehingga penulis dapat menyelesaikan Skripsi dengan baik.
- ❖ Terutuk Lelaki Terhebat, yaitu Papa Amelia yang sangat penulis sayangi. Terimakasih untuk segala dukungan dan semangat sehingga penulis dapat menyelesaikan study dengan baik.



- ❖ Teruntuk Wanita Tangguh Keluarga, yaitu Mama Gusniarti Indra. Terimakasih untuk segala dukungan, repetan, wejangan, dan segudang list planning kehidupan. Terimakasih untuk tetap sehat dan menyaksikan sebuah karya sederhana yang penulis sajikan untukmu.
- ❖ Teruntuk beberapa Kurcaci Keluarga 'GINDRA', Terimakasih untuk Intan Nabila Putri Amelia selaku adek cewek satu-satunya yang selalu memberikan kalimat "SEMANGAT SKRIPSIANNYA KAK, JANGAN MELALA MULUK", Untuk Dhailallul Khairat yang selalu melontarkan pertanyaan "KAKAK KAPAN WISUDA" Untuk M. Ikhsan Nur Hakim malaikat kecilnya keluarga. Terimakasih untuk semua hal ribet dan kebahagiaan yang kalian buat.
- ❖ Teruntuk Keluarga Besar Alm. Ruslan Abbas, Terimakasih untuk semua dukungan yang diberikan hingga penulis bisa menyelesaikan perkuliahan dengan baik.
- ❖ Teruntuk Teman-Teman Satu Bimbingan, yaitu Sri Devi, Maulana Sarowis, Yoga Kurniawan, Jhody Damara dan Jakfar Shodiq terimakasih atas do'a, semangat, hal-hal merepotkan serta kerja keras team yang luar biasa.
- ❖ Teruntuk wanita-wanita terbaik sepanjang masa, yaitu Sri Devi, Sri Handayani, Nur'ainaani Darma, Rezeki Olanda Hermila, Farahdiva Assyfa Andrin dan Firma Hilda Wahyuni Terimakasih banyak untuk semua drama yang ada, untuk semua canda tawa, untuk semua momen yang tercipta dan untuk keputusan kalian yang selalu tetap ada.
- ❖ Teruntuk Lelaki-Lelaki KUSUT, yaitu Wawan, Tamma, Ari, Wahyu dan Rasyid, Terimakasih sudah mengizinkan seorang Happy menjadi salah satu wanita waras diantara 5 laki-laki yang sedikit kurang waras , terimakasih untuk setiap momen liburan yang selalu dadakan.

- ❖ Teruntuk Manusia-Manusia IDIOT, yaitu Fikri, Edwin, Dika, Ervin, Om Tray, Fajar, Ayu, Cippy, Tiara, Mayang, dan Ulyy Terimakasih sudah memberikan history lucu semasa drama kuliah yang panjang.
- ❖ Teruntuk wanita Satu Zodiak Capricorn, Yaitu Farhana Aulia R. Terimakasih sudah menambahkan banyak drama selama perkuliahan.
- ❖ Teruntuk Teman-Teman Kelas A Angkatan 2014 yang tidak bisa disebutkan satu-persatu. Terimakasih sudah memberikan momen terbaik selama drama perkuliahan.
- ❖ Teruntuk Teman-Teman Kelas E Angkatan 2015 yang tidak bisa disebutkan satu-persatu. Terimakasih sudah menerima seorang happy untuk menjadi anggota lokal, walau Cuma dua Semester.
- ❖ Teruntuk beberapa Tokoh, Terimakasih untuk yang pernah menetap ataupun memilih untuk pergi.
- ❖ Serta Terimakasih untuk segala pihak yang ikut membantu, baik secara langsung maupun tidak langsung.

Sebagai manusia biasa penulis menyadari dalam penulisan Tugas Akhir ini masih jauh dari kesempurnaan, oleh karena itu saran dan kritik untuk kemajuan penulis sangat diharapkan demi pengembangannya di masa yang akan datang.

Amin....

***Wassalamualaikum Wr.Wb.***

Pekanbaru, 08 Mei 2020  
Mahasiswa Ybs,

**HAPPY FIRSTY AMELIA**

## KATA PENGANTAR

*Assalamu'alaikum Wr. Wb.*

Dengan menyebut nama Allah SWT yang Maha Pengasih lagi Maha Panyayang, Penulis ucapkan puji syukur atas kehadiran-Nya, yang telah melimpahkan rahmat, hidayah, dan inayah-Nya, sehingga Penulis dapat menyelesaikan skripsi yang berjudul **“Pengelompokan Tematik Al-Qur’an Menggunakan Metode LSA dengan Pembobotan Raw Count dan IDF-Smooth”** ini tepat pada waktunya, untuk memenuhi salah satu syarat dan memperoleh Gelar Sarjana Teknik Pada Fakultas Teknik Universitas Islam Riau.

Dalam penyusunan skripsi ini, Penulis telah banyak mendapat hambatan dan halangan yang berarti. Untuk itu, Penulis menyampaikan ucapan terima kasih kepada semua pihak yang telah banyak membantu dalam pembuatan skripsi ini.

1. Bapak Dr. Eng. Muslim, ST.,MT selaku Dekan Fakultas Teknik.
2. Ibu Ause Labellapansa, ST., M.Kom selaku ketua Program Studi Teknik Informatika.
3. Bapak Dr. Arbi Haza Nasution, M.IT., B.IT selaku pembimbing yang telah ikhlas dan sabar memberikan bimbingan, motivasi dan arahan di tengah kesibukan beliau.
4. Seluruh Dosen Prodi Teknik Informatika serta staff tata usaha.

5. Kedua orang tua dan keluarga, yang selalu mendo'akan, serta memberikan dukungannya.

Terlepas dari semua itu, Penulis menyadari bahwa skripsi ini masih jauh dari kesempurnaan baik dalam bentuk penyusunan maupun materinya. Kritik konstruktif dari pembaca sangat penulis harapkan untuk dapat menyempurnakan skripsi ini. Akhir kata semoga skripsi ini dapat memberikan manfaat yang sebesar-besarnya.

*Wassalamu'alaikum Wr. Wb.*

Pekanbaru, 08 Mei 2020

**HAPPY FIRSTY AMELIA**



# AL-QUR'AN THEMATIC CLUSTERING USING METHOD *LSA* WITH *RAW COUNT* AND *IDF-SMOOTH*

Happy Firsty Amelia  
Faculty of Engineering  
Informatic Engineering Study Program  
Islamic University of Riau  
Email : [happyfirstyamelia@student.uir.ac.id](mailto:happyfirstyamelia@student.uir.ac.id)

## ABSTRACT

The development of the current era of globalization is always experiencing very significant progress. Plus the information dissemination is very fast to various directions. Technology itself requires us to solve a problem with speed, accuracy and accuracy. One of them is understanding the content contained in the Al-Qur'an by analyzing and visualizing thematic Al-Qur'an. This study presents a thematic grouping of the Qur'an using the latent semantic analysis (LSA) method and weighting of Raw count and IDF-Smooth. As an example of how to analyze a verse with the theme of faith in the books by gathering the whole verses of the Qur'an that contain the meaning of the theme, then an analysis is done one by one from the verse, whether the verse falls into the category of believers in the books the book according to view in the Qur'an Amazing Cordoba. Data that has been analyzed by the system produces 22 suras and false positive verses which will then be re-evaluated by a Qur'anic thematic expert. The data that has been evaluated results in 5 contributions of surahs and verses with an accuracy rate of 44-47% in the sub-theme "Obligations and Virtues of Belief in the Books".

**Keywords:** *Al-Qur'an*, *latent semantic analysis (LSA)*, *Raw Count* and *IDF-Smooth*



**PENGELOMPOKAN TEMATIK AL-QUR'AN MENGGUNAKAN  
METODE LSA DENGAN PEMBOBOTAN  
RAW COUNT DAN IDF-SMOOTH**

Happy Firsty Amelia  
Fakultas Teknik  
Program Studi Teknik Informatika  
Universitas Islam Riau  
Email : [happyfirstyamelia@student.uir.ac.id](mailto:happyfirstyamelia@student.uir.ac.id)

**ABSTRAK**

Perkembangan era globalisasi saat ini selalu mengalami kemajuan yang sangat signifikan. Ditambah lagi penyebaran informasi yang sangat cepat ke berbagai penjuru. Teknologi itu sendiri menuntut kita untuk menyelesaikan sebuah masalah dengan kecepatan, ketepatan dan keakuratan. Salah satunya memahami isi kandungan yang terdapat di dalam Al-Qur'an dengan cara melakukan analisa dan visualisasi tematik Al-Qur'an. Penelitian ini menyajikan sebuah pengelompokan tematik Al-Qur'an dengan menggunakan metode latent semantic analysis (LSA) dan pembobotan Raw count dan IDF-Smooth. Sebagai contoh bagaimana cara menganalisa sebuah ayat yang bertemakan beriman kepada kitab-kitab dengan cara mengumpulkan keseluruhan ayat Al-Qur'an yang mengandung arti tema tersebut, kemudian dilakukan sebuah analisa satu persatu dari ayat tersebut, apakah ayat tersebut masuk kedalam kategori beriman kepada kitab-kitab menurut pandangan dalam Al-Qur'an Amazing Cordoba. Data yang sudah dianalisis oleh sistem menghasilkan 22 data surah dan ayat tidak tepat (false positive) yang kemudian akan di evaluasi kembali oleh seorang pakar tematik Al-Qur'an. Data yang sudah dievaluasi tersebut menghasilkan 5 kontribusi surah dan ayat dengan tingkat akurasi sebesar 44-47% pada subtema "Kewajiban dan Keutamaan Beriman pada Kitab-Kitab".

**Kata Kunci :** *Al-Qur'an, latent semantic analysis (LSA), Raw Count dan IDF-Smooth*

## DAFTAR ISI

	Hal
<b>KATA PENGANTAR.....</b>	<b>ix</b>
<b>DAFTAR ISI.....</b>	<b>xiii</b>
<b>DAFTAR TABEL .....</b>	<b>xvi</b>
<b>DAFTAR GAMBAR.....</b>	<b>xviii</b>
<b>DAFTAR LAMPIRAN .....</b>	<b>xix</b>
<b>BAB I PENDAHULUAN.....</b>	<b>1</b>
1.1 Latar Belakang.....	1
1.2 Identifikasi Masalah .....	3
1.3 Rumusan Masalah .....	4
1.4 Batasan Masalah.....	4
1.5 Tujuan Penelitian.....	5
1.6 Manfaat Penelitian.....	5
<b>BAB II LANDASAN TEORI .....</b>	<b>6</b>
2.1 Studi Kepustakaan .....	6
2.2 Dasar Teori .....	7
2.2.1 Al-Qur'an .....	7
2.2.2 Tematik .....	7
2.2.3 Beriman Kepada Kitab-Kitab.....	8
2.2.4 Nama Kitab-Kitab Allah SWT.....	10

2.2.5	Subtema Kitab-Kitab.....	12
2.2.6	<i>Text Mining</i> .....	13
2.2.7	<i>Preprocessing</i> .....	13
2.2.8	<i>Bag Of Word</i> .....	14
2.2.9	<i>Latent Semantic Analysis</i> .....	15
2.2.10	TF-IDF .....	15
2.2.11	<i>Singular Value Decomposition</i> .....	18
2.2.12	<i>Evaluation Measure</i> .....	19
2.2.13	<i>Python</i> .....	20
2.2.14	<i>Flowchart</i> .....	20
<b>BAB III METODOLOGI PENELITIAN.....</b>		<b>23</b>
3.1	Alat dan Bahan Penelitian.....	23
3.1.1	Teknik Pengumpulan Data.....	23
3.1.2	Spesifikasi Kebutuhan Perangkat Keras ( <i>Hardware</i> ) .....	24
3.1.3	Spesifikasi Kebutuhan Perangkat Lunak ( <i>Software</i> ).....	24
3.2	Gambar Alur Penelitian.....	25
3.3	Perancangan Sistem .....	26
3.3.1	Tahapan Preprocessing.....	26
3.3.2	<i>Bag Of Word</i> .....	30
3.3.3	<i>Latent Semantic Analysis</i> .....	31
3.3.4	TF-IDF .....	33

3.3.5	<i>Singular Value Decomposition</i> .....	38
3.3.6	<i>Evaluasi</i> .....	40
<b>BAB IV HASIL DAN PEMBAHASAN .....</b>		<b>41</b>
4.1	Data Inputan .....	41
4.1.1	Data Tematik Al-Qur'an <i>Cordoba</i> .....	41
4.2	Tahapan <i>Preprocessing</i> .....	43
4.3	Pembobotan Term .....	45
4.3.1	Term Frequency ( <i>TF</i> ) .....	45
4.3.2	Invers Document Frequency- <i>Smooth</i> ( <i>IDF-Smooth</i> ).....	46
4.3.3	<i>TF-IDF Smooth</i> .....	47
4.4	<i>Singular Value Decomposition</i> ( <i>SVD</i> ).....	49
4.5	<i>Evaluasi</i> .....	53
4.5.1	Data Inputan <i>Gold Standard</i> .....	53
4.5.2	Hasil Evaluasi Terhadap Al-Qur'an <i>Amazing Cordoba</i> .....	55
4.5.3	Hasil Evauasi <i>Data False Positive</i> Al-Qur'an.....	58
4.5.4	Hasil Evaluasi Terhadap Pakar.....	58
<b>BAB V KESIMPULAN DAN SARAN .....</b>		<b>61</b>
5.1	Kesimpulan.....	61
5.2	Saran.....	62
<b>DAFTAR PUSTAKA .....</b>		<b>63</b>



## DAFTAR TABEL

	Hal
Tabel 2.1 Kategori Jumlah Ayat, Surah dan Subtema Kitab-Kitab .....	12
Tabel 2.2 Contoh Pembobotan Term Frequency .....	16
Tabel 2.3 Pembobotan Invers Document Frequency .....	17
Tabel 2.4 Simbol dan Fungsi <i>Flowchart</i> .....	21
Tabel 3.1. Contoh Proses <i>Tokenizing</i> .....	28
Tabel 3.2. Contoh Proses <i>Stemming</i> .....	28
Tabel 3.3. Contoh Proses <i>Case Folding</i> .....	29
Tabel 3.4. Contoh Proses <i>Punctuation</i> .....	29
Tabel 3.5. Contoh Proses <i>Stopword</i> .....	30
Tabel 3.6. Contoh <i>Bag Of Word</i> Tema Beriman Kepada Kitab-Kitab .....	31
Tabel 3.7. Dokumen yang Akan Dihitung .....	33
Tabel 3.8. Kata Kunci <i>Bag Of Words</i> .....	35
Tabel 3.9. Kemunculan <i>Bag Of Word</i> TF .....	35
Tabel 3.10. Hasil Perhitungan IDF .....	36
Tabel 3.11. Hasil Perhitungan TF-IDF <i>Smooth</i> .....	37
Tabel 3.12. Contoh Data Awal SVD ( <i>U</i> ) .....	38
Tabel 3.13. Contoh Pengelompokan SVD ( <i>U</i> ) .....	39
Tabel 3.14. Contoh Pembentukan Konsep .....	40
Tabel 4.1. Contoh Data Tematik Cordoba Kitab-Kitab .....	41

