

**KLASIFIKASI MULTI-LABEL PADA TERJEMAHAN AYAT AL-  
QUR'AN MENGGUNAKAN ALGORITMA MULTINOMIAL NAIVE  
BAYES**

**SKRIPSI**



disusun oleh

**Muhammad Zulfakar**

**16.11.0500**

**PROGRAM SARJANA  
PROGRAM STUDI INFORMATIKA  
FAKULTAS ILMU KOMPUTER  
UNIVERSITAS AMIKOM YOGYAKARTA  
YOGYAKARTA**

**2020**

**KLASIFIKASI MULTI-LABEL PADA TERJEMAHAN AYAT  
AL-QUR'AN MENGGUNAKAN ALGORITMA MULTINOMIAL NAIVE  
BAYES**

**SKRIPSI**

untuk memenuhi sebagian persyaratan  
mencapai gelar Sarjana  
pada Program Studi Informatika



disusun oleh

**Muhammad Zulfakar**

**16.11.0500**

**PROGRAM SARJANA  
PROGRAM STUDI INFORMATIKA  
FAKULTAS ILMU KOMPUTER  
UNIVERSITAS AMIKOM YOGYAKARTA  
YOGYAKARTA**

**2020**

**PERSETUJUAN**

**SKRIPSI**

**KLASIFIKASI MULTI-LABEL PADA TERJEMAHAN AYAT  
AL-QUR'AN MENGGUNAKAN ALGORITMA MULTINOMIAL NAIVE  
BAYES**

yang dipersiapkan dan disusun oleh

**Muhammad Zulfakar**

**16.11.0500**

telah disetujui oleh Dosen Pembimbing Skripsi  
pada tanggal 5 Desember 2019

**Dosen Pembimbing,**

**Sumarni Adi, S.Kom,M.Cs**

**NIK. 190302256**

**PENGESAHAN**  
**SKRIPSI**  
**KLASIFIKASI MULTI-LABEL PADA TERJEMAHAN AYAT**  
**AL-QUR'AN MENGGUNAKAN ALGORITMA MULTINOMIAL NAIVE**  
**BAYES**

yang dipersiapkan dan disusun oleh

**Muhammad Zulfakar**

**16.11.0500**

telah dipertahankan di depan Dewan Penguji  
pada tanggal 16 Juli 2020

**Susunan Dewan Penguji**

**Nama Penguji**

**Tanda Tangan**

**Supriatin, M.Kom**  
**NIK. 190302239**

\_\_\_\_\_

**Dina Maulina, M.Kom**  
**NIK. 190302250**

\_\_\_\_\_

**Sumarni Adi, S.Kom, M.Cs**  
**NIK. 190302256**

\_\_\_\_\_

Skripsi ini telah diterima sebagai salah satu persyaratan  
untuk memperoleh gelar Sarjana Komputer  
Tanggal 23 Juli 2020

**DEKAN FAKULTAS ILMU KOMPUTER**

**Krisnawati, S.Si., M.T.**  
**NIK. 190302038**

## PERNYATAAN

Saya yang bertandatangan dibawah ini menyatakan bahwa, skripsi ini merupakan karya saya sendiri (ASLI), dan isi dalam skripsi ini tidak terdapat karya yang pernah diajukan oleh orang lain untuk memperoleh gelar akademis di suatu institusi pendidikan tinggi manapun, dan sepanjang pengetahuan saya juga tidak terdapat karya atau pendapat yang pernah ditulis dan/atau diterbitkan oleh orang lain, kecuali yang secara tertulis diacu dalam naskah ini dan disebutkan dalam daftar pustaka.

Segala sesuatu yang terkait dengan naskah dan karya yang telah dibuat adalah menjadi tanggung jawab saya pribadi.

Yogyakarta, 11 September 2020



**Muhammad Zulfakar**

**16.11.0500**

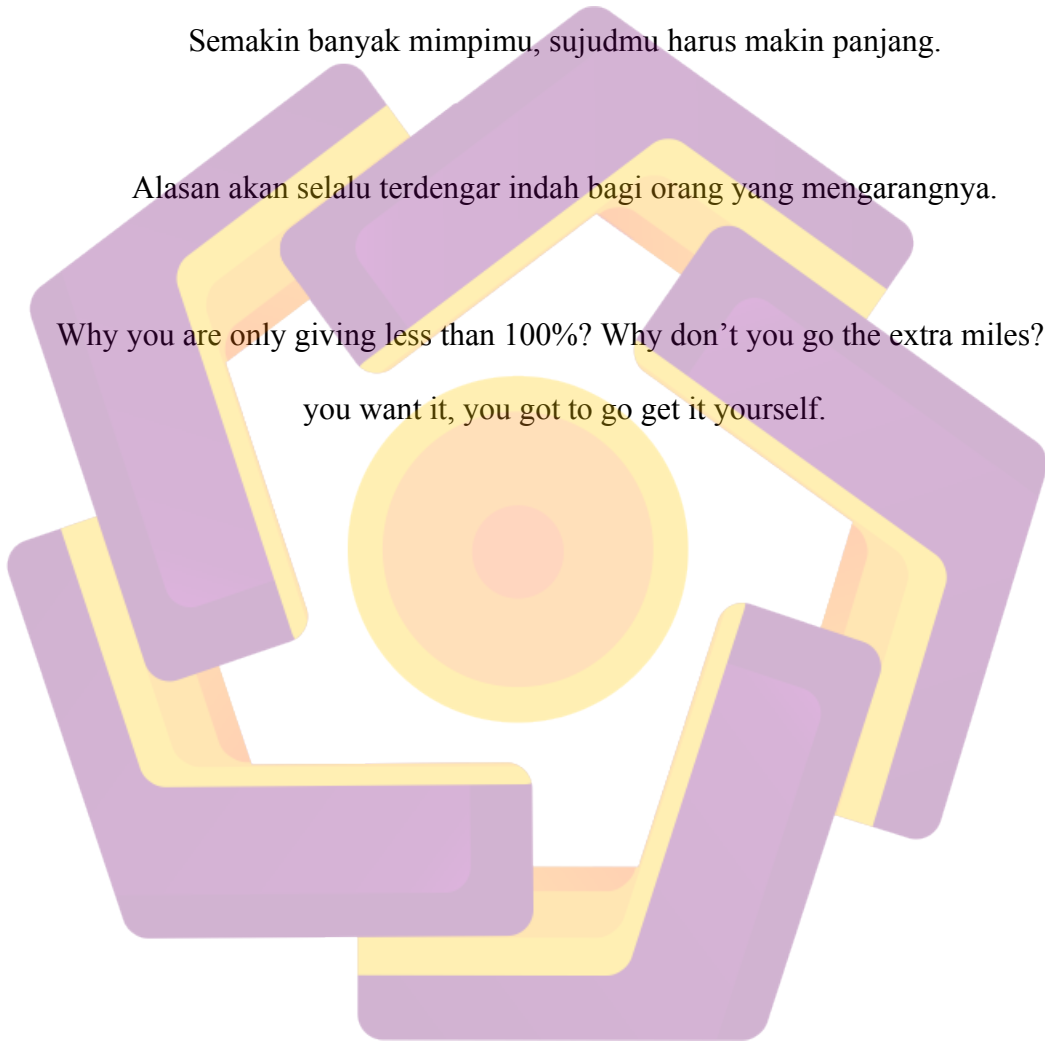
## MOTTO

Selama ikhtiar dan doa menjadi teman setia, maka jangan khawatir dengan masa depanmu.

Semakin banyak mimpimu, sujudmu harus makin panjang.

Alasan akan selalu terdengar indah bagi orang yang mengarangnya.