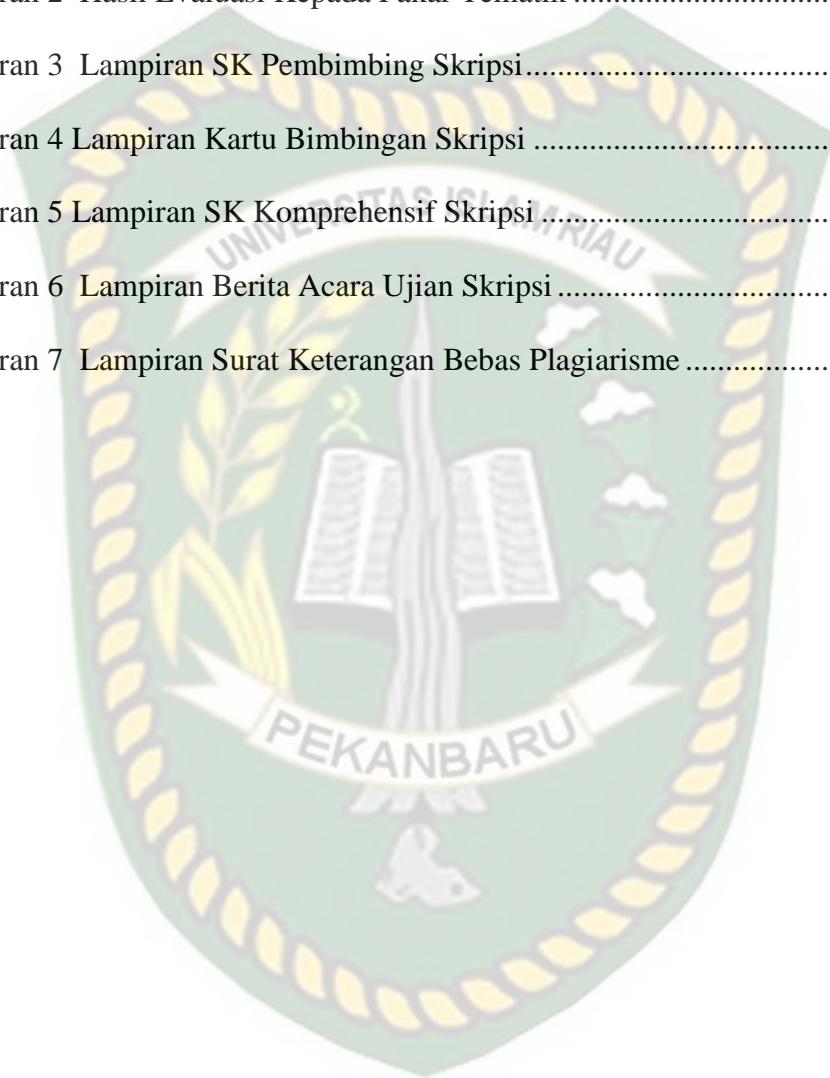
Tabel 4.2. Tampilan Langkah-Langkah dan Hasil <i>Preprocessing</i> .....	43
Tabel 4.3. Penambahan <i>Stopword</i> .....	44
Tabel 4.4. <i>Bag Of Word</i> Tema Beriman Kepada Kitab-Kitab .....	45
Tabel 4.5. Mengeset Nilai BoW Menjadi 0 .....	46
Tabel 4.6. Hasil Menghitung Jumlah Kemunculan BoW .....	46
Tabel 4.7. Hasil Tampilan Perhitungan IDF- <i>Smooth</i> .....	47
Tabel 4.8. Hasil Tampilan Perhitungan TF-IDF .....	47
Tabel 4.9. Hasil Pengelompokan <i>Surah &amp; Ayat</i> Perkonsep.....	51
Tabel 4.10. Hasil Penggabungan <i>Surah, Ayat, Bow</i> dan TF-IDF .....	53
Tabel 4.11. Inputan Data <i>Gold Standartd</i> .....	54
Tabel 4.12. Hasil Evaluasi Terhadap Al-Qur'an.....	55
Tabel 4.13. Hasil Evaluasi Terbaik Terhadap Al-Qur'an .....	57
Tabel 4.14. Data yang tidak Tepat ( <i>False Positive</i> ).....	58
Tabel 4.15. Hasil Data Evaluasi Terhadap Pakar.....	59
Tabel 4.16. Hasil Evaluasi Terhadap Al-Qur'an dan Pakar.....	60

## DAFTAR GAMBAR

Gambar 3.1	Alur Perancangan Sistem .....	25
Gambar 3.2	<i>Flowchart</i> Preprocessing.....	27
Gambar 3.3	<i>Flowchart Latent Semantic Analysis</i> .....	32
Gambar 4.1	Visualisasi dari TF-IDF- <i>Smooth</i> .....	48
Gambar 4.2	Grafik Presentasi <i>Surah</i> dan <i>Ayat</i> berdasarkan <i>Bag Of Word</i> .....	49
Gambar 4.3	Matrik Perkalian TF-IDF .....	50
Gambar 4.4	Proses Pengelompokan <i>U</i> .....	51
Gambar 4.5	Grafik Hasil Evaluasi Al-Qur'an Terbaik .....	57

## DAFTAR LAMPIRAN

Lampiran 1	Data Tema beriman Kepada Kitab-Kitab .....
Lampiran 2	Hasil Evaluasi Kepada Pakar Tematik .....
Lampiran 3	Lampiran SK Pembimbing Skripsi.....
Lampiran 4	Lampiran Kartu Bimbingan Skripsi .....
Lampiran 5	Lampiran SK Komprehensif Skripsi .....
Lampiran 6	Lampiran Berita Acara Ujian Skripsi .....
Lampiran 7	Lampiran Surat Keterangan Bebas Plagiarisme .....





# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Perkembangan era globalisasi saat ini selalu mengalami kemajuan yang sangat signifikan. Ditambah lagi penyebaran informasi yang sangat cepat ke berbagai penjuru. Sejalan dengan berkembangnya globalisasi yang sangat pesat tentunya bertambah juga berbagai permasalahan yang kita hadapi di kehidupan sehari-hari. Seiring majunya perkembangan teknologi kita dituntut untuk menyelesaikan permasalahan yang ada dengan memanfaatkan kecepatan, ketepatan, keakuratan dan kecanggihan dari teknologi tersebut untuk mendapatkan serta memberi informasi yang ada dengan hasil yang optimal.

Salah satunya dengan visualisasi tematik Al-Qur'an adalah sebuah metode yang dilakukan dengan cara memahami isi kandungan beserta tema-tema yang terdapat di dalam Al-Qur'an dengan cara melakukan sebuah analisa ayat Al-Qur'an dengan tema yang sama. Sebagai contoh apabila kita ingin menganalisa sebuah ayat yang bertemakan beriman kepada kitab-kitab. Maka dapat dilakukan dengan cara mengumpulkan keseluruhan ayat Al-Qur'an yang mengandung arti tentang beriman kepada kitab-kitab, kemudian dilakukan sebuah analisa satu persatu dari ayat tersebut, apakah ayat tersebut masuk kedalam kategori beriman kepada kitab-kitab menurut pandangan dalam Al-Qur'an.

Al-Qur'an adalah firman Allah berupa wahyu yang diturunkan Allah SWT kepada nabi Muhammad SAW melalui perantara malaikat Jibril yang dalam bahasa arab, kemudian disampaikan ke umat manusia secara Al-Tawatur (langsung dari Nabi Muhammad SAW ke banyak orang) yang tercantum mulai dari surat Al-Fatihah di akhiri dengan surat An-Nas. Al-Qur'an terdiri dari 30 juz, 114 surat, 86 makiyah dan 38 surat Madaniyah. Al-Qur'an dijadikan pedoman manusia untuk hidup yang jauh lebih baik lagi hingga akhir zaman. Al-Qur'an dijadikan pegangan hidup selama menjalankan kehidupan di dunia. Dalam bahasa Arab, yang sampai kepada umat manusia secara *al-tawatur* (langsung dari nabi Muhammad SAW kepada orang banyak) yang kemudian termaktub dalam bentuk mushaf, dimulai dari surat Al-Fatihah dan ditutup dengan surat Al-Nas.

Tafsir Qur'an memiliki banyak tema, salah satunya adalah berima kepada kitab – kitab. Ayat yang terdapat di dalam Al-Qur'an mengharuskan seorang muslim untuk membaca, memahami, mempelajari kandungan yang terdapat didalamnya untuk kehidupan dunia dan akhirat. Setiap muslim yang beriman tentunya diwajibkan untuk memahami akan arti dari kitab yang diturunkan Allah SWT. Mempercayai dan meyakini dalam hati bahwa Allah menurunkan wahyu kepada rasulnya berupa kitab-kitab sebagai pedoman untuk hidup kaumnya. kitab – kitab yang Allah SWT turunkan dan Rasul Penerimaannya :

- a. Kitab *Taurat* sebuah kitab yang diturunkan kepada **Nabi Musa AS**
- b. Kitab *Zabur* sebuah kitab yang diturunkan kepada **Nabi Daus AS**
- c. Kitab *Injil* sebuah kitab yang diturunkan kepada **Nabi Isa AS**

d. Kitab *Al-Qur'an* sebuah kitab yang diturunkan kepada Nabi terakhir **Muhammad SAW** sebagai petunjuk hidup umatnya.

Berdasarkan uraian diatas yang menjadi permasalahan adalah bagaimana kita membahas dan memecahkan suatu masalah lebih rinci tentang tematik Al-Qur'an serta bagaimana cara melakukan pembagian ayat-ayat tersebut dalam satu tema yang sama. Dengan demikian dapat disimpulkan mesin pencarian menjadi salah satu alternatif yang baik untuk seseorang belajar dan memahani isi kandungan dalam Al-Qur'an. Teknologi tematik sendiri dipilih sebagai suatu solusi untuk mempermudah dalam proses pencarian informasi. Atas sadar latar belakang tersebut maka proposal skripsi ini mengambil judul: **“Pengelompokan Tematik Al-Qur'an Menggunakan Metode LSA dengan Pembobotan *Raw Count* dan *IDF-Smooth*”**.

## 1.2 Identifikasi Masalah

Adapun identifikasi masalah yang dapat diambil dari latar belakang tersebut adalah sebagai berikut :

1. Pengelompokan Tematik Al-Qur'an masih dilakukan secara manual yaitu dengan cara mengumpulkan para ulama dan ahli tafsir.
2. Pengelompokan Tematik Al-Qur'an membutuhkan waktu lama yang dirasa kurang maksimal.
3. Belum tersedianya *korpus* atau kamus data mengenai pengelompokan tematik Al-Qur'an dengan tema beriman kepada kitab-kitab.

### 1.3 Rumusan Masalah

Berdasarkan uraian latar belakang diatas, maka dapat diambil rumusan masalah yaitu, bagaimana cara melakukan pengelompokan tematik Al-Qur'an secara otomatis dengan tema beriman kepada kitab-kitab.

### 1.4 Batasan Masalah

Agar permasalahan tidak menyimpang dan lebih terarah maka peneliti memberikan batasan pengembangannya. Batasan – batasan yang terdapat pada penelitian ini adalah :

1. Pengelompokan tematik Al-Qur'an ini hanya membahas tema beriman kepada kitab-kitab dengan menggunakan beberapa *bag of word* yang telah ditentukan.
2. Pengelompokan tematik Al-Qur'an ini menggunakan data terjemahan tematik tema beriman kepada kitab-kitab bersumber dari Al-Qur'an *Cordoba Amazing* yang di inputkan secara manual dalam berbentuk format csv.
3. Pengelompokan tematik Al-Qur'an ini menggunakan keseluruhan data Qur'an dari <http://qurandatabase.org/>
4. Pengelompokan tematik Al-Qur'an ini menghasilkan data inputan hanya sebatas tema beriman kepada kitab-kitab.
5. Pengelompokan tematik Al-Qur'an ini hanya menggunakan 6 subtema dari tema beriman kepada kitab-kitab berdasarkan Al-Qur'an *Cordoba Amazing*.



### 1.5 Tujuan Penelitian

Adapun tujuan dari penelitian ini adalah :

1. Mempermudah pengelompokan tematik secara menyeluruh dengan tema beriman kepada kitab-kitab.
2. Membuat sebuah pengelompokan dari hasil *bag of word* untuk menghasilkan surah dan ayat Al-Qur'an berdasarkan subtema pada tema beriman kepada kitab-kitab.
3. Menggunakan metode LSA untuk membangun sebuah tematik Al-Qur'an dalam pencarian ayat-ayat yang sesuai dengan tema beriman kepada kitab-kitab sehingga dapat menunjang pemahaman dan pembelajaran terhadap Al-Qur'an secara keseluruhan.

### 1.6 Manfaat Penelitian

Adapun manfaat dari penelitian ini adalah sebagai berikut :

1. Mempermudah ahli tafsir untuk melakukan pengelompokan surah dan ayat tema beriman kepada kitab-kitab.
2. Membantu peneliti selanjutnya untuk memperoleh kamus data tentang pengelompokan sebuah tematik Al-Qur'an tema beriman kepada kitab-kitab.

## BAB II

### LANDASAN TEORI

#### 2.1 Studi Kepustakaan

Dalam sebuah penelian yang dilakukan oleh (Wicaksono & Irawan, 2014) dengan penelitian yang berjudul “*Sistem Deteksi Kemiripan Antar Dokumen Teks Menggunakan Model Bayesian Pada Term Latent Semantic Analysis (LSA)*”. Pada penelitian ini bertujuan untuk mempermudah pengguna menemukan beberapa tindakan yang berpotensi melakukan penjiplakan atau *plagiat* dengan cara melakukan pemeriksaan kemiripan antar dokumen tersebut dengan menggunakan metode LSA (*Latent Semantic Analysis*) dengan memanfaatkan frekuensi kemunculan dari beberapa term yang dihasilkannya.

Penelitian kedua dilakukan oleh (Ismail, 2018) dengan penelitian yang membahas tentang “*Analisis Semantik Pada Kata Ahzab dan Berivasinya dalam Al-Qur’an*”. Penelitian ini bertujuan untuk untuk memudahkan pengguna dalam memahami pemaknaan yang tepat terhadap kata *Ahzab* dalam Al-Qur’an. Metode yang digunakan peneliti dalam pembahasan kali ini adalah menggunakan metode *analisis komponen semantik dan analisis kombinasi semantik ahzab*. Yang bertujuan untuk mengetahui konsep ahzab dalam al-qur’an berdasarkan tinjauan semantik.

Penelitian selanjutnya dilakukan oleh (Luthfiarta, Zeniarja, & Salam, 2013), dengan penelitian berjudul “*Algoritma Latent Semantic Analysis (LSA) pada Peringkat Dokumen Otomatis Untuk Proses Clustering Dokumen*”, penelitian ini

memiliki peranan yang sangat signifikan dalam kemajuan teknologi informasi, bertujuan untuk memudahkan dan menampilkan informasi, diantaranya memiliki peranan penting dalam pengembangan web dibidang akurasi kategorisasi *keyword* otomatis pada sebuah *search engine*.

## 2.2 Dasar Teori

### 2.2.1 Al-Qur'an

Al-Qur'an adalah "*bacaan*" sedangkan asal kata Al-Quran berbentuk masdar dengan arti isim maful, yaitu *marfu'* (dibaca). Adapun pengertian Al-Qur'an secara terminologis adalah Kalam Allah SWT yang merupakan sebuah mukjizat yang diturunkan (*diwahyukan*) kepada nabi Muhammad SAW yang dituliskan di *mushaf* dan diriwayatkan secara *mutawatir* dan membacanya adalah sebuah ibadah (Setiaji, 2014). Ada juga yang mengatakan Al-Qur'an adalah *kalamullah* yang diturunkan kepada Nabi Muhammad, dengan bahasa Arab, yang sampai kepada kita secara mutawatir, yang ditulis di dalam *mushaf*, dimulai dari surah al-fatihah dan diakhiri dengan surah an-nas, membacanya berfungsi sebagai ibadah, sebagai mukjizat bagi nabi Muhammad dan sebagai hidayah atau petunjuk bagi umat manusia untuk dijadikan petunjuk dan pedoman hidup bagi setiap umat Islam yang ada di muka bumi.

### 2.2.2 Tematik

Metode Tematik Al-Qur'an adalah sebuah metode memahami makna dai keagamaan ataupun sebuah term-term dalam Al-Qur'an dengan cara menganalisis

keseluruhan ayat-ayat Al-Qur'an yang memiliki kesamaan. (Rahmat,2015) Menurut Zahir bin Awadh, Tafsir Maudhu'i yaitu : suatu metode pengumpulan ayat-ayat Al-Quran yang terpisah-pisah dari berbagai surat dalam Al-Quran yang berhubungan dengan topik (tema) yang sama baik secara lafas maupun hukum, dan menafsirkannya sesuai dengan tujuan-tujuan Al-Quran. Sementara itu ulama Baqir Al-Sadr memberikan pengertian, bahwa Tafsir Maudhu'i yaitu : suatu metode Tafsir yang berupaya menghimpun ayat-ayat Al-Quran dari berbagai surat dan yang berkaitan pula dengan persoalan atau tema yang ditetapkan sebelumnya, kemudian membahas dan menganalisa kandungan ayat-ayat tersebut sehingga menjadi suatu kesatuan yang utuh.

Dari berbagai pengertian yang dikemukakan tersebut di atas, maka dapat diambil suatu kesimpulan bahwa Tafsir Maudhu'i yaitu suatu metode penafsiran Al-Quran dimana para mufassir berupaya mengumpulkan ayat-ayat Al-Qur'an dari berbagai surat yang memiliki kesamaan tema, sehingga mengarah kepada suatu pengertian dan tujuan yang sama pula.

### **2.2.3 Beriman Kepada Kitab-Kitab**

Diantara keenam rukum-rukun iman yang wajib di yakini oleh seorang mukmin ialah beriman kepada kitab-kitab yang telah Allah turunkan kepada rasulnya. Dimana Allah berfirman :



آمَنَ الرَّسُولُ بِمَا أُنزِلَ إِلَيْهِ مِنْ رَبِّهِ وَالْمُؤْمِنُونَ كُلٌّ آمَنَ بِاللَّهِ وَمَلَائِكَتِهِ وَكُتُبِهِ وَرُسُلِهِ  
لَا نُفَرِّقُ بَيْنَ أَحَدٍ مِنْ رُسُلِهِ وَقَالُوا سَمِعْنَا وَأَطَعْنَا غُفْرَانَكَ رَبَّنَا وَإِلَيْكَ الْمَصِيرُ

"Rasul telah beriman kepada al-Qur'an yang diturunkan kepadanya dari Tuhannya, demikian pula orang-orang yang beriman. semuanya beriman kepada Allah, malaikat-malaikat -Nya, kitab-kitab -Nya dan rasul-rasul -Nya". (QS al-Baqarah: 285).

Demikian pula ditegaskan oleh Allah SWT dalam firman-Nya :

كَانَ النَّاسُ أُمَّةً وَاحِدَةً فَبَعَثَ اللَّهُ النَّبِيِّينَ مُبَشِّرِينَ وَمُنذِرِينَ وَأَنْزَلَ مَعَهُمُ الْكِتَابَ بِالْحَقِّ  
لِيَحْكُمَ بَيْنَ النَّاسِ فِي مَا اختلفوا فِيهِ

"Manusia itu adalah umat yang satu. (setelah timbul perselisihan), Maka Allah mengutus para Nabi, sebagai pemberi peringatan, dan Allah menurunkan bersama mereka kitab yang benar, untuk memberi keputusan di antara manusia tentang perkara yang mereka perselisihkan". (QS al-Baqarah: 213).

Imam Ibnu Abil Izz Al-Hanafi dalam bukunya Syarh Thahawiyah menjelaskan, "Adapun beriman kepada kitab-kitab suci 5 yang diturunkan kepada para rasul, maka kami mengimani nama-nama yang telah Allah ta'ala berikan dalam Al-Qur'an seperti Taurat, Injil dan Zabur. Adapun beriman pada Al-Qur'an maka hal itu dengan cara menetapkan Al-Qur'an, serta mengikuti petunjuknya, yang mana dua perkara ini merupakan kelebihan Al-Qur'an dari pada keimanan kepada kitab-kitab suci lainnya. Maka wajib atas kita untuk beriman bahwa kitab-kitab yang diturunkan pada para

rasul (dahulu) semuanya datang dari sisi Allah dengan benar, membawa petunjuk, cahaya, penjelas, serta penawar hati".

Adapun berian kepada kitab-kitab suci yang telah diturunkan oleh rasul, maka kami mengimani nama-nama yang telah Allah berikan dala Al-Qur'an seperti : Taurat, Injil, Zabur. Beriman kepada kitab-kitab Allah rasal dari bahasa arab yaitu *Jannah* dengan akar katanya yaitu *Janna*. Kata tersebut berasal dari kata *Janana* pada asalnya berarti tertutup, yaitu tidak dapat dijangkau oleh panca indera manusia.

#### 2.2.4 Nama Kitab-Kitab Allah SWT

1. *Taurat* yaitu kitab yang diturunkan kepada Nabi Musa AS sebagai pedoman hidup bagi kaum Bani Israil. Firman Allah SWT (QS.Al-Isra [17]:2).

وَأَتَيْنَا مُوسَى الْكِتَابَ وَجَعَلْنَاهُ هُدًى لِّبَنِي إِسْرَائِيلَ إِلَّا تَتَّخِذُوا مِنِّي دُونِي وَكَيْلًا

“Dan Kami berikan kepada Musa, Kitab (Taurat) dan Kami jadikannya petunjuk bagi Bani Israil (dengan firman), “Janganlah kamu mengambil (pelindung) selain Aku.”(QS.Al-Isra [17]:2)

Adapun isi pokok Kitab Taurāt adalah: (1) jangan ada padamu Tuhan lain dihadirat-Ku. (2) Jangan membuat patung ukiran dan jangan pula menyembah patung karena Aku Tuhan Allāh mu. (3) Jangan kamu menyebut Tuhan Allāh mu dengan sia-sia. (4) Ingatlah akan hari sabat (sabtu), supaya kamu sucikan dia. (5) Berilah hormat kepada bapak ibumu. (6) Jangan membunuh sesama manusia. (7) Jangan berzina. (8) Jangan mencuri. (9) Jangan menjadi saksi palsu. (10) Jangan berkeinginan memiliki hak orang lain.

2. *Zabur* yaitu kitab yang diturunkan kepada Daud AS sebagai pedoman hidup bagi kaumnya. Firman Allah SWT (QS.AI-Isra [17]:55).

وَرَبُّكَ أَعْلَمُ بِمَنْ فِي السَّمَاوَاتِ وَالْأَرْضِ ۗ وَقَدْ فَضَّلْنَا بَعْضَ النَّبِيِّنَ عَلَى بَعْضٍ ۖ  
وَأَتَيْنَا دَاوُودَ زَبُورًا

“Dan Tuhan-mu lebih mengetahui siapa yang (ada) di langit dan di bumi. Dan sesungguhnya telah Kami lebihkan sebagian nabi-nabi itu atas sebagian (yang lain), dan Kami berikan *Zabur* kepada Daud.”(QS.AI-Isra [17]:55)

3. *Injil* yaitu kitab yang diturunkan kepada Isa AS sebagai pedoman hidup bagi kaum Bani Israil. Firman Allah SWT (QS.AI-Maidah [5]:46).

وَقَفَّيْنَا عَلَىٰ آثَارِهِم بِعِيسَى ابْنِ مَرْيَمَ مُصَدِّقًا لِمَا بَيْنَ يَدَيْهِ مِنَ التَّوْرَةِ ۗ وَأَتَيْنَاهُ الْإِنْجِيلَ فِيهِ  
هُدًى وَنُورٌ وَمُصَدِّقًا لِمَا بَيْنَ يَدَيْهِ مِنَ التَّوْرَةِ وَهُدًى وَمَوْعِظَةً لِّلْمُتَّقِي

"Dan Kami iringkan jejak mereka (nabi nabi Bani Israil) dengan Isa putera Maryam, membenarkan Kitab yang sebelumnya, yaitu: *Taurat*. Dan Kami telah memberikan kepadanya Kitab *Injil* sedang didalamnya (ada) petunjuk dan dan cahaya (yang menerangi), dan membenarkan kitab yang sebelumnya, yaitu Kitab *Taurat*. Dan menjadi petunjuk serta pengajaran untuk orang-orang yang bertakwa”.

Isi pokok Kitab *Injil* adalah ajaran untuk hidup dengan lebih mencintai akhiat (zuhud) dan menjauhi kerakusan dan ketamakan dunia. Ini dimaksudkan untuk meluruskan kehidupan orang-orang Yahudi yang materialistis.

4. *Al-Qur'an* adalah kitab yang diturunkan Kepada Nabi, Muhammad SAW sebagai petunjuk hidup umatnya. Berbeda dengan kitab-kitab sebelumnya yang hanya terbatas untuk satu kaum, Al-Qurān tidak hanya diturunkan untuk bangsa Arab, melainkan untuk seluruh umat. Firman Allah SWT (QS.Yusuf [12]:2).

إِنَّا أَنْزَلْنَاهُ قُرْآنًا عَرَبِيًّا لَعَلَّكُمْ تَعْقِلُونَ

“*Sesungguhnya Kami menurunkannya berupa Al Qurān dengan berbahasa Arab, agar kamu memahaminya.*” (QS. Yūsuf [12]:2)

### 2.2.5 Subtema Kitab – Kitab

Berdasarkan sumber yang terdapat pada Al-Qur'an *Cordoba Amazing* (Tafsir Qur'an lengkap) terdapat 299 surat dan ayat serta 6 subtema (*terlampir*). Berikut adalah subtema beriman kepada kitab-kitab :

**Tabel 2.1** Kategori Jumlah Ayat, Surat dan Sub Tema Kitab-Kitab

No	Sub Tema	Surah & Ayat
1	Kewajiban dan Keutamaan Beriman pada Kitab-Kitab.	27
2	Pengakuan antara satu kitab dengan lainnya	29
3	Adanya perubahan dalam beberapa kitab samawi	6
4	Al-Qur'an terpelihara dari penyelewengan.	5
5	Hikmah penurunan kitab-kitab samawai.	98
6	Sikap manusia terhadap sikap samawi	135
TOTAL		300 <i>Surah dan Ayat</i>



### 2.2.6 Text Mining

Menurut (Luh, Sri, Ginantra, & Wardani, 2019) Text Mining merupakan sebuah proses pengamilan data dalam berupa text dari sumber berupa sebuah dokumen. Dengan *text mining* kita dapat mencari kata kunci yang dapat mewakili konten dokumen kemudian melakukan dianalisis dan melakukan pencocokan antara dokumen dan kata kunci yang telah dibuat.

### 2.2.7 Preprocessing

Preprocessing adalah tahap awal sebelum sebelum kumpulan dokumen diklasterisasi. Pada tahap ini, persiapan proses untuk dokumen dan data dilakukan demikian bahwa dokumen / data siap diproses dengan benar. Dengan kata lain tujuan dari preprocessing adalah menyeragamkan *dataset* yang ada agar mudah diolah. (Amalia, Lydia, Fadilla, & Huda, 2018). Adapun juga tahapan dalam preprocessing teks yaitu :

1. *Tokenizing* tahap pemecahan dokumen menjadi sebuah kata per kata. Sebagai contoh: "**dengan menyebut nama allah**", itu akan menjadi dipecah menjadi: '**dengan**', '**menyebut**', '**nama**', '**allah**'.
2. *Stemming* tahapan dimana proses untuk mendapatkan sebuah kata-kata dasar dari sebuah kata lalu menghilangkan akhiran dari kedua awalan, dan akhiran. Contoh "**mendustakan**" menjadi "**dusta**".
3. *Case Folding* tahap menyeragamkan semua jenis huruf pada dokumen. semua karakter dalam teks diproses lalu diubah menjadi huruf kecil.

Misalnya: **‘DENGAN MENYEBUT NAMA ALLAH’** akan diubah menjadi **‘dengan menyebut nama allah’**.

4. *Punctuation* tahapan untuk menghapus semua karakter non *alphabet* misalnya simbol, spasi, titik dan lain-lain.
5. *Stopword* tahapan menghapus kata-kata dimana kata-kata tersebut sering muncul dan yang tidak penting akan dihapus atau dibuang. Contoh kata yang dipertanyakan adalah kata depan, kata sambung, kata keterangan dan kata pengganti, seperti: "**yang**", "**ke**", "**di**", "**sebuah**", "**pada**", "**oleh**", "**ini**", "**dari**".

### 2.2.8 Bag Of Words

*Bag Of Words* merupakan sebuah metode untuk untuk merepresentasikan data set yang berbasis teks seperti dokumen, dan artikel ke dalam suatu matriks. *Bag of word* menyimpan informasi frekuensi kemunculan suatu kata dalam suatu dokumen dalam bentuk matriks 2 dimensi. *Bow* mengansumsikan tidak adanya hubungan antar kata yang terjadi, jadi cara kita melihat letergantungan kata dengan model linier berbasis fitur. Hal pertama yang harus diperhatikan adalah struktur dokumen diabaikan dalam representasi *bag of word*, tetapi dilihat dalam pencarian web. Kedua, dilihat dari representasi terstruktur yang digunakan dalam sistem *database* dalam sebuah mesin pencarian. Hal yang ketiga lihat bagaimana struktur *query* yang digunakan untuk menjawab beberapa pertanyaan dan akhirnya *bag of words* didasarkan pada sebuah kata dan banyaknya aplikasi.(Mardiana & Nyoto, 2015)

### 2.2.9 *Latent Semantic Analysis (LSA)*

Latent Semantic Analysis adalah sebuah metode yang berguna untuk menemukan hubungan, keterkaitan, dan kemiripan antar beberapa dokumen-dokumen yang penggalan dari dokumen dan kata-kata yang muncul pada tiap dokumen dimanfaatkan komputasi statistik untuk merepresentasikan yang digunakan sebagai sebagai arti kata untuk sebuah *corpus* yang besar. *Corpus* itu sendiri memiliki arti kumpulan beberapa teks yang memiliki kesamaan dari tema dan subjek. LSA itu sendiri dapat diterapkan kesemua kumpulan dokumen yang dibersihkan dari beberapa sintaksis dan gramatikal. Prinsip kinerja *Latent Semantic Analysis* dengan cara melakukan pendekatan dari peringkat terendah ke matriks dokumen-dokumen untuk melakukan penghapusan beberapa informasi yang tidak relevan. (Wicaksono & Irawan, 2014).

### 2.2.10 *Term Frequency dan Inverse Documents Frequency (TF-IDF)*

TF-IDF merupakan sebuah metode untuk memberikan bobot dari suatu kata (term) terhadap dari sebuah dokumen. Metode ini penggabungan dua konsep untuk perhitungan bobot yaitu, frekuensi kemunculan kata dalam sebuah dokumen dan inverse frekuensi dokumen yang mengandung kata tersebut. Frekuensi kemunculan kata didalam dokumen yang diberikan menunjukkan seberapa penting kata tersebut didalam dokumen tersebut. Frekuensi dokumen yang mengandung kata tersebut menunjukkan seberapa umum kata tersebut. Sehingga bobot hubungan antara sebuah kata dan sebuah dokumen akan tinggi apabila frekuensi kata tersebut tinggi didalam

dokumen dan frekuensi keseluruhan dokumen yang mengandung kata tersebut yang rendah pada kumpulan dokumen.(Intan & Defeng, 2006).

1. *Term Frequency* (TF) menyatakan banyaknya jumlah kemunculan sebuah term (*TF Tinggi*) dalam dokumen yang saling bersangkutan. Semakin besar jumlah kemunculan suatu term dalam dokumen, semakin besar pula bobot menghasilkan nilai kesesuaian yang semakin besar. Berikut ini Jenis-jenis formula untuk melakukan pembobotan *Term frequency* (TF) :

**Tabel 2.2** Contoh Pembobotan Term Frequency

Skema Pembobotan	TF
Binary	0,1
Raw Count	$f_{t,d}$
Term Frequency	$f_{t,d} / \sum_{t \in d} f_{t',d}$
Log Normalization	$\text{Log}(1 + F_{t,d})$
Double Normalization 0.5	$0,5 + 0,5 * \frac{f_{t,d}}{\max\{t \in d\}} f_{t,d}$

Sumber : (Poletini, 2004)

Keterangan :

(ft,d) : jumlah kata kunci yang muncul dalam satu dokumen

t : kata kunci

d : dokumen



$\max\{t' \in d\}$  : Jumlah kata kunci yang paling sering muncul pada satu dokumen

2. *Invers Document Frequency* (IDF) berfungsi untuk mengurangi dominasi dari term yang sering muncul diberbagai dokumen. Hal tersebut dibutuhkan karena *term* banyak yang muncul diberbagai dokumen yang dapat dianggap sebagai *common term* sehingga menjadikan dia tidak penting nilainya (term umum). Tetapi sebaliknya factor dari jarang munculnya term tersebut harus diperhatikan dalam pemberian bobot. Berikut ini Jenis-jenis formula untuk melakukan pembobotan *Invers Document Frequency* (IDF) :

**Tabel 2.3** Pembobotan Term Invers Document Frequency

Skema Pembobotan	IDF
Binary	1
<i>Inverse Document Frequency</i>	$\log \frac{N}{nt} = -\log \frac{nt}{N}$
<i>Inverse Document Frequency Smooth</i>	$\log \left( \frac{N}{1+nt} \right)$
<i>Inverse Document Frequency max</i>	$\log \left( \frac{\max\{t' \in d\} nt'}{1+nt} \right)$
<i>Probabilistic Inverse Document Frequency</i>	$\log \frac{N-nt}{nt}$

Sumber : (Poletini, 2004)

Keterangan :

N : Jumlah seluruh dokumen

Nt : Jumlah kata yang muncul pada seluruh dokumen

$\max\{t' \in d\} nt'$  : Jumlah kata yang paling banyak muncul pada seluruh dokumen.

### 2.2.11 Singular Value Decomposition (SVD)

*Singular Value Decomposition (SVD)* merupakan sebuah metode saat ini yang dipergunakan untuk mendeteksi objek dalam *image processing*. SVD dimulai dari sebuah faktorisasi pada sebuah matriks tak nol sehingga menghasilkan tiga buah matriks tak nol. Satu diantara tiga buah matriks tersebut akan mengandung nilai-nilai singular dari matriks awalnya. (Ghina et al., 2019)

SVD memiliki tiga sifat yaitu matriks *hanger* (matriks  $U$ ), *aligner* (matriks  $V$ ) dan *stretcher* (matriks  $S$ ). Pada matriks  $U$  dan  $V$  jika dikalikan maka bentuk dari matriks tidak akan berubah, karena sifat kedua matriks tersebut menjaga bentuk objek. Sedangkan pada matriks *Stretcher* ( $S$ ) merupakan matriks diagonal, hasil yang didapatkan jika sumbu matriks dikalikan dengan matriks ( $S$ ) terhamparnya semua nilai pada kurva sepanjang sumbu  $x$  dan  $y$ .