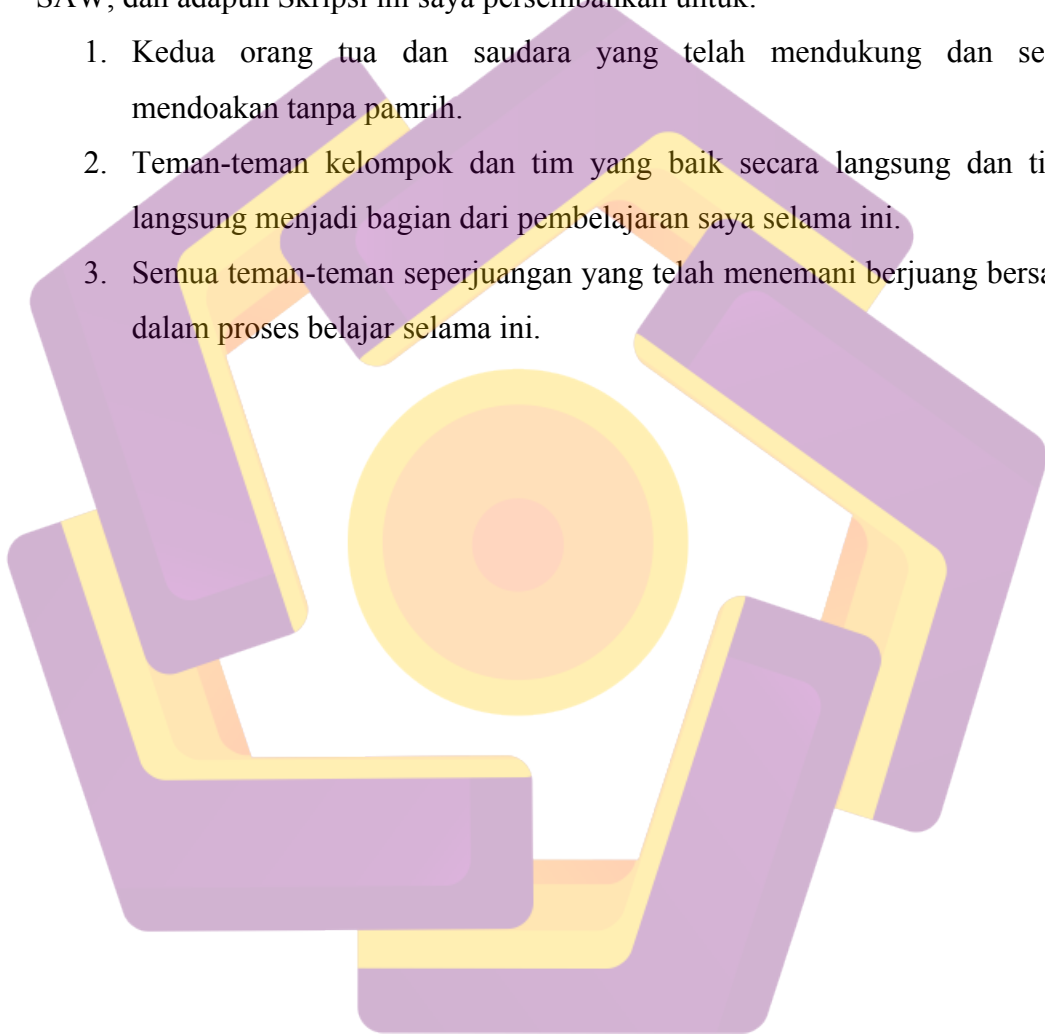
Why you are only giving less than 100%? Why don't you go the extra miles? If you want it, you got to go get it yourself.



## PERSEMBAHAN

Alhamdulillah, Puji syukur saya panjatkan kehadirat Allah SWT atas kemudahan yang diberikan untuk menyelesaikan Skripsi ini, shalawat serta salam semoga selalu tercurah kepada manusia terbaik sepanjang zaman, Muhammad SAW, dan adapun Skripsi ini saya persembahkan untuk:

1. Kedua orang tua dan saudara yang telah mendukung dan selalu mendoakan tanpa pamrih.
2. Teman-teman kelompok dan tim yang baik secara langsung dan tidak langsung menjadi bagian dari pembelajaran saya selama ini.
3. Semua teman-teman seperjuangan yang telah menemani berjuang bersama dalam proses belajar selama ini.



## KATA PENGANTAR

Puji syukur penulis panjatkan kepada Allah SWT yang telah melimpahkan rahmat dan hidayah-Nya, sehingga penulis dapat menyelesaikan penulisan skripsi ini dengan judul “Klasifikasi Multi-Label Pada Terjemahan Ayat Al-Qur’an Menggunakan Algoritma Multinomial Naive Bayes”, serta shalawat dan salam penulis curahkan kepada manusia terbaik, Muhammad SAW yang pernah diutus untuk menyempurnakan akhlak seluruh manusia.

Penulis menyadari sepenuhnya bahwa terselesaikannya skripsi ini karena adanya bantuan dan bimbingan dari berbagai pihak. Dalam pembuatan skripsi ini masih banyak sekali kekurangan-kekurangan dan kelemahan-kelemahannya. Oleh karena itu penulis berharap kepada semua pihak agar dapat menyampaikan kritik dan saran yang membangun untuk menambah kesempurnaan skripsi ini, namun penulis tetap berharap skripsi ini akan bermanfaat bagi semua pihak yang membacanya.

Yogyakarta, 11 September 2020

Penulis

**Muhammad Zulfakar**

**16.11.0500**



## DAFTAR ISI

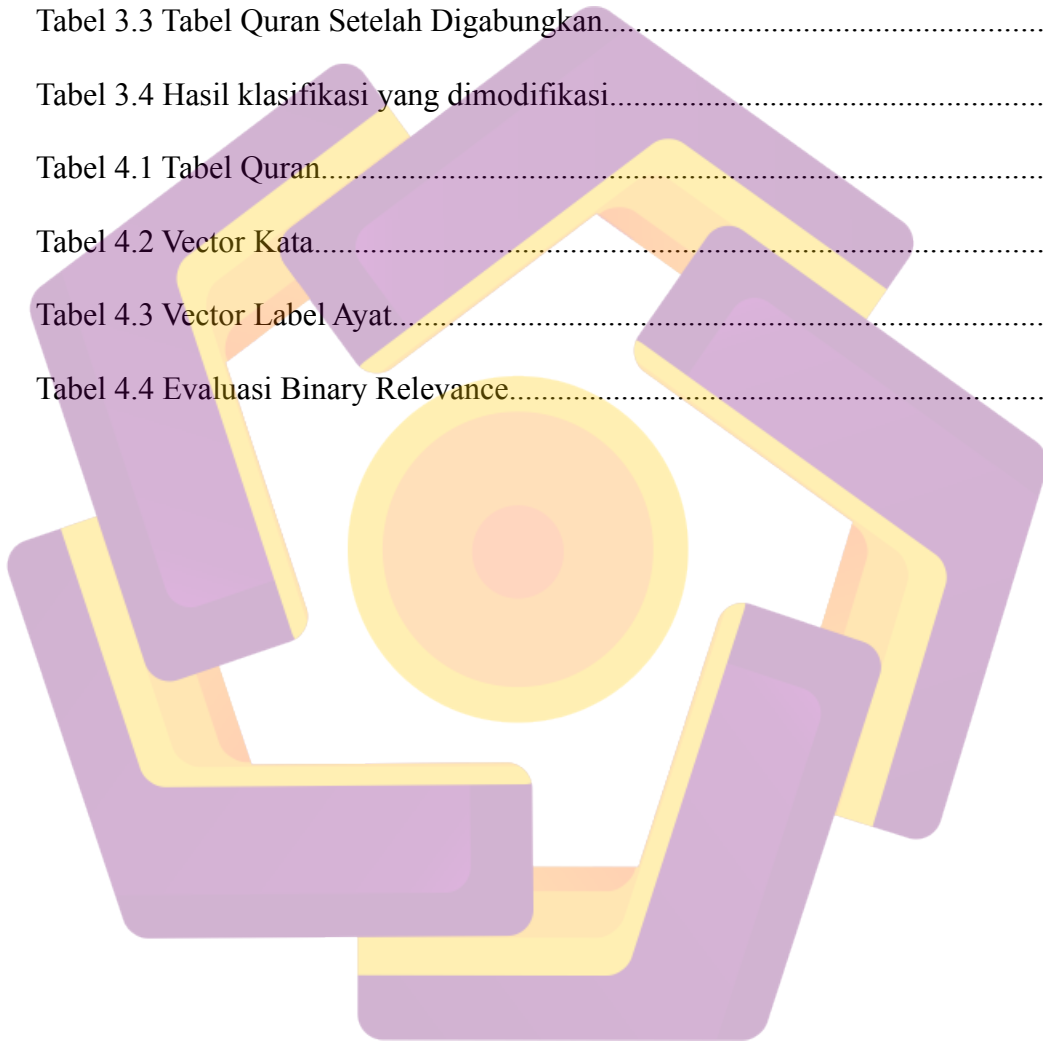
<b>JUDUL.....</b>	<b>I</b>
<b>PERSETUJUAN.....</b>	<b>III</b>
<b>PENGESAHAN.....</b>	<b>IV</b>
<b>PERNYATAAN.....</b>	<b>V</b>
<b>MOTTO.....</b>	<b>VI</b>
<b>PERSEMBAHAN.....</b>	<b>VII</b>
<b>KATA PENGANTAR.....</b>	<b>VIII</b>
<b>DAFTAR ISI.....</b>	<b>IX</b>
<b>DAFTAR TABEL.....</b>	<b>XII</b>
<b>DAFTAR GAMBAR.....</b>	<b>XIII</b>
<b>INTISARI.....</b>	<b>XVII</b>
<b>ABSTRAK.....</b>	<b>XVIII</b>
<b>BAB I PENDAHULUAN.....</b>	<b>1</b>
1.1 LATAR BELAKANG.....	1
1.2 RUMUSAN MASALAH.....	3
1.3 BATASAN MASALAH.....	3
1.4 MAKSUD DAN TUJUAN PENELITIAN.....	4
1.5 METODOLOGI PENELITIAN.....	4
1.5.1 Metode Pengumpulan Data.....	5
1.5.2 Metode Eksperimental.....	5
1.5.3 Metode Evaluasi.....	5
1.6 SISTEMATIKA PENULISAN.....	6
<b>BAB II LANDASAN TEORI.....</b>	<b>8</b>
2.1 KAJIAN PUSTAKA.....	8
2.2 DASAR TEORI.....	14