Adapun rumus dari SVD adalah sebagai berikut :

$$A_{td} \approx U_{tm} S_{mm} V_{md}^T \dots \dots \dots (2.1)$$

Keterangan :

- A : matriks asal
- U : matriks eigenvector dari  $AA^T$
- S : matriks diagonal
- $V^T$  : transpose dari matriks  $V$
- t : jumlah baris matriks

- d : jumlah kolom matriks  
 m : rank, dimana  $\text{rank}(\leq \min(t,d))$

### 2.2.12 Evaluation Measure

Menurut (Santoso, Virginia, & Lukito, 2017) Evaluasi dilakukan untuk menguji hasil *klasifikasi* dengan mengukur nilai kebenaran dari sistem. Parameter yang digunakan untuk mengukur nilai kebenaran adalah *precision*, *recall*, dan *F-Measure*. *Precision* berfungsi untuk mengidentifikasi sebuah kualitas dari pengelompokan sistem, *Recall* mengidentifikasi kualitas dari sistem, sedangkan *F-Measure* adalah pengukuran kualitas dari akurasi sebuah pengelompokan. Adapun rumus perhitungan evaluasi sebagai berikut :

$$Precision = \frac{TP}{TP+FP} \dots\dots\dots (2.2)$$

$$Recall = \frac{TP}{TP+FN} \dots\dots\dots (2.3)$$

$$F-1 = 2 * \frac{Precision \times recall}{precision+recall} \dots\dots\dots (2.4)$$

Keterangan :

*True Positive* (TP) = suatu kondisi dimana sistem mendeteksi kelas positif dan faktanya pun positif.

*True Negative* (TN) = suatu kondisi dimana sistem mendeteksi kelas negatif dan faktanya pun negatif.

*False Positive* (FP) = suatu kondisi dimana sistem mendeteksi kelas positif dan faktanya pun negatif.

*False Negative* (FN) = suatu kondisi dimana sistem mendeteksi kelas negatif dan faktanya pun positif.

### 2.2.13 Python







Python adalah sebuah bahasa pemrograman yang bersifat open source. Python dioptimalisasikan untuk *software quality*, *developer productivity*, *program portability*, dan *component integration*. Python dipercaya untuk mengembangkan berbagai macam perangkat lunak, seperti internet scripting, systems programming, user interfaces, product customization, numeric programming. Python itu sendiri telah menduduki peringkat 4 atau 5 untuk software yang sering digunakan diseluruh dunia. (Harismawan, Kharisma, & Afirianto, 2018)




### 2.2.14 Flowchart

*Flowchart* adalah Bagian yang mempunyai arus yang menggambarkan langkah-langkah penyelesaian suatu masalah. *Flowchart* merupakan cara penyajian dari suatu algoritma bagan-bagan yang mempunyai arus yang menggambarkan langkah-langkah penyelesaian suatu masalah. Simbol *flowchart* dan fungsinya dapat dilihat pada tabel sebagai berikut (Syarif & Jurusan Teknik Informatika Fakultas Teknologi Industri, 2016).



Tabel 2.4 Simbol dan Fungsi Flowchart

No	Simbol	Nama	Fungsi
1		Terminator	Permulaan / pengakhiran program
2		Flow Line	Arah aliran program
3		Preparation	Proses inisialisasi/pemberian nilai awal
4		Process	Proses pengolahan data
5		Input/Output Data	Proses input/output data, parameter, informasi
6		Predefined Process	Permulaan sub program / proses menjalankan sub program

No	Simbol	Nama	Fungsi
7		Decision	Perbandingan, pernyataan, penyeleksian data yang memberikan pilihan untuk langkah selanjutnya
8		On Page Connector	Penghubung pada bagian-bagian flowchart yang berada pada satu halaman
9		Off page Connector	Penghubung, bagianagian flowchart yang berada pada halaman berbeda

## BAB III

### METODOLOGI PENELITIAN

#### 3.1 Alat dan Bahan Penelitian

Alat dan bahan yang digunakan dalam penelitian ini adalah sebagai berikut:

##### 3.1.1 Teknik pengumpulan data

Metodologi penelitian merupakan tahapan-tahapan yang akan dilalui peneliti untuk memperoleh gambaran yang jelas mengenai penelitian, maka penyusun metodologi penelitian sebagai berikut :

##### a. *Data Collecting*

Data yang dikumpulkan yaitu semua kumpulan surat, ayat dan terjemahan yang tertera didalam Tafsir Al-Qur'an yang kemudian di sortir kembali berdasarkan tema beriman kepada kitab-kitab. Data tersebut diperoleh dengan melakukan pengambilan data terhadap Al-Qur'an *Cordoba Amazing* yang terpercaya.

##### b. *Studi Literatur*

Studi literatur dapat dilakukan dengan cara mengumpulkan serta mempelajari segala macam informasi yang berhubungan dengan Pengelompokan Tematik Al-Qur'an Menggunakan Metode LSA dan Pembobotan *Raw Count* dan *IDF-Smooth*.

### c. Penyusunan Laporan Penelitian

Penyusunan laporan dilakukan pada tahap sebagai dokumentasi yang bertujuan untuk menjelaskan alur dari Pengelompokan Tematik Al-Qur'an Menggunakan Metode LSA dengan Pembobotan *Raw Count* dan *IDF-Smooth* agar memudahkan pengimplementasian kepada peneliti selanjutnya.

#### 3.1.2 Spesifikasi Kebutuhan Perangkat Keras (*Hardware*)

Personal komputer atau laptop digunakan untuk perancangan sistem dengan spesifikasi *hardware* sebagai berikut :

1. *Processor* : Intel Core i3-4030U
2. *Ram* : 2,00 GB
3. *Hardisk* : 500 GB
4. *Sysitem Type* : 64-bit *Operating System*

#### 3.1.3 Spesifikasi Kebutuhan Perangkat Lunak (*Softwate*)

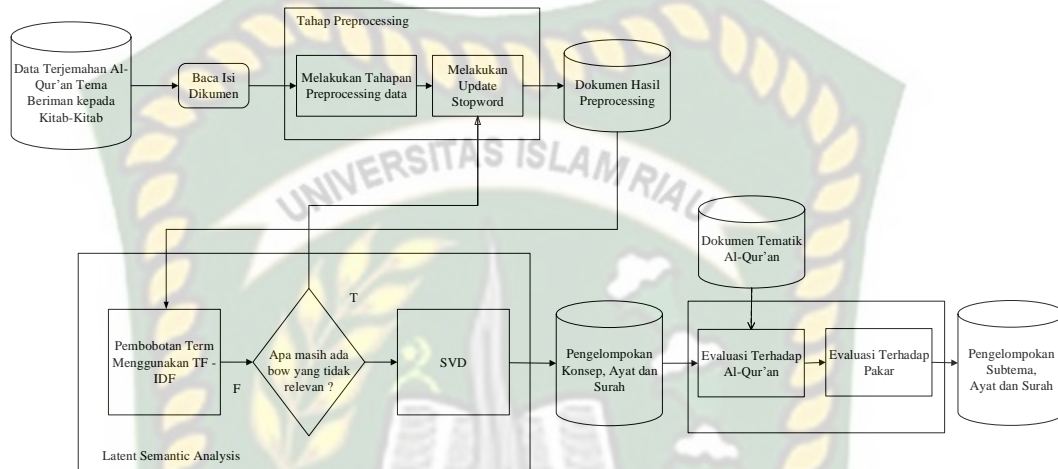
Perangkat lunak (*software*) yangdigunakan dalam pembuatan pengelompokan tematik Al-Qur'an menggunakan metode LSA dengan pembobotan *Raw Count* dan *IDF-Smooth* ini adalah sebagai berikut :

1. Microsoft Windows 10 Professional (64-bit)
2. Python 3.7 (64-bit)
3. Jupyter Notebook
4. *Microsoft Office Visio* 2013



### 3.2 Gambaran Alur Penelitian

Alur pengelompokan Tematik Al-Qur'an dapat digambarkan melalui rancangan yang akan dibuat pada gambar 3.1 dibawah ini ;



**Gambar 3.1 Alur Perancangan Sistem**

Pada alur perencanaan diatas dapat dijelaskan sebagai berikut :

1. Data terjemahan Al-Qur'an tema beriman kepada kitab – kitab menjadi inputan dasar sistem sudah berupa *format csv*.
2. Kemudian data yang telah di baca oleh sistem akan dilakukan tahap *preprocessing* hingga menghasilkan *bag of word*.
3. Hasil dari tahap *preprocessing* berupa *bag of word* akan diproses mengguakan metode LSA sampai menghasilkan pengelompokan *bag of word*, surah dan ayat.
4. Pada tahapan *latent semantic analysis* akan dilakukan pembobotan *bag of word* dari dokumen hasil *preprocessing*. Kemudian akan dilakukan pengelompokan dari masing-masing hasil nilai *bag of word* terhadap

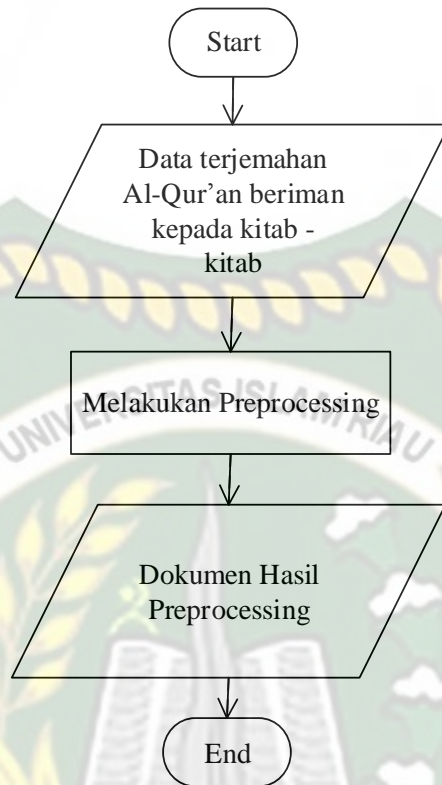
dokumen. Nilai tersebut yang akan menjadi acuan dalam proses pengelompokan atau *singular value decomposition* (SVD).

5. Pada Tahap akhir, dilakukan evaluasi terhadap data *ayat, surah* yang sudah sesuai dengan subtema yang di ambil dari Al-Qur'an *Cordoba Amazing*, sehingga data dapat di *akurasi* kecocokan dari evaluasi yang dilakukan. Pada tahapan ini juga menghasilkan surah dan ayat yang ada didalam hasil SVD tetapi tidak terdapat didalam Al-Qur'an *Cordoba*, yang kemudian akan di evaluasi kembali oleh seorang *ahli pakar* tematik Al-Qur'an. Pada penelitian ini penulis menggunakan kuisisioner sebagai alat mediasi untuk kemudian diisi oleh pakar tematik tersebut. Adapun pakar yang digunakan untuk menilai hasil evaluasi pengelompokan terdiri dari 2 orang yaitu bapak *Dr. Hamzah, M.Ag* selaku dosen dan wakil dekan II di prodi pendidikan agama islam (Universitas Islam Riau) dan bapak *Musaddad Harahap, M.Pdi* yang merupakan dosen serta sekretaris prodi agama islam serta dosen di (Universitas Islam Riau).

### **3.3 Perancangan Sistem**

#### **3.3.1 Tahapan Preprocessing**

Tahapan dimana sebuah proses preprocessing melakukan seleksi data atau pembersihan yang akan diproses pada terjemahan Al-Qur'an dengan tema beriman kepada kitab – kitab. Adapun tahapan *preprocessing* adalah sebagai berikut aplikasi dengan cara :



**Gambar 3.2** *Flowchart Preprocessing*

Dari alur *flowchart* diatas dapat dijelaskan bahwa :

1. *Tokenizing* tahapan dimana melakukan pemotongan string inputan berdasarkan tiap kata dan penyusunannya. Tahap ini sangat dibutuhkan untuk pencarian *bag of word* dengan dokumen yang sedang di proses. Contoh tahap tokenizing dapat dilihat pada tabel 3.1 dibawah ini.

**Tabel 3.1** Contoh Proses Tokenizing

Sebelum	Sesudah
ya allah berkahilah kami dalam rezeki yang telah engkau berikan kepada kami dan peliharalah kami dari siksa api neraka	ya    allah    berkahilah    kami    dalam    rezeki    yang    telah    engkau    berikan    kepada    kami    dan    peliharalah    kami    dari    siksa    api    neraka

2. *Stemming* tahapan dimana untuk memperkecil jumlah indeks yang berbeda dari suatu dokumen menjadi kata dasar (*root word*), juga untuk melakukan pengelompokan kata-kata yang memiliki kata dasar dan arti yang berbeda namun memiliki arti yang serupa. Contoh tahap stemming dapat dilihat pada tabel 3.2 dibawah ini.

**Tabel 3.2** Contoh Proses Stemming

Sebelum	Sesudah
Ya Allah berkahilah kami dalam rezeki yang telah Engkau berikan kepada kami dan peliharalah kami dari siksa api neraka	Ya    Allah    berkah    kami    dalam    rezeki    telah    Engkau    beri    kepada    kami    pelihara    kami    siksa    api    neraka

3. *Case Folding* adalah pengkonversian seluruh teks dalam dokumen menjadi bentuk standar (*huruf kecil atau lowercase*). Case folding biasanya dilakukan untuk mengubah dokumen yang memiliki huruf kapital menjadi huruf kecil. Tahapan ini dilakukan untuk lebih mempermudah saat melakukan pencarian *bag of word*. Contoh tahap case holding dapat dilihat pada tabel 3.3 dibawah ini.



**Tabel 3.3** Contoh Proses Case Folding

Sebelum	Sesudah
“Ya Allah” berkahilah kami dalam rezeki yang telah Engkau berikan kepada kami dan peliharalah kami dari siksa api neraka.	“    ya    allah    “    berkahilah    kami    dalam    rezeki    yang    telah    engkau    berikan    kepada    kami    dan    peliharalah    kami    dari    siksa    api    neraka

4. *Punctuation* merupakan tahapan dimana melakukan penghapusan tanda baca setelah dilakukan proses *case folding*. Menghilangkan tanda baca dapat membantu proses pencocokan *bag of word* terhadap surah dan ayat sehingga lebih akurat. Contoh tahapan *punctuation* dapat dilihat pada tabel 3.4 dibawah ini :

**Tabel 3.4** Contoh Proses Punctuation

Sebelum	Sesudah
“Ya Allah” berkahilah kami dalam rezeki yang telah Engkau berikan kepada kami dan peliharalah kami dari siksa api neraka.	“    ya    allah    “    berkahilah    kami    dalam    rezeki    yang    telah    engkau    berikan    kepada    kami    dan    peliharalah    kami    dari    siksa    api    neraka    .

5. *Stopword Removal*

Tahapan yang dilakukan dengan cara mengambil kata-kata yang dianggap penting dari hasil tokenization atau membuang kata-kata yang dianggap tidak perlu didalam proses text-mining untuk mencari sebuah *bag of word* di dalam dokumen. Contoh tahap *stopword* dapat dilihat pada tabel 3.5 dibawah ini.