2.2.1 <i>Text Mining</i> .....	14
2.2.2 <i>Klasifikasi</i> .....	16
2.2.3 <i>Klasifikasi Multi-Label</i> .....	16
2.2.4 <i>Multinomial Naïve Bayes</i> .....	20
2.3 <b>PENGUJIAN</b> .....	21
2.3.1 <i>Hamming Loss</i> .....	21
2.4 <b>AL-QUR'AN</b> .....	22
2.4.1 <i>Aqidah</i> .....	22
2.4.2 <i>Akhlak</i> .....	22
2.4.3 <i>Syariah</i> .....	23
2.4.4 <i>Ilmu</i> .....	23
2.4.5 <i>Kisah</i> .....	23
2.4.6 <i>Alam Dunia</i> .....	24
2.4.7 <i>Alam Ghaib</i> .....	25
2.4.8 <i>Alam Akhirat</i> .....	26
2.5 <b>ANDROID</b> .....	26
2.6 <b>FLOWCHART</b> .....	27
2.7 <b>ANTARMUKA</b> .....	27
<b>BAB III METODE PENELITIAN</b> .....	<b>28</b>
3.1 <b>TINJAUAN UMUM</b> .....	28
3.2 <b>ALAT DAN BAHAN PENELITIAN</b> .....	29
3.2.1 <i>Hardware</i> .....	29
3.2.2 <i>Software</i> .....	30
3.3 <b>ALUR PENELITIAN</b> .....	30
3.3.1 <i>Data Acquisition</i> .....	32
3.3.2 <i>Preprocessing</i> .....	35
3.3.3 <i>Pembagian Dataset</i> .....	37
3.3.4 <i>Klasifikasi</i> .....	37
3.3.5 <i>Evaluasi Hasil Klasifikasi</i> .....	38
3.3.6 <i>Implementasi Hasil Klasifikasi</i> .....	38