Tabel 3.5 Contoh Stopword

Sebelum	Sesudah
Ya Allah berkahilah kami dalam rezeki <b>yang</b> telah Engkau berikan kepada kami <b>dan</b> peliharalah kami <b>dari</b> siksa api neraka	Ya    Allah    berkahilah    kami    dalam    rezeki    telah    Engkau    berikan    kepada    kami    peliharalah    kami    siksa    api    neraka

### 3.3.2 Bag Of Word

Gambaran sederhana untuk pengolahan bahasa alami dan pencarian informasi, representasi sebuah teks yang sangat penting untuk mendukung proses analisis data statistik di dalamnya. Data teks yang tidak terstruktur dapat direpresntasikan secara sederhana menggunakan sekumpulan set kata. *Bag of words* itu sendiri nantinya akan di implementasikan untuk melakukan pembobotan menggunakan metode *latent semantic analysis*. Beberapa kata *bag of word* yang terdapat didalam tema beriman kepada kitab-kitab yang telah dianalisis oleh penulis menggunakan terjemahan yang terdapat dalam rujukan Al-Qur'an *cordobaamazing* sesuai dengan studi kasus penulis yaitu "Beriman Kepada Kitab-Kitab". Berikut adalah hasil dari contoh dari *bag of word* tema beriman kepada kitab-kitab setelah dilakukannya tahapan *preprocessing*. Berikut adalah contoh *bag of word* pada tema beriman kepada kitab-kitab pada tabel 3.6 :

**Tabel 3.6** Contoh Bag Of Word Tema Beriman  
Kepada Kitab-Kitab

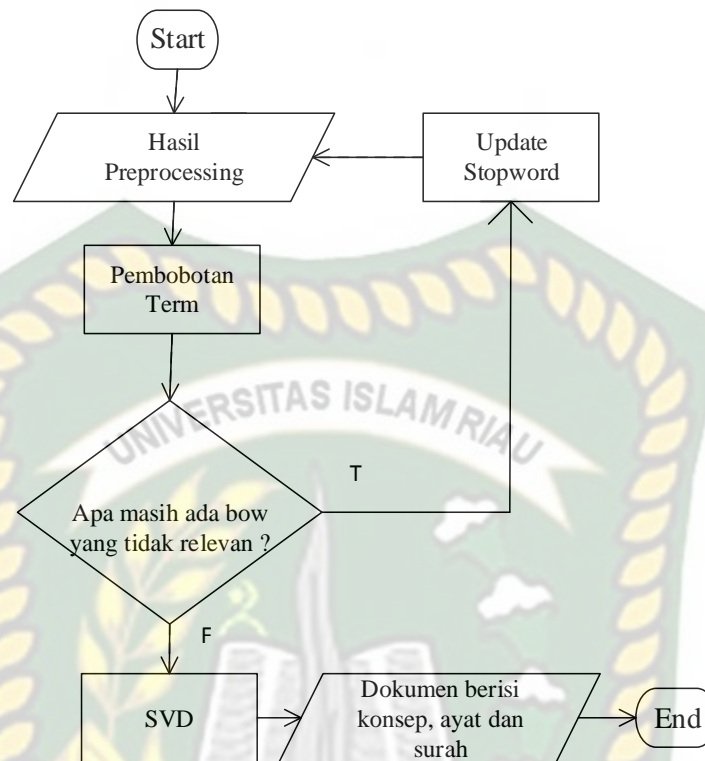
No	Bag Of Words
1	Shalat
2	Augerah
3	Taurat
4	Mukjizat
5	Isa
6	Musa
7	Zabur

### 3.3.3 Latent Semantic Analysis

Sebuah metode yang digunakan untuk menemukan hubungan, keterkaitan, kemiripan antar dokumen-dokumen dan kata-kata yang sering kali muncul pada dokumen dengan memanfaatkan komputasi statistik yang ada, untuk menggali dan merepresentasikan konteks yang digunakan sebagai sebuah arti kata untuk kumpulan sejumlah teks yang memiliki kesamaan tema dan subjek (*corpus*).

Metode LSA menerima masukan (*input*) berupa dokumen teks yang selanjutnya akan dibandingkan kata-kata unik yang digunakan atau yang ada pada dokumen kemudian direpresentasikan sebagai matriks, dimana indeks dokumen-dokumen yang dibandingkan merupakan kolom matriks, kata unik (*term*) merupakan baris matriks, dan nilai dari matriks tersebut adalah banyaknya atau frekuensi kemunculan sebuah kata (*term*) di setiap dokumen

Berikut ini adalah gambaran dan penjelasan dari alur Latent Semantic Analysis yang dapat dilihat dari Gambar 3.3 sebagai berikut :



**Gambar 3.3** Flowchart Latent Semantic Analysis

Berdasarkan alur flowchar pada gambar 3.3 di atas dapat dijelaskan bahwa setelah hasil *preprocessing* dilakukan maka selanjutnya melakukan pembobotan untuk setiap kata atau *bag of word* yang akan diproses. Selajutnya, kata yang sudah di diberi bobot akan dicek kembali apakah kata atau *bag of word* belum relevan atau sesuai, jika belum maka proses akan kembali ke *update stopwords* untuk menemukan kembali *bag of word* yang sesuai. Tetapi, Jika sudah relevan maka langkah selanjutnya adalah kata yang sudah dibobotkan tadi akan diproses menggunakan *singular value decomposition* (SVD) untuk mengelompokan perkonsep. Setelah



proses SVD selesai maka akan menghasilkan sebuah dokumen berisi konsep, ayat dan surah.

### 3.3.4 TF-IDF

Sebuah metode yang dilakukan untuk melakukan pembobotan (term) terhadap sebuah dokumen. Metode ini memiliki cara penggabungan dua buah konsep. Pada tahap ini penulis akan melakukan pembobotan term dengan menggunakan TF-IDF. Untuk TF menggunakan pembobotan *TF-Raw Count* dengan formula :

$$\text{Row Count} = f_{t,d} \dots\dots\dots (3,1)$$

Berikut adalah sebagian contoh data terjemahan dari 299 data tema beriman kepada kitab-kitab yang peneliti gunakan untuk menghitung jumlah kemunculan *bag of word* pada tabel 3.7 dibawah ini :

**Tabel 3.7** Dokumen yang akan dihitung

No	Dokumen	BoW	Terjemahan
1	D1	[shalat, anugerah]	(yaitu) mereka yang beriman kepada yang ghaib, yang mendirikan <b>shalat</b> , dan menafkahkan sebahagian rezeki yang Kami <b>anugerah</b> kan kepada mereka.
2	D2	[mukjizat, mukjizat, zabor]	Jika mereka mendustakan kamu, maka sesungguhnya rasul-rasul sebelum kamupun telah didustakan (pula), mereka membawa <b>mukjizat-mukjizat</b> yang nyata, <b>zabor</b> dan kitab yang memberi penjelasan yang sempurna.

No	Dokumen	BoW	Terjemahan
3	D3	[taurat, musa, mukjizat, isa]	Dan sesungguhnya Kami telah mendatangkan Al Kitab ( <b>taurat</b> ) kepada <b>musa</b> , dan Kami telah menyusulinya (berturut-turut) sesudah itu dengan rasul-rasul, dan telah Kami berikan bukti-bukti kebenaran ( <b>mukjizat</b> ) kepada <b>isa</b> putera Maryam dan Kami memperkuatnya dengan Ruhul Qudus. Apakah setiap datang kepadamu seorang rasul membawa sesuatu (pelajaran) yang tidak sesuai dengan keinginanmu lalu kamu menyombong; maka beberapa orang (diantara mereka) kamu dustakan dan beberapa orang (yang lain) kamu bunuh?
4	D4	Taurat	Tidakkah kamu memperhatikan orang-orang yang telah diberi bahagian yaitu Al Kitab ( <b>taurat</b> ), mereka diseru kepada kitab Allah supaya kitab itu menetapkan hukum diantara mereka; kemudian sebahagian dari mereka berpaling, dan mereka selalu membelakangi (kebenaran).

Untuk kemunculan *bag of word* pada tabel 3.8 didapatkan dari kemunculan *bag of word* pada contoh dokumen yang kan dihitung, Contoh kata *bag of word* dapat dilihat pada tabel 3.8 dibawah ini :

**Tabel 3.8** Kata Kunci (Bag Of Word)

No	Bag Of Words
1	Shalat
2	Augerah
3	Taurat
4	Mukjizat
5	Isa
6	Musa
7	Zabur

Pada tahapan setelah kata kunci sudah ditentukan maka dilakukan pengecekan kemunculan kata kunci bag of word dalam tiap-tiap dokumen yang telah ditentukan. Contoh kemunculan bag of word dapat dilihat pada tabel 3.9 dibawah ini :

**Tabel 3.9** Kemunculan Bag Of Words TF

<i>Term</i>	shalat	anugerah	taurat	musa	mukjizat	isa	zabur
Terjemahan ['shalat', 'anugerah']	1	1	0	0	0	0	0
['mukjizat', 'mukjizat', 'zabur']	0	0	0	0	2	0	1
['taurat', 'musa', 'mukjizat', 'isa']	0	0	1	1	1	1	0
['taurat']	0	0	1	0	0	0	0

1. Berdasarkan ketentuan yang berlaku terhadap formula *TF Raw Count* ( $f_{t,d}$ ) yaitu angka 1 menyatakan adanya kemuculan kata kunci sedangkan angka 0 menyatakan tidak terdapatnya *bag of word* didalam dokumen. Apabila kata kunci yang ditentukan muncul lebih dari satu kali didalam dokumen yang sama maka, kata kunci tersebut tetap dihitung satu kali kemunculan saja perdokumen.

2. Formula yang dipakai untuk menentukan *Inverse Documents Frequency* (IDF)-*Smooth* adalah pada  $\mathbf{IDF} = \log\left(\frac{N}{1+nt}\right)$  dengan keterangan **N** menyatakan jumlah seluruh dokumen sedangkan **nt** menyatakan jumlah kemunculan kata kunci pada seluruh dokumen. Untuk penghitungan bobot *bag of words* pada *Invers Document Frequency* dilakukan dengan cara menghitung jumlah kemunculan *bag of words* pada keseluruhan dokumen, dibagi dengan satu, kemudian ditambah dengan jumlah kata *bag of words* yang muncul. Berikut adalah contoh tabel kemunculan kata kunci pada *Invers Document Frequency* :

**Tabel 3.10** Hasil Perhitungan IDF

anugerah	isa	taurat	mukjizat	shalat	musa	zabur
0,778	0,778	0.845	0.845	0,778	0,778	0,778

Pada tabel 3.10 dapat dilakukan perhitungan dengan cara seperti berikut :

$$\text{anugerah} = \text{IDF} = \log\left(\frac{5}{1+1}\right) = 0.778$$

$$\text{isa} = \text{IDF} = \log\left(\frac{5}{1+1}\right) = 0.778$$

$$\text{taurat} = \text{IDF} = \log\left(\frac{5}{1+2}\right) = 0.845$$

$$\text{mukjizat} = \text{IDF} = \log\left(\frac{5}{1+2}\right) = 0.845$$

$$\text{shalat} = \text{IDF} = \log\left(\frac{5}{1+1}\right) = 0.778$$



$$\text{musa} = \text{IDF} = \log\left(\frac{5}{1+1}\right) = 0.778$$

$$\text{zabur} = \text{IDF} = \log\left(\frac{5}{1+1}\right) = 0.778$$

3. Berbeda dengan cara perhitungan TF dan IDF untuk cara menentukan bobot *bag of words* pada TF-IDF dilakukan dengan cara mengalikan nilai akhir dari masing-masing dari TF dan IDF.

**Tabel 3.11** Hasil Perhitungan TF-IDF Smooth

	shalat	anugerah	taurat	musa	mukjizat	isa	zabur
0	0.778	0.778	0	0	0	0	0
1	0	0	0	0	1.690	0	0.778
2	0	0	0.845	0.778	0.845	0.778	0
3	0	0	0.845	0	0	0	0

Terakhir untuk mencari nilai bobot pada  $TF \times IDF$  dapat dilakukan dengan rumus  $TF \times IDF$ . Misalkan :

**shalat** = (“shalat”, jumlah kata perdokumen (**Ftd**)) x **Log**(“shalat”, jumlah seluruh dokumen (**N**)/1+ jumlah kata “shalat” diseluruh dokumen (**nt**)) = 1 x 0.778 = **0.778**

$$\text{augerah} = \text{TF} \times \text{IDF} = 1 \times 0.778 = 0.778$$

$$\text{taurat} = \text{TF} \times \text{IDF} = 1 \times 0.845 = 0.845$$

$$\text{TF} \times \text{IDF} = 1 \times 0.845 = 0.845$$

$$\text{musa} = \text{TF} \times \text{IDF} = 1 \times 0.778 = 0.778$$

$$\text{mukjizat} = \text{TF} \times \text{IDF} = 1 \times 0.845 = 0.845$$

$$\text{TF} \times \text{IDF} = 2 \times 0.845 = 1.690$$

$$\text{isa} = \text{TF} \times \text{IDF} = 1 \times 0.778 = 0.778$$

$$\text{zabur} = \text{TF} \times \text{IDF} = 1 \times 0.778 = 0.778$$

### 3.3.5 Singular Value Decomposition (SVD)

Dalam tahapan ini, nilai *Singular Value Decomposition* (SVD) digunakan untuk mempermudah melakukan pengolahan data serta memperkecil ukuran matriks dari hasil perhitungan TF-IDF. SVD juga mempunyai kemampuan untuk melakukan pemodelan, sehingga dapat meningkatkan nilai akurasi dan menghilangkan bagian yang dirasa tidak penting didalam dokumen. SVD diproses oleh sistem secara otomatis menggunakan *library python*, sehingga akan menghasilkan sebuah pengelompokan berdasarkan jumlah subtema yang sudah terbentuk. Selain itu pada proses SVD akan menghasilkan surah dan ayat secara otomatis berdasarkan hasil dari nilai SVD. Adapun untuk *ilustrasi SVD* dapat dilihat pada tabel 3.12 dibawah ini :

**Tabel 3.12** Contoh Data Awal SVD (*U*)

Dokumen \ BOW	D1	D2	D3	D4	D5
Anugerah					
Taurat					
Mukjizat					
Isa					
Musa					

Pada tabel 3.12 merupakan contoh dari data awal SVD, dimana *bag of word* dinyatakan dalam kolom yang berwarna orange muda. Bow yang sudah muncul kemudian dikelompokkan berdasarkan dokumen. Dapat dilihat pada tabel 3.13 dibawah ini :

**Tabel 3.13** Contoh Pengelompokan SVD ( $U$ )

Dokumen \ BOW	D2	D4	D5	D1	D3
Anugerah					
Taurat					
Mukjizat					
Isa					
Musa					

Pada tabel 3.13 merupakan proses dari sebuah pengelompokan SVD matrik  $U$  dimana pengelompokan dokumen berdasarkan dari kemunculan sebuah *bag of word*, yang kemudian akan saling berkelompok secara otomatis. Langkah terakhir adalah pembentukan sebuah konsep berdasarkan dokumen yang sudah terkelompok. Maka hasil dari pengelompokan dapat dilihat pada tabel 3.14 :

**Tabel 3.14** Contoh Pembentukan Konsep SVD ( $U$ )

Dokumen \ BOW	D2	D4	D5	D1	D3
Anugerah					K1
Taurat		K2			
Mukjizat				K3	
Isa		K4			
Musa	K5				

Berdasarkan pada tabel 3.14 dapat disimpulkan bahwa terbentuk 5 buah konsep dari hasil pengelompokan dokumen berdasarkan *bag of word*.

### 3.3.6 Evaluasi

Tahapan yang dilakukan untuk menghitung sebuah persentasi dari ketepatan dan kecocokan dari hasil keseluruhan antara surah, dan ayat masing masing sub tema, pada pengelompokan yang dilakukan dengan tahapan *latent semantic analysis* juga merujuk kepada yang tertera dalam Al-Qur'an *Cordoba Amazing*. Pada penelitian kali ini penulis menggunakan teknik F-measure untuk mengukur suatu kinerja dari model yang diusulkan. Pengukuran dilakukan berdasarkan pada nilai *Precision Recall*. Apabila semakin tinggi nilai dari *Precision Recall* maka akan semakin bagus untuk tingkat akurasi. Data hasil evaluasi sistem, kemudian akan di evaluasi kembali oleh seorang ahli pakar. Pada tahapan ini pakar tersebut menggunakan buku tafsir "*index Al-Qur'an*" dengan pengarang H. Suhardi (*telampir*).



## BAB IV

### HASIL DAN PEMBAHASAN

#### 4.1 Data Inputan

##### 4.4.1 Data Tematik Al-Qur'an Cordoba

Dari data tematik Al-Qur'an yang ada saat ini peneliti menggunakan *tafsir* Al-Qur'an *Cordoba Amazing* sebagai acuan data yang digunakan selama penelian berlangsung. Adapun tema yang dibahas adalah beriman kepada kitab-kitab. Tema beriman kepada kitab-kitab itu sendiri terdiri 6 sub tema, 278 surah, ayat serta terjemahannya (*terlampir*). Data tersebut dijadikan inputan berupa data *csv* yang kemudian akan diproses menggunakan bahasa pemrograman *python*. Berikut adalah sebagian contoh data tema beriman kepada kitab-kitab berdasarkan tabel 4.1 dibawah ini :

**Tabel 4.1** Contoh Data Tematik Cordoba Tema Kitab-Kitab

No	Surah	Ayat	Terjemahan
1	2	3	(yaitu) mereka yang beriman kepada yang ghaib, yang mendirikan shalat, dan menafkahkan sebahagian rezeki yang Kami anugerahkan kepada mereka.
2	2	4	dan mereka yang beriman kepada Kitab (Al Quran) yang telah diturunkan kepadamu dan Kitab-kitab yang telah diturunkan sebelumnya, serta mereka yakin akan adanya (kehidupan) akhirat.