3.4 PERANCANGAN SISTEM.....	38
<b>BAB IV HASIL DAN PEMBAHASAN.....</b>	<b>41</b>
4.1 PREPROCESSING.....	41
4.1.1 <i>Cleansing dan Stemming</i> .....	43
4.1.2 <i>Stopwords-Removal dan Tokenization</i> .....	43
4.1.3 <i>Feature Extraction</i> .....	44
4.2 PEMBAGIAN DATASET.....	45
4.3 KLASIFIKASI.....	46
4.3.1 <i>Implementasi Pendekatan Binary Relevance</i> .....	48
4.3.2 <i>Implementasi Pendekatan Classifier Chains</i> .....	48
4.3.3 <i>Implementasi Pendekatan Label Powerset</i> .....	49
4.4 EVALUASI HASIL KLASIFIKASI.....	50
4.4.1 <i>Evaluasi Binary Relevance</i> .....	50
4.4.2 <i>Evaluasi Classifier Chains</i> .....	52
4.4.3 <i>Evaluasi Label Powerset</i> .....	54
4.5 IMPLEMENTASI CLASSIFIER.....	55
4.6 PERANCANGAN APLIKASI.....	59
4.6.1 <i>Implementasi Antarmuka</i> .....	59
4.6.2 <i>Implementasi Fungsi Aplikasi</i> .....	63
4.7 PENGUJIAN APLIKASI.....	68
4.7.1 <i>Blackbox Testing</i> .....	68
<b>BAB V KESIMPULAN DAN SARAN.....</b>	<b>71</b>
5.1 KESIMPULAN.....	71
5.2 SARAN.....	72
<b>DAFTAR PUSTAKA.....</b>	<b>73</b>

## DAFTAR TABEL

Tabel 2.1 Perbandingan Penelitian Terdahulu.....	10
Tabel 3.1 Spesifikasi Hardware.....	28
Tabel 3.2 Quran.....	31
Tabel 3.3 Tabel Quran Setelah Digabungkan.....	32
Tabel 3.4 Hasil klasifikasi yang dimodifikasi.....	37
Tabel 4.1 Tabel Quran.....	39
Tabel 4.2 Vector Kata.....	43
Tabel 4.3 Vector Label Ayat.....	44
Tabel 4.4 Evaluasi Binary Relevance.....	47



## DAFTAR GAMBAR

Gambar 2.1: Contoh Dataset Multi-Label.....	15
Gambar 2.2: Proses Binary Relevance.....	16
Gambar 2.3: Dataset Classifier Chains.....	16
Gambar 2.4: Proses Classifier Chains.....	17
Gambar 2.5: Dataset Label Powerset.....	17
Gambar 2.6: Label Transformasi Powerset.....	18
Gambar 3.1 Alur penelitian.....	30
Gambar 3.2 Distribusi ayat berdasarkan topik.....	32
Gambar 3.3 Distribusi ayat berdasarkan banyak label.....	33
Gambar 3.4: Desain antarmuka aplikasi android.....	38
Gambar 4.1: Implementasi Cleansing dan Stemming.....	40
Gambar 4.2: Hasil Cleansing dan Stemming.....	40
Gambar 4.3: Implementasi Stopwords-Removal.....	41
Gambar 4.4: Hasil Stopwords-Removal dan Tokenization.....	41
Gambar 4.5 Bag of words diubah menjadi vector.....	42
Gambar 4.6: Pembagian dataset.....	42
Gambar 4.7 Implementasi Binary Relevance.....	45
Gambar 4.8 Implementasi Classifier Chains.....	45
Gambar 4.9 Implementasi Label Powerset.....	46
Gambar 4.10: Grafik kecepatan binary relevance.....	48
Gambar 4.11: Grafik hamming loss binary relevance.....	48
Gambar 4.12: Grafik kecepatan classifier chains.....	49

Gambar 4.13: Grafik hamming loss classifier chains.....	50
Gambar 4.14: Grafik kecepatan Label Powerset.....	51
Gambar 4.15: Grafik hamming loss label powerset.....	52
Gambar 4.16 Grafik hamming loss semua clasifier.....	53
Gambar 4.17: Import dataset Al-Quran yang belum diklasifikasi.....	53
Gambar 4.18: Transformasi data ke dalam vector bag of word.....	54
Gambar 4.19: Klasifikasi dataset.....	54
Gambar 4.20: Hasil klasifikasi.....	55
Gambar 4.21: Hasil akhir klasifikasi.....	55
Gambar 4.22: Hasil akhir klasifikasi lengkap.....	56
Gambar 4.23: Ekspor hasil klasifikasi.....	56
Gambar 4.24 Query alter table.....	57
Gambar 4.25 Implementasi halaman Discover.....	58
Gambar 4.26 Implementasi antarmuka per-topik.....	58
Gambar 4.27 Implementasi antarmuka halaman TopicDetail.....	59
Gambar 4.28 Implementasi antarmuka per-ayat.....	60
Gambar 4.29 Menyalin file database.....	61
Gambar 4.30 Pembacaan file database oleh aplikasi.....	62
Gambar 4.31 Data class Topik.....	62
Gambar 4.32 Dao Topic.....	63
Gambar 4.33 Data class verse.....	64
Gambar 4.34 Dao verse.....	65
Gambar 4.35 Pengujian fitur topik.....	66