No	Surah	Ayat	Terjemahan
3	2	87	Dan sesungguhnya Kami telah mendatangkan Al Kitab (taurat) kepada musa, dan Kami telah menyusulinya (berturut-turut) sesudah itu dengan rasul-rasul, dan telah Kami berikan bukti-bukti kebenaran (mukjizat) kepada isa putera Maryam dan Kami memperkuatnya dengan Ruhul Qudus. Apakah setiap datang kepadamu seorang rasul membawa sesuatu (pelajaran) yang tidak sesuai dengan keinginanmu lalu kamu menyombongmaka beberapa orang (diantara mereka) kamu dustakan dan beberapa orang (yang lain) kamu bunuh?
4	2	136	Katakanlah (hai orang-orang mukmin): "Kami beriman kepada Allah dan apa yang diturunkan kepada kami, dan apa yang diturunkan kepada Ibrahim, Isma'îl, Ishaq, Ya'qub dan anak cucunya, dan apa yang diberikan kepada Musa dan Isa serta apa yang diberikan kepada nabi-nabi dari Tuhannya. Kami tidak membeda-bedakan seorangpun diantara mereka dan kami hanya tunduk patuh kepada-Nya".
5	2	176	Yang demikian itu adalah karena Allah telah menurunkan Al Kitab dengan membawa kebenaran; dan sesungguhnya orang-orang yang berselisih tentang (kebenaran) Al Kitab itu, benar-benar dalam penyimpangan yang jauh (dari kebenaran).

(Data Terlampir)

## 4.2 Tahapan Preprocessing

Pada tahapan ini akan dilakukan pengolahan seluruh data terjemahan tematik Al-Qur'an *Cordoba Amazing* tema beriman kepada kitab-kitab. Untuk melakukan *preprocessing* ada beberapa tahapan, yaitu *tokenizing*, *stemming*, *case folding*, *punctuation* dan *stopwords*. Berikut adalah langkah-langkah serta hasil dari *preprocessing* yang dijelaskan pada tabel 4.2 dibawah ini :

**Tabel 4.2** Tampilan langkah-langkah dan hasil preprocessing

	Surah	Ayat	Terjemahan	Tokenizing	Stem_CaseFal	Panctuation	Stopword	Bag Of Word
1	2	3	(yaitu) mereka yang beriman kepada yang ghaib...	[(, yaitu, ), mereka, yang, beriman, kepada, y..	[, yaitu, , mereka, yang, iman, kepada, yang, ..	[yaitu, mereka, yang, iman, kepada, yang, ghai..	[shalat, anugerah]	Shalat anugerah
2	2	4	dan mereka yang beriman kepada Kitab (Al Quran...	[dan, mereka, yang, beriman, kepada, Kitab, (...	[dan, mereka, yang, iman, kepada, kitab, , al..	[dan, mereka, yang, iman, kepada, kitab, al..	[ ]	
3	2	87	Dan sesungguhnya Kami telah mendatangkan Al Ki...	[Dan, sesungguhnya , Kami, telah, mendatangka n,.	[dan, sungguh, kami, telah, datang, al, kitab,..	[dan, sungguh, kami, telah, datang, al, kitab,..	[taurat, musa, mukjizat, isa]	taurat musa mukjizat isa
4	2	136	Katakanlah (hai orang-orang mukmin): "Kami ber..	[Katakanlah, (, hai, orang-orang, mukmin, ), :...	[kata, , hai, orang, mukmin, kami, iman,..	[kata, hai, orang, mukmin, kami, iman, kepada,..	[ibrahim, musa, isa]	Ibrahim musa isa
5	2	176	Yang demikian itu adalah karena Allah telah me...	[Yang, demikian, itu, adalah, karena, Allah, t...	[yang, demikian, itu, adalah, karena, allah,..	[yang, demikian, itu, adalah, karena, allah, t...	[ ]	

Dari tabel 4.2 terdapat langkah-langkah dan hasil *preprocessing* dari data terjemahan Al-Qur'an *Cordoba* tema beriman kepada kitab-kitab. Pada tahapan pertama *tokenizing* merupakan sebuah proses pemisahan kalimat dari terjemahan Al-

Qur'an *Cordoba* bertemakan beriman kepada kitab-kitab. Kemudian, kalimat yang sudah dipisahkan tersebut akan dilakukan tahap selanjutnya yaitu *Stemming case folding* yaitu proses membuang kata ber-imbuan menjadi kata dasar serta pengubahan beberapa huruf kapital menjadi huruf kecil. Setelah semua kalimat berubah menjadi huruf kecil, maka tahapan selanjutnya adalah *punctuation* dimana tahapan yang dilakukan untuk menghilangkan tanda baca yang berada didalam kalimat. Setelah tanda baca sudah dihapuskan maka tahapan selanjutnya adalah *stopword* penghapusan kata yang tidak masuk dalam kategori tema beriman kepada kitab-kitab, penghapusan penambahan tersebut dilakukan berdasarkan *stopword* bahasa Indonesia dan kebutuhan yang ada. Contoh penambahan *stopword* tema beriman kepada kitab-kitab dapat dilihat pada tabel 4.3 :

**Tabel 4.3** Penambahan Stopword

Penambahan Stopword
jika, nya, orang, al, kurangnya, baiknya, berkali, kali, mata, olah, israil, buruk, ayat, azab, kabar, siksa, ragu, jalan, maha, sombong, bukti, nabi, sihir, putus, sisi, neraka, lurus, tatkala, al, saksi, sentuh, pedih, bohong, rupa, adakan, umat, atur, tangan, utus, semesta, ambil, mudah, kali, takwa, akal, laku, barangsiapa, tanggung, acuh, wahyu, zalim, simpang, pelihara, mulia, kelak, ingkar, kaum, sombong, pedih, masuk, acuh, agung, zakat, gembira, kabarnya, paham, arab, bahasa, temu, tolong, hati, adakannya, mengada, yatim, acuh, tuhan, orang, dusta, manusia, terang, kafir, rasul, baca, hadap, iman, huni, turun, agung, larang, akhirat, ikan, selisih, rahmat, sempurna, zalim, nyata, dusta, tiada, bawa, alam, dongeng, golong, dengar, niscaya, kabar, ragu, buah, utus, neraka, nikmat, bantah, haram, lihat, hidup, jahannam, gila, kisah, hai, serah, tajam, tuduh, maha, balas, sesat, perkasa, hamba, cahaya, berat, agama, memecah, utus, pecah, anut, maha, sihir, akibat, surat, al, turun, hukum, sebah, wajib, utus, sekurang,



### Penambahan Stopword

setidak, tama, tidaknya, quran, kitab, turun, simpang, sungguh, bawa, selisih, berita, gembira, izin, musuh, alir, bangkit, barang, bumi, ampun, hitung, kuasa, lahir, surga, langit, niscaya, sembunyi, hati, adil, bapa, biarpun, enggan, hawa, kaum, kaya, amal, ismail, mukmin, kerja, maslahat, hadap, miskin, nafsu, putar, tegak, air, kerabat, kecuai, nyala, acuh, ada, agung, ahmad, aikah, ajam, ajar, alas, alangkah, as, alif, alim, ancam, andaikata, anggap, angsur, hikmah, allah, aniaya, api, asing, atur, bahagian, balau, ban, bangga, bangkang, saksi, kalah, kalam, kalang, kalimat, kam, kampung, karunia, kasih, kebaji, keledai, kemari, kena, kendati, keras, keridhaan, kertas, khabar, khawatir, khianat, ku, kuat, kulit, kumpul, kutuk, laam, laki, laknat.

Setelah melakukan tahapan *preprocessing* pada bagian *stopword* akan menghasilkan sebuah *bag of word* atau sekumpulan kata-kata seperti pada tabel 4.4 :

**Tabel 4.4** Bag Of Word Tema Beriman Kepada Kitab-Kitab

Bag Of Word
'ibrahim', 'anugerah', 'ummat', 'muhammad', 'injil', 'taurat', 'malaikat', 'perintah', 'musa', 'jibril', 'berkah', 'kiamat', 'isa', 'zabur'

## 4.3 Pembobotan Term

### 4.3.1 Term Frequency

Pada tahapan ini dilakukan untuk mencari nilai bobot BoW per dokumen. Pada penelitian ini penulis menggunakan jenis perhitungan *TF-Raw Count*, yaitu tahapan yang terlebih dahulu melakukan perhitungan kemunculan kata atau *bag of word* terhadap semua terjemahan Al-Qur'an tema beriman kepada kitab-kitab. Pada proses ini peneliti melakukan 2 tahapan yaitu, mengeset nilai setiap dokumen menjadi nol, dan menghitung kemunculan dari setiap dokumen. Beberapa tahap dapat dilihat pada tabel dibawah ini :

**Tabel 4.5** Mengeset nilai BoW menjadi nol

<i>Term</i> Terjemahan	shalat	anugerah	taurat	musa	mukjizat	isa	zabur
['shalat', 'anugerah']	0	0	0	0	0	0	0
['mukjizat', 'mukjizat', 'zabur']	0	0	0	0	0	0	0
['taurat', 'musa', 'mukjizat', 'isa']	0	0	0	0	0	0	0
['taurat']	0	0	0	0	0	0	0

Pada tabel 4.5 merupakan sebuah proses mengeset nilai *bag of word* menjadi 0. Dimana tahapan tersebut bertujuan untuk memudahkan dalam proses menghitung kemunculan *Bag Of Word*.

**Tabel 4.6** Hasil Menghitung Jumlah Kemunculan *Bag Of Word*

<i>Term</i> Terjemahan	shalat	anugerah	taurat	musa	mukjizat	isa	zabur
['shalat', 'anugerah']	1	1	0	0	0	0	0
['mukjizat', 'mukjizat', 'zabur']	0	0	0	0	2	0	1
['taurat', 'musa', 'mukjizat', 'isa']	0	0	1	1	1	1	0
['taurat']	0	0	1	0	0	0	0

Pada tabel 4.6 merupakan contoh dari proses untuk menghitung kemunculan BoW pada dokumen menggunakan *TF-Raw Count*.

### 4.3.2 *Invers Document Frequency- Smooth*

*Invers document frequency (IDF)* merupakan perhitungan untuk menghitung kemunculan kata atau *bag of word* pada keseluruhan terjemahan Al-Qur'an Cordoba tema beriman kepada kitab-kitab. Untuk hasil perhitungan IDF dapat dilihat pada tabel 4.7 :

**Tabel 4.7** Hasil Tampilan Perhitungan IDF-Smooth

anugerah	isa	taurat	mukjizat	shalat	musa	zabur
0,778	0,778	0.845	0.845	0,778	0,778	0,778

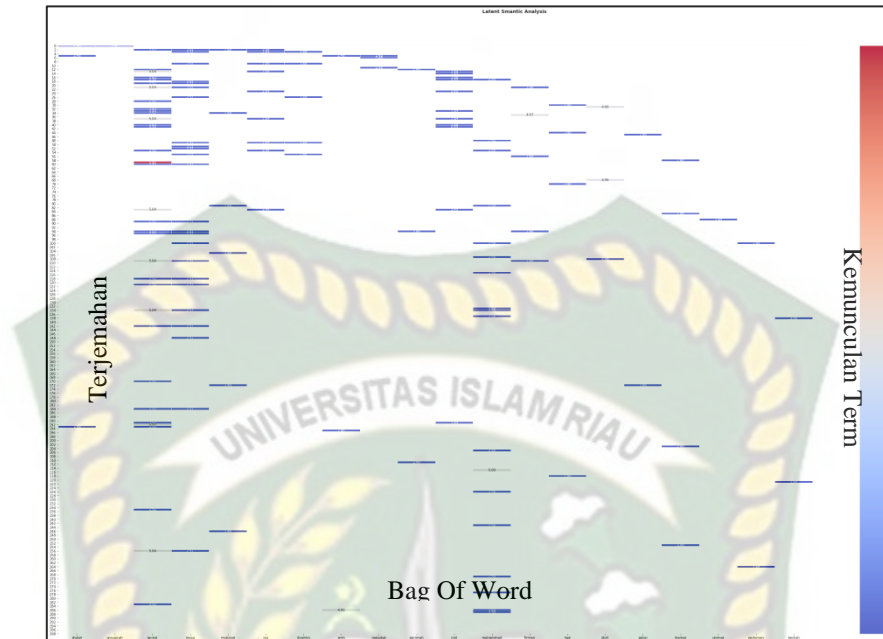
### 4.3.3 TF – IDF Smooth

TF-IDF-*Smooth* merupakan proses perhitungan antara perkalian TF-IDF untuk menghasilkan seberapa tingkat kepentingan kata (*bag of word*) terhadap setiap terjemahan Al-Qur'an *Cordoba*. Hasil perhitungan TF-IDF dapat dilihat pada tabel 4.8 dibawah ini :

**Tabel 4.8** Hasil Tampilan Perhitungan TF-IDF

	shalat	anugerah	taurat	musa	mukjizat	isa	zabur
0	0.778	0.778	0.000	0.000	0.000	0.000	0.000
1	0.000	0.000	0.000	0.000	1.690	0.000	0.778
2	0.000	0.000	0.845	0.778	0.845	0.778	0.000
3	0.000	0.000	0.845	0.000	0.000	0.000	0.000

Pada tabel 4.8 merupakan hasil dari perkalian TF-IDF yang kemudian hasilnya akan diproses kembali menggunakan *singular value decomposition* (SVD). Dalam pengelompokan ini dinilai bukan hanya seberapa banyak kemunculan *term* yang muncul dalam terjemahan itu sendiri, tetapi seberapa sering *term* itu muncul untuk keseluruhan terjemahan Al-Qur'an *Cordoba*. *Visualisasi* kemunculan *bag of word* pada TF-IDF dapat dilihat pada gambar 4.1 dibawah ini :

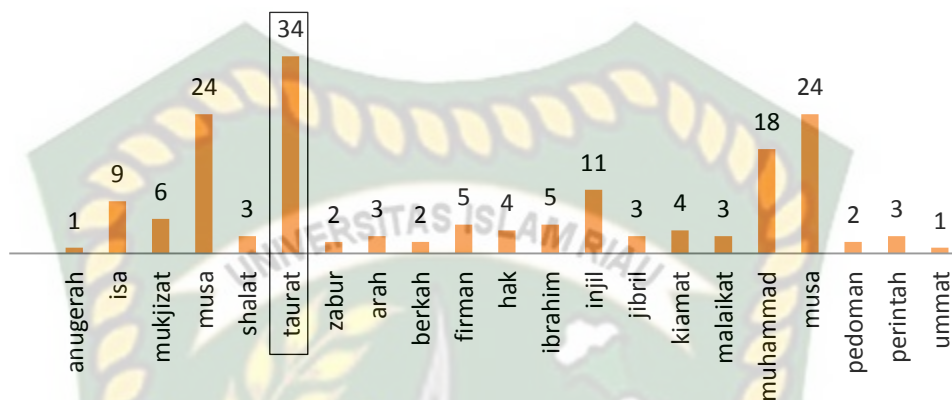


**Gambar 4.1** Visualisasi dari TF-IDF-*Smooth*

Pada gambar 4.1 terdapat kemunculan *term* yang ditandai dengan warna biru sampai warna merah. Semakin sering kemunculan *term* terhadap dokumen terjemahan tersebut akan ditandai dengan warna merah dan sebaliknya. Untuk memperjelas kemunculan *bag of word* terhadap keseluruhan terjemahan maka dapat dilihat pada gambar 4.2 dibawah ini :



### Presentasi Kemunculan Kata Terhadap Seluruh Terjemahan



**Gambar 4.2** Grafik Presentasi Surah & Ayat Berdasarkan *Bag Of Word*

Dari gambar 4.2 sudah menjelaskan secara keseluruhan *bag of word* yang saling berkaitan yang kemudian dihitung menggunakan TF-IDF, dari grafik tergambar jelas bahwa “**taurat**” yang lebih dominan dengan persentase 34 kali kemunculan dari seluruh terjemahan.