Gambar 4.36 Hasil Pengujian Fitur Ayat.....67







## INTISARI

Al-Qur'an adalah firman Allah SWT yang menjadi kitab suci sekaligus pedoman bagi umat Islam di seluruh dunia yang setiap ayatnya mengandung makna dan hikmah. Ayat-ayat Al-Qur'an bisa diklasifikasikan ke dalam beberapa topik, diantara ayat-ayat tersebut terdapat ayat yang dapat diklasifikasikan ke dalam satu topik saja, namun ada juga yang dapat diklasifikasikan ke dalam beberapa topik yang berbeda yang dalam dunia Artificial Intelligence hal ini disebut dengan klasifikasi multi-label.

Penelitian ini bertujuan untuk mengklasifikasikan dokumen terjemahan Al-Qur'an berbahasa Indonesia menjadi 8 kategori yaitu Aqidah, Syariah, Akhlak, Ilmu, Kisah, Alam Dunia, Alam Ghaib, dan Alam Akhirat. Dalam penelitian ini dokumen yang digunakan yaitu, data set terjemahan Al-Qur'an berbahasa Indonesia yang diambil dari *qurandatabase.org*. Klasifikasi ini dilakukan dengan metode Naïve Bayes yang merupakan metode probabilistik memiliki beberapa keunggulan kesederhanaan dalam komputasi.

Dalam penelitian ini proses pengklasifikasian dilakukan dengan tiga pendekatan yaitu Binary Relevance, Classifiers Chain, dan Label Powerset. Kemudian akan dilakukan perbandingan dari ketiga pendekatan dan dipilih satu dengan hasil terbaik untuk dijadikan sebagai model untuk mengklasifikasikan ayat-ayat yang belum memiliki topik. Hasil dari penelitian ini nantinya berupa sebuah tabel database Al-Qur'an lengkap dengan topik untuk masing-masing ayat yang nantinya bisa diimplementasikan ke sistem secara luas, namun pada penelitian ini diimplementasikan ke aplikasi android.

**Kata Kunci:** Al-Qur'an, Klasifikasi, Dokumen, Naive Bayes, Android

## ABSTRAK

*Al-Qur'an is the word of Allah SWT which is the holy book as well as a guideline for Muslims around the world, each verse contains meaning and wisdom. The verses of the Al-Qur'an can be classified into several topics, among them there are verses that can be classified into only one topic, and some can be classified into several different topics. This case is called a multi-label classification in the world of artificial intelligence.*

*This study aims to classify the document translated into the Indonesian language Al-Qur'an into 8 categories, namely Aqidah (Aqidah), Syariah (Sharia), Akhlak (Morals), Ilmu (Knowledge), Kisah (Stories), Alam Dunia (World), Alam Ghaib (Unseen Nature), and Alam Akhirat (the Afterlife). The dataset in this study was taken from [alquranalhadi.com](http://alquranalhadi.com) with web scraping technique then continued with data processing. Then the data from the translation of the Al-Qur'an was taken from [qurandatabase.com](http://qurandatabase.com). This classification is carried out using the Naïve Bayes method, which is a probabilistic method that has several advantages of simplicity in computation.*

*In this study, the classification process is carried out with three approaches, namely Binary Relevance, Classifier Chains, and Label Powerset. Then a comparison of the three approaches will be carried out and the one with the best results will be selected to serve as a model for classifying verses that do not yet have a topic. The results of this research will be in the form of a complete Al-Qur'an database table with topics for each verse which later can be implemented to a wide information system, but in this study it was implemented into an android application.*

**Keywords:** *Al-Qur'an, Classification, Document, Naive Bayes, Android*