#### 4.4 *Singular Value Decomposition (SVD)*

Pada tahapan ini akan dilakukan sebuah pengelompokan surah dan ayat berdasarkan hasil dari perkalian TF-IDF. Sebelum dilakukannya proses *SVD*, terlebih dahulu ubah nilai dari pembobotan menjadi nilai matriks agar bias dibaca oleh *SVD*. Tahapan untuk melakukan *SVD* adalah membuat hasil perkalian dari TF-IDF menjadi bentuk matrik, lalu melakukan proses pengelompokan  $U$  dengan menampilkan ayat, dan surat dalam bentuk konsep. Pada pengelompokan *SVD* dilakukan proses menggunakan *library* yang sudah disediakan oleh *python*, sehingga untuk

pengelompokan dapat dilakukan dengan otomatis. Hasil dari proses SVD dapat dilihat pada gambar 4.3 dibawah ini :

(0, 0)	2.4800069429571505
(0, 1)	2.4771212547196626
(2, 2)	2.5224442335063197
(2, 3)	2.509202522331103
(2, 4)	2.484299839346786
(2, 5)	2.4885507165004443
(3, 3)	2.509202522331103
(3, 5)	2.4885507165004443
(3, 6)	2.482873583608754
(5, 0)	2.4800069429571505
(5, 7)	2.4800069429571505
(5, 8)	2.4800069429571505
(6, 8)	2.4800069429571505
(9, 3)	2.509202522331103
(9, 5)	2.4885507165004443
(9, 6)	2.482873583608754
(11, 8)	2.4800069429571505
(12, 2)	2.5224442335063197
(13, 2)	5.0448884670126395
(13, 5)	2.4885507165004443
(13, 10)	2.4913616938342726
(14, 10)	2.4913616938342726
(16, 2)	2.5224442335063197
(16, 10)	2.4913616938342726
: :	

**Gambar 4.3** Matrik perkalian TF-IDF

Setelah mendapatkan hasil perkalian dari TF-IDF, tahapan selanjutnya melakukan pengelompokan matrik  $U$  dengan cara mengelompokan ayat dan surat yang saling terkait dengan *bag of word* yang menjadi konsep. Hasil dari pengelompokan  $U$  dapat dilihat pada gambar 4.4 dibawah ini :

[ [ 1.37852910e-03 -1.42185518e-03 3.48198489e-03 -8.54011666e-03
3.93725582e-16 -1.21517644e-02]
[-5.13931361e-34 8.17008526e-19 1.11019296e-16 -6.24500451e-17
-6.46201747e-28 -6.93692488e-18]
[ 1.60174663e-01 7.40602372e-02 -1.26912777e-01 1.58774361e-01
4.90778884e-16 -8.93062342e-02]
...
[-0.00000000e+00 0.00000000e+00 0.00000000e+00 -0.00000000e+00
0.00000000e+00 0.00000000e+00]
[-0.00000000e+00 0.00000000e+00 0.00000000e+00 -0.00000000e+00
0.00000000e+00 0.00000000e+00]
[-0.00000000e+00 0.00000000e+00 0.00000000e+00 -0.00000000e+00
0.00000000e+00 0.00000000e+00]]]

**Gambar 4.4** Proses Pengelompokan U

Setelah melakukan *preprocessing*, pembobotan *term* pada *singular value decomposition* (SVD) maka langkah selanjutnya adalah setiap kata *bag of word* akan dikelompokkan. Pengelompokan kata perkonsep ini menghasilkan sebanyak 6 konsep berdasarkan dari banyak subtema beriman kepada kitab-kitab yang di ambil dari Al-Qur'an *cordoba* yang telah ditetapkan pada saat proses SVD. Berikut adalah hasil pengelompokan *bag of word* yang terdapat pada tabel 4.9 :

**Tabel 4.9** Hasil Pengelompokan Surah & Ayat Perkonsep

Konsep	Surat & Ayat	Jumlah
0	(3, 93), (32, 23), (17, 2), (41, 45), (5, 46), (7, 169), (2, 87), (7, 145), (40, 53), (6, 91), (6, 154), (7, 154), (23, 49), (21, 48), (5, 68), (61, 6), (7, 157), (3, 3), (9, 111), (5, 66), (3, 50), (5, 44), (7, 170), (28, 49), (2, 101), (3, 23), (2, 41), (62, 5), (2, 136), (3, 84), (42, 13), (11, 17), (87, 19), (46, 12), (28, 43), (46, 30), (5, 81), (16, 103), (57, 27), (4, 174), (3, 81), (5, 47), (68, 44), (32, 3), (10, 38), (20, 99), (52, 33), (46, 8), (16, 89), (37, 37), (23, 70), (68, 51), (36, 70), (2, 75), (3, 184), (35, 25), (16, 44), (13, 36), (2, 177), (2, 3), (45, 20), (7, 182), (21, 105), (4, 136), (2, 285), (2, 97), (16, 102), (2, 4), (2, 176), (6, 90)	70 Surah & Ayat

Konsep	Surat & Ayat	Jumlah
1	(16, 103), (68, 44), (11, 17), (4, 174), (68, 51), (52, 33), (16, 89), (36, 70), (10, 38), (37, 37), (46, 8), (32, 3), (23, 70), (20, 99), (61, 6), (5, 68), (32, 23), (3, 84), (42, 13), (2, 136), (2, 177), (7, 182), (87, 19), (28, 43), (5, 81), (46, 30), (46, 12), (3, 184), (45, 20), (16, 44), (35, 25), (57, 27), (2, 87), (4, 136), (2, 285), (21, 105), (38, 29), (21, 50), (2, 91), (2, 185), (18, 56), (21, 24), (3, 119)	43 Surah & Ayat
2	(2, 136), (42, 13), (3, 84), (87, 19), (28, 43), (5, 81), (46, 30), (46, 12), (2, 87), (11, 17), (7, 145), (40, 53), (23, 49), (6, 91), (7, 154), (21, 48), (6, 154), (17, 2), (41, 45), (3, 81), (2, 75), (45, 20), (32, 23), (57, 27), (3, 184), (35, 25), (16, 44), (13, 36), (21, 105), (2, 4), (2, 97), (16, 102), (3, 119), (6, 90)	34 Surah & Ayat
3	(57, 27), (5, 46), (2, 136), (3, 84), (42, 13), (5, 47), (9, 111), (3, 3), (7, 157), (5, 66), (2, 87), (5, 68), (61, 6), (87, 19), (3, 184), (16, 44), (35, 25), (4, 174), (21, 105), (3, 50), (2, 4), (41, 40), (10, 15), (2, 174), (6, 51), (18, 56), (21, 24), (2, 185), (2, 91), (2, 176), (3, 53), (3, 7), (3, 119), (6, 90)	34 Surah & Ayat
4	(3, 81), (7, 145), (2, 75), (17, 2), (57, 27), (5, 47), (5, 68), (5, 46), (3, 3), (5, 66), (9, 111), (7, 157), (13, 36), (68, 44), (3, 84), (42, 13), (2, 136), (5, 44), (7, 182), (2, 177), (16, 103), (61, 6), (37, 37), (23, 70), (46, 8), (36, 70), (10, 38), (20, 99), (16, 89), (52, 33), (32, 3), (68, 51), (2, 285), (4, 136), (2, 97), (16, 102), (21, 50), (38, 29), (3, 119), (6, 90)	40 Surah & Ayat
5	(68, 44), (2, 177), (7, 182), (2, 3), (7, 170), (2, 285), (4, 136), (2, 136), (42, 13), (3, 84), (87, 19), (5, 46), (41, 45), (40, 53), (7, 154), (23, 49), (6, 154), (21, 48), (6, 91), (3, 93), (57, 27), (7, 157), (9, 111), (5, 66), (3, 3), (46, 30), (5, 81), (28, 43), (46, 12), (7, 169), (5, 47), (62, 5), (2, 101), (28, 49), (2, 41), (3, 23), (2, 4), (2, 176), (38, 29), (21, 50), (3, 53), (3, 119), (6, 90)	43 Surah & Ayat

Pada tabel 4.9 sudah terlihat keseluruhan hasil pengelompokan dari matrik *U* yang menghasilkan 6 konsep. Hasil dari pengelompokan inilah yang akan dievaluasi pada tahap berikutnya untuk mencari kecocokan antara subtema yang ada didalam Al-Qur'an *Amazing Cordoba*.



## 4.5 Evaluasi

Pada tahapan ini akan dilakukan evaluasi sebanyak dua kali yaitu dengan menggunakan sistem dan ahli pakar *tematik Qur'an*. Pakar tematik itu sendiri akan mengevaluasi dengan cara mengisi kuisisioner yang sudah disediakan oleh peneliti untuk mencari kecocokan apakah surah dan ayat yang sudah dievaluasi oleh sistem sudah sesuai dengan tafsir *qur'an*, atau ada sebuah kontribusi diluar dari *tafsir qur'an* itu sendiri. Sebelum dilakukannya dua tahapan evaluasi tersebut. Ada beberapa langkah yang harus dilalui untuk mendapatkan hasil akhir. Berikut adalah sebagian contoh data dari hasil penggabungan surah, ayat, *bag of word* dan TF-IDF yang dapat dilihat pada tabel 4.10 dibawah ini :

**Tabel 4.10** Hasil Penggabungan Surah, Ayat, BoW, dan TF-IDF

Index	Bag Of Word	Surah & Ayat	Tf-Idf
0	Shalat	( 2, 3)	2.480.007
1	Anugerah	( 2, 3)	2.477.121
2	Taurat	(2, 87)	2.522.444
3	Musa	(2, 87)	2.509.203
4	Mukjizat	(2, 87)	2.484.300
...	...	....	....
121	Arah	(68, 44)	4.900498
122	Muhammad	(68, 44)	2.471292
123	Muhammad	(68, 51)	2.471292

### 4.5.1 Data Inputan *Gold Standard*

Data inputan *gold standard* merupakan data keseluruhan yang digunakan untuk mengevaluasi data akhir *Latent Semantic Analysis* yang terdiri dari 6 subtema dan

300 data berisikan, surah, ayat, dan terjemahan yang sesuai dengan tema beriman kepada kitab-kitab berdasarkan Al-Qur'an Cordoba. Beberapa contoh data evaluasi pada penelitian ini dapat dilihat pada tabel 4.11 :

**Tabel 4.11** Inputan Data *Gold Standard*

No	Sub Tema	Surah dan Ayat	Jumlah
1	Kewajiban dan Keutamaan Beriman pada Kitab-Kitab	(2, 3), (2, 4), (2, 87), (2, 136), (2, 176), (2, 177), (2, 285), (3, 7), (3, 53), (3, 84), (3, 119), (4, 136), (5, 44), (5, 46), (5, 47), (5, 59), (5, 66), (5, 68), (5, 81), (6, 154), (6, 156), (17, 2), (29, 46), (57, 27), (64, 8), (87, 18), (87, 19)	27 Surah & Ayat
2	Pengakuan Antara Satu Kitab dengan lainnya.	(2, 4), (2, 41), (2, 89), (2, 91), (2, 97), (2, 101), (3, 3), (3, 50), (3, 81), (4, 47), (5, 46), (5, 48), (6, 92), (7, 157), (9, 111), (10, 37), (12, 111), (21, 24), (21, 105), (27, 76), (35, 31), (37, 37), (42, 13), (46, 10), (46, 12), (46, 30), (61, 6), (87, 18), (87, 19)	29 Surah & Ayat
3	Adanya Perubahan dalam Beberapa Kitab Samawi	(2, 75), (2, 79), (2, 174), (3, 93), (6, 91), (30, 32)	6 Surah & Ayat
4	Al-Quran Terpelihara dari Penyelewengan	(15, 9), (18, 27), (22, 52), (75, 19), (85, 22)	5 Surah & Ayat
5	Hikmah Penurunan Kitab-Kitab Samawi	(2, 2), (2, 97), (2, 129), (2, 185), (2, 213), (2, 231), (2, 242), (3, 4), (3, 101), (3, 103), (3, 108), (3, 138), (3, 164), (4, 105), (4, 174), (5, 16), (5, 46), (6, 19), (6, 51), (6, 55), (6, 70), (6, 90), (6, 154), (6, 157), (7, 2), (7, 35), (7, 52), (7, 145), (7, 154), (7, 174), (7, 203), (9, 33), (10, 57), (11, 17), (12, 2), (12, 104), (12, 111), (14, 1), (16, 44), (16, 64), (16, 89), (16, 102), (17, 2), (17, 41), (18, 2), (19, 97), (20, 2), (20, 3), (20, 99), (20, 113), (21, 10), (21, 48), (21, 106), (22, 16), (23, 49), (24, 1), (24, 34), (25, 1), (25, 33), (25, 50), (26, 209), (27, 2), (28, 2), (28, 43), (28, 86), (29, 51), (31, 3), (32, 3), (32, 23), (34, 50), (36, 6), (36, 70), (38, 29), (38, 87), (39, 28), (39, 71), (40, 53), (40, 54), (43, 3), (43, 44), (44, 58), (45, 11), (46, 12), (50, 45), (54, 17), (54, 22), (54, 32), (54, 40), (65, 10), (65, 11), (68, 52), (69, 48), (73, 19), (74, 54), (76, 29), (80, 11), (80, 12), (81, 27)	98 Surah & Ayat
6	Sikap Manusia Terhadap Kitab Samawi	(2, 170), (2, 176), (2, 213), (3, 4), (3, 7), (3, 19), (3, 23), (3, 78), (3, 184), (4, 55), (4, 56), (4, 140), (5, 10), (5, 86), (6, 7), (6, 25), (6, 49), (6, 66), (6, 70), (6, 89), (6, 91), (6, 150), (6, 157), (7, 9), (7, 36), (7, 40), (7, 146), (7, 157), (7, 169), (7, 170), (7, 177), (7, 182), (8, 2), (8, 31), (9, 9), (9, 65), (9, 124), (9, 125), (9, 127), (10, 15), (10, 17), (10, 38), (10, 39), (10, 40), (10, 95), (13, 1), (13, 19), (13, 36), (15, 91), (16, 24), (16, 30), (16, 103), (16, 104), (16, 105), (18, 56), (18, 57), (19, 73), (21, 50), (21, 93), (22, 72), (23, 53), (23, 66), (23, 70), (23, 71), (23, 105), (25, 30), (26, 5), (26, 198), (26, 199), (26, 200), (26, 201), (28, 49), (28, 50), (28, 52), (30, 10), (30, 32), (31, 6), (31, 7), (31, 21), (32, 3), (34, 31), (34, 43), (35, 25), (35, 43), (37, 15), (39, 32), (39, 59), (40, 70), (41, 26), (41, 40), (41, 41), (41, 44), (41, 45), (41, 52), (43, 24), (43, 78), (44, 9), (45, 6), (45, 8), (45, 9), (45, 20), (45, 25), (45, 31), (45, 35), (46, 7), (46, 8), (46, 10), (46, 11), (47, 9), (50, 5), (50, 12), (50, 14), (51, 8), (52, 33), (53, 59), (53, 60), (53, 61), (56, 81), (57, 16), (62, 5), (64, 10), (68, 15), (68, 44), (68, 51), (69, 49), (74, 16), (74, 18), (74, 21), (74, 22), (74, 23), (74, 24), (74, 25), (74, 49), (78, 28), (83, 13)	135 Surah & Ayat

#### 4.5.2 Hasil Evaluasi Terhadap Al-Qur'an *Amazing Cordoba*

Pada tahapan ini evaluasi yang dilakukan terhadap *Al-Qur'an amazing cordoa*, yaitu melakukan pengecekan data hasil nilai LSA dengan data inputan evaluasi yang terdiri dari subtema, surah, dan ayat. Evaluasi dilakukan dengan cara menghitung nilai *precision*, *recall*, dan *f1-Score* yang akan menghasilkan beberapa evaluasi yang dapat dilihat pada tabel 4.12 dibawah ini :

**Tabel 4.12** Hasil Evaluasi Terhadap Al-Qur'an

Pengelompokan	No	Total Data	Subtema	TP	FP	FN	P	R	FS
<b>Konsep 0</b> (Jumlah Data: 70)	1	27	Kewajiban dan Keutamaan Beriman pada Kitab-Kitab.	21	59	6	26%	77%	39%
	2	29	Pengakuan antara satu kitab dengan lainnya	18	62	11	22%	62%	33%
	3	6	Adanya perubahan dalam beberapa kitab samawi	4	76	2	5%	66%	9%
	4	5	Al-Qur'an terpelihara dari penyelewengan.	0	80	5	0%	0%	0%
	5	97	Hikmah penurunan kitab-kitab samawai.	22	58	75	27%	22%	24%
	6	135	Sikap manusia terhadap sikap samawi	27	53	108	33%	20%	25%
<b>Konsep 1</b> (Jumlah data: 43)	1	27	Kewajiban dan Keutamaan Beriman pada Kitab-Kitab.	14	36	13	28%	51%	36%
	2	29	Pengakuan antara satu kitab dengan lainnya	9	41	20	18%	31%	22%
	3	6	Adanya perubahan dalam beberapa kitab samawi	2	48	4	4%	33%	7%
	4	5	Al-Qur'an terpelihara dari penyelewengan.	0	50	5	0%	0%	0%
	5	97	Hikmah penurunan kitab-kitab samawai.	17	33	80	34%	17%	23%
	6	135	Sikap manusia terhadap sikap samawi	16	34	119	32%	11%	17%
<b>Konsep 2</b> (Jumlah data: 34)	1	27	Kewajiban dan Keutamaan Beriman pada Kitab-Kitab.	12	34	15	26%	44%	32%
	2	29	Pengakuan antara satu kitab dengan lainnya	11	41	20	18%	37%	29%
	3	6	Adanya perubahan dalam beberapa kitab samawi	1	45	5	25%	16%	3%
	4	5	Al-Qur'an terpelihara dari penyelewengan.	0	46	5	0%	0%	0%
	5	97	Hikmah penurunan kitab-kitab samawai.	9	37	88	19%	9%	12%
	6	135	Sikap manusia terhadap sikap samawi	19	27	116	41%	14%	20%

Pengelompokan	No	Total Data	Subtema	TP	FP	FN	P	R	FS
Konsep 3 (Jumlah data: 34)	1	27	Kewajiban dan Keutamaan Beriman pada Kitab-Kitab.	14	22	13	38%	51%	44%
	2	29	Pengakuan antara satu kitab dengan lainnya	13	23	16	36%	44%	40%
	3	6	Adanya perubahan dalam beberapa kitab samawi	1	35	5	2%	16%	4%
	4	5	Al-Qur'an terpelihara dari penyelewengan.	0	36	5	0%	0%	0%
	5	97	Hikmah penurunan kitab-kitab samawai.	8	28	89	22%	8%	12%
	6	135	Sikap manusia terhadap sikap samawi	8	28	127	22%	5%	9%
Konsep 4 (Jumlah data: 40)	1	27	Kewajiban dan Keutamaan Beriman pada Kitab-Kitab.	11	43	16	20%	40%	27%
	2	29	Pengakuan antara satu kitab dengan lainnya	9	45	20	16%	31%	21%
	3	6	Adanya perubahan dalam beberapa kitab samawi	2	52	4	3%	55%	6%
	4	5	Al-Qur'an terpelihara dari penyelewengan.	0	54	5	0%	0%	0%
	5	97	Hikmah penurunan kitab-kitab samawai.	17	37	80	31%	17%	22%
	6	135	Sikap manusia terhadap sikap samawi	21	33	114	38%	15%	22%
Konsep 5 (Jumlah data: 43)	1	27	Kewajiban dan Keutamaan Beriman pada Kitab-Kitab.	0	33	17	23%	37%	28%
	2	29	Pengakuan antara satu kitab dengan lainnya	11	32	18	25%	37%	30%
	3	6	Adanya perubahan dalam beberapa kitab samawi	2	41	4	4%	33%	8%
	4	5	Al-Qur'an terpelihara dari penyelewengan.	0	43	5	0%	0%	0%
	5	97	Hikmah penurunan kitab-kitab samawai.	12	31	85	27%	12%	17%
	6	135	Sikap manusia terhadap sikap samawi	14	29	121	32%	10%	15%

Keterangan :

*TP* : True Positive

*P* : Precision

*FP* : False Positive

*R* : Recall

*FN* : False Negative

*FS* : F1-Score

Pada tabel 4.12 sudah terlihat beberapa hasil perhitungan evaluasi terhadap Al-Qur'an *Amazing Cordoba*. Langkah selanjutnya adalah melakukan analisis dari keseluruhan konsep kemudian mengambil nilai *f1-score* terbaik. Hasil evaluasi Al-Qur'an *amazing cordoba* terbaik dapat dilihat pada tabel 4.13 dibawah ini :



**Tabel 4.13** Hasil Evaluasi Terbaik Terhadap Al-Qur'an

Konsep	TP	FP	FN	Precision (%)	Recall (%)	F1-Score (%)
Konsep 3	14	22	13	38.88	51.85	44.44%

Keterangan :

*TP* : True Potitive

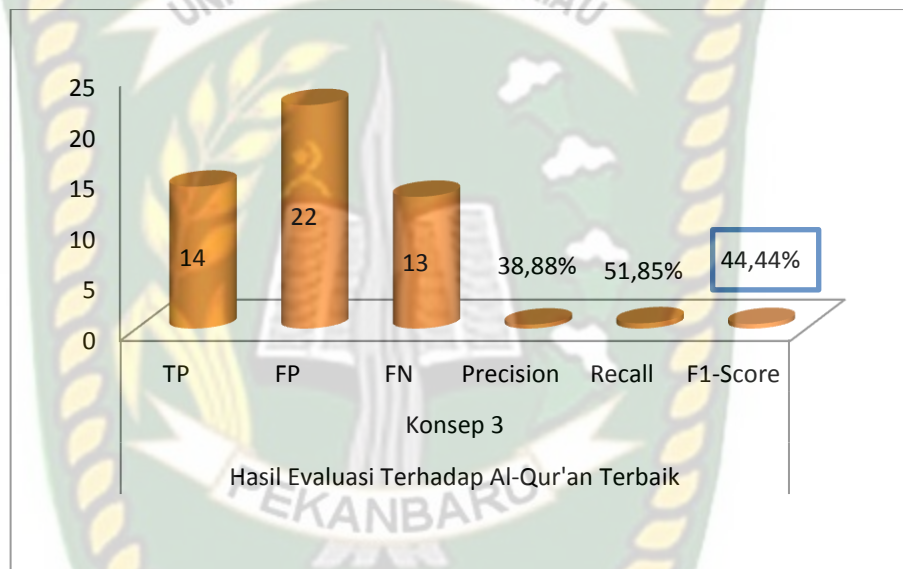
*FP* : False Positive

*FN* : False Ngative

*P* : Precission

*R* :Recall

*FS* : F1-Score

**Gambar 4.5** Grafik Hasil Evaluasi Al-Qur'an Terbaik

Pada tabel 4.13 dan gambar 4.5 diatas menunjukkan hasil akhir dari perhitungan *f1-score* terbaik, Adapun hasil *f1-score* terbaik terdapat pada konsep 3, sub tema **“Kewajiban dan Keutamaan Beriman pada Kitab-Kitab”** dengan nilai *f1-score* **44,44%**.

### 4.5.3 Hasil Evaluasi Data *False Positive* Al-Qur'an

Evaluasi terhadap Al-Qur'an menghasilkan sebuah data tidak tepat atau *false positive*. Data tidak tepat tersebut akan dievaluasi kembali oleh ahli pakar, yang bertujuan untuk mengoreksi kebenaran, apakah surah dan ayat tersebut digolongkan kepada subtema “Kewajiban dan Keutamaan Beriman pada Kitab-Kitab” atau tidak. Apabila menurut ahli pakar data tersebut cocok dengan subtema yang menghasilkan data tidak cocok tersebut maka akan terjadi penambahan data kembali pada subtema hingga menghasilkan sebuah kontribusi atau peningkatan akurasi data pada subtema tersebut. Pada tema beriman kepada kitab-kitab menghasilkan data *false positive* berjumlah 22 surah dan ayat yang dapat dilihat pada tabel 4.13. Berikut adalah sebagian contoh data *false positive* dapat dilihat pada tabel 4.14 dibawah ini :

**Tabel 4.14** Data Tidak Tepat (*False Positive*)

SURAH & AYAT YANG TIDAK TEPAT				TOTAL
(42, 13)	(61, 6)	(9, 111)	(7, 157)	22 Surah & Ayat
(3, 3)	(3, 184)	(35, 25)	(16, 44)	
(46, 12)	(46, 30)	(21,105)	(2, 97)	
(16, 102)	(10, 15)	(2, 174)	(6, 51)	
(41, 40)	(2, 185)	(2, 91)	(18, 56)	
(21, 24)	(6, 90)			

### 4.5.4 Hasil Evaluasi Terhadap Pakar

Data hasil evaluasi pakar didapatkan dari hasil evaluasi terhadap data Al-Qur'an (*false positive*) berjumlah 22 surah dan ayat tidak tepat yang dapat dilihat pada tabel 4.13. Data yang dihasilkan tersebut akan dievaluasi kembali oleh pakar. Disini peneliti menggunakan dua orang pakar yang akan mengevaluasi kembali data

hasil evaluasi Al-Qur'an yaitu, Bapak Dr. Hamzah, M.Ag selaku Dosen dan Wakil Dekan II di Prodi Pendidikan Agama Islam (Universitas Islam Riau) dan Bapak Musaddad Harahap, M.Pdi yang merupakan Dosen serta Sekretaris Prodi Agama Islam di (Universitas Islam Riau). Adapun hasil evaluasi dari kedua pakar mengenai data data tidak tepat tersebut dapat dilihat pada tabel 4.15 dibawah ini :

**Tabel 4.15** Hasil Evaluasi Data Terhadap Pakar

NO	Surah & Ayat	Respon Pakar 1		Respon Pakar 2		Jumlah Penambahan (Surah & Ayat)
		(Dr.Hamzah, M.Ag)		(Musaddad Harahap, M.Pdi)		
		IYA	TIDAK	IYA	TIDAK	
1	(42, 13)		√		√	<b>5 Kontribusi</b>
2	(61, 6)	√		√		
3	(9, 111)		√		√	
4	(7, 157)	√			√	
5	(3, 3)	√		√		
6	(3, 184)		√		√	
7	(35, 25)		√		√	
8	(16, 44)	√		√		
9	(46, 12)		√		√	
10	(46, 30)		√		√	
11	(21, 105)		√		√	
12	(2, 97)		√		√	
13	(16, 102)		√		√	
14	(10, 15)		√		√	
15	(2, 174)		√		√	
16	(6, 51)		√		√	
17	(41, 40)		√		√	
18	(2, 185)	√			√	
19	(2, 91)	√		√		
20	(18, 56)		√		√	
21	(21, 24)		√		√	
22	(6, 90)	√		√		

Pada tabel diatas didapatkan kontribusi sebanyak 5 surah dan ayat hasil dari evaluasi pakar yaitu surah 6 ayat 61, surah 3 ayat 3, surah 16 ayat 44, surah 2 ayat 91 dan surah 6 ayat 90 kedalam subtema kewajiban dan keutamaan beriman pada kitab-kitab. Muncul sebuah pertanyaan, kenapa hanya menghasilkan 5 kontribusi dari 22 data yang telah dievaluasi oleh pakar ? Menurut penjelasan dari kedua pakar, dalam pengelompokan surah dan ayat dalam Al-Qur'an tidak bisa dilihat hanya dari terjemahannya saja, dikarenakan banyaknya bahasa istilah dan perumpamaan kata. Jadi diperlukan beberapa pendukung lain seperti hadist, tafsir dan asbabunnuzul ayat itu sendiri. Adapun data yang sudah dievaluasi pakar kemudian akan ditambahkan pada data inputan evaluasi sehingga akan menghasilkan perubahan pada data hasil evaluasi terhadap Al-Qur'an. Perubahan data dapat dilihat pada tabel 4.16 dibawah ini :

Tabel 4.16 Hasil Evaluasi Terhadap Al-Qur'an dan Pakar

Konsep	Pengelompokkan	F1-Score		Selisih (%)	Peningkatan (%)
		Cordoba (%)	Pakar (%)		
konsep 3	Kewajiban dan Keutamaan Beriman pada Kitab-Kitab	44.46	55.88	11.42	25.68
<b>Rata-rata Peningkatan (%)</b>					<b>25.68</b>

Pada tabel diatas terjadi peningkatan *f1-score* setelah dilakukannya evaluasi oleh pakar. Dikareakan adanya kontribusi surat dan ayat sehingga akan adanya perubahan data pada evaluasi terhadap Al-Qur'an. Adapun cara untuk menghitung nilai selisih dengan cara nilai *cordoba* dikurang nilai *pakar*, sedangkan nilai *peningkatan* didapatkan dari nilai *selisih* dibagi nilai *cordoba* dikalikan 100.



## BAB V

### KESIMPULAN DAN SARAN

#### 5.1 Kesimpulan

Berdasarkan hasil dari penelitian yang telah dilakukan diatas, dapat ditarik beberapa kesimpulan sebagai berikut :

1. Metode yang digunakan pada pengelompokan tematik bertemakan beriman kepada kitab-kitab ini menghasilkan 6 subtema yang sudah terkelompokan dengan menggunakan *bag of word* berdasarkan *tafsir* dan *hadist*.
2. Terciptanya beberapa kamus data atau *bag of word* yang saling berkaitan satu dan yang lainnya dengan tema beriman kepada kitab-kitab agar nantinya dapat dipergunakan sebagai bahan penelitian selanjutnya.
3. Adanya penambahan *kontribusi* sebanyak 5 surah dan ayat pada subtema “Kewajiban dan Keutamaan Beriman pada Kitab-Kitab”.
4. Pada implementasi pengelompokan tematik bertemakan beriman kepada kitab-kitab hanya muncul satu subtema yaitu “Kewajiban dan Keutamaan Beriman pada Kitab-Kitab” dengan metode pengujian *Precesion*, *Recall* an *F1-Score* menghasilkan 44-47%.

## 5.2 Saran

Pada penelitian ini masih jauh dari kata sempurna sehingga penulis menyarankan bagi peneliti selanjutnya agar dapat mengembangkan penelitian ini, diantaranya sebagai berikut :

1. Menggunakan metode selain *latent semantic analysis* (LSA) untuk menemukan pengelompokan yang jauh lebih baik lagi.
2. Pada penelitian selanjutnya agar dapat menganalisis keseluruhan tema - tema yang terdapat didalam Al-Qur'an sehingga pengelompokan tidak hanya dilakukan pada satu tema saja.
3. Melakukan pengelompokan tidak hanya berdasarkan *bag of word* saja, melainkan juga berdasarkan beberapa *tafsir* dan *hadist*.

## DAFTAR PUSTAKA

- Amalia, A., Lydia, M. S., Fadilla, S. D., & Huda, M. (2018). *Perbandingan Metode Klaster dan Preprocessing Untuk Dokumen Berbahasa Indonesia. Jurnal Rekayasa Elektroika*, Vol 14(1), 35–42.
- Ghina, C. F., Purnamasari, R., Zakiawati, D., Elektro, F. T., Telkom, U., Gigi, F. K., & Bandung, U. P. (2019). "Identifikasi pola sidik bibir pada identitas manusia menggunakan metode singular value decomposition (SVD) dan Klasifikasi learning vector quantization (LVQ) untuk aplikasi ibidang biometrik forensic identification of lip print pattern on human identity using singular value decomposition(SVD)metod and learning vector quantization LVQ) classification for forensic biometric" Vol (2), 3974–3979.
- Harison. (2016). *Model pengembangan sistem diperlukan sebagai metode, best practices, deliverables, dan alatotomasi yang digunakan stakeholder untuk mengembangkan dan meningkatkan sistem informasi dan perangkat lunak. Proses perangkat lunak terdapat beberapa tahap*, Vol 4 (2), 76–81.
- Harismawan, A. F., Kharisma, A. P., & Afirianto, T. (2018). Analisis Perbandingan Performa Web Service Menggunakan Bahasa Pemrograman Python, PHP , dan Perl pada Client Berbasis Android. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer (J-PTIHK) Universitas Brawijaya*, Vol 2, 237–245.
- Ismail, E. (2018). *Analisis Semantik Pada Kata Ahzāb Dan Derivasinya Dalam Al-Quran*. Al-Bayan: Jurnal Studi Ilmu Al- Qur'an Dan Tafsir, Vol(1)

- Intan, R., & Defeng, A.(2006). Hard : Subject-Based Search Engine Menggunakan Tf-Idf Dan Jaccard ' S Coefficient. *Jurnal Teknik Industri*, Vol 8 (1), 61–72.
- Luh, N., Sri, W., Ginantra, R., & Wardani, N.W.(2019). *Implementasi Metoda Naïve Bayes dan Vector Space Model dalam Deteksi Kesamaan artikel Jurnal Berbahaya*. Vol 4(2).
- Luthfiarta, A., Zeniarja, J., & Salam, A. (2013). *Algoritma Latent Semantic Analysis (LSA) Pada Peringkasan Dokumen Otomatis Untuk Proses Clustering Dokumen*. Seminar Nasional Teknologi Informasi & Komunikasi Terapan 2013 (SEMANTIK 2013), 2013(November), 13–18.
- Metisen, B. M., & Sari, H. L. (2015). *Analisis clustering menggunakan metode K-Means dalam pengelompokan penjualan produk pada Swalayan Fadhila*. *Jurnal Media Infotama*, Vol 11(2), 110–118.
- Mardiana, T., & Nyoto, R. D. (2015). Kluster Bag of Word Menggunakan Weka. *Jurnal Edukasi Dan Penelitian Informatika (JEPIN)*, Vol 1(1), 1–5.
- Pembelajaran, M., Agama, P., Di, I., & Inklusi, S. (n.d.). (2016) “*Model Pembelaaran Pedidika Agama Islam di Sekolah Inklusi*” Universitas Pendidikan Indonesia
- Polettine, N. (2004). *The Vector Space Model in Information Retrieval - Term Weighting Problem*, 1-4.
- Rahmat, M. (2015).”Implementasi Metode Tematik Al-Quran untuk Memahami Makna Beriman kepada Malaikat-malaikatNya Allah”. *Taklim – Jurnal Pendidikan Agama Islam*, Vol13(No.1 Maret 2015), 79–92.



Santoso, V. I., Virginia, G., & Lukito, Y. (2017). "Penerapan Sentiment Analysis Pada Hasil Evaluasi Dosen Dengan Metode Support Vector Machine". Jurnal Transformatika, Vol 14(2), 72.

Setiaji, A. C. dan B. (2014). "Al-quran dan al-sunnah sebagai sumber ajaran islam". Vol 26(2), 89–110.

Wicaksono, D. W., & Irawan, M. I. (2014). Sistem Deteksi Kemiripan Antar Dokumen Teks Menggunakan Model Bayesian Pada Term Latent Semantic Analysis (LSA). Jurnal Sains Dan Seni POMITS, Vol 3(2), 1–6.

