



**KLASIFIKASI AYAT AL-QURAN TERJEMAHAN MENGGUNAKAN METODE
*SUPPORT VECTOR MACHINE DAN K-NEAREST NEIGHBORS***

AAN SAPUTRA

161410277

**Skripsi ini Diajukan Sebagai Syarat Memperoleh Gelar
Sarjana Komputer**

PROGRAM STUDI SISTEM INFORMASI

FAKULTAS ILMU KOMPUTER

UNIVERSITAS BINA DARMA

PALEMBANG

2020

HALAMAN PENGESAHAN

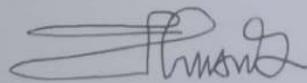
Klasifikasi Ayat Al-Quran Terjemahan Menggunakan
Metode Support Vector Machine dan K-Nearest Neighbors

AAN SAPUTRA

161410277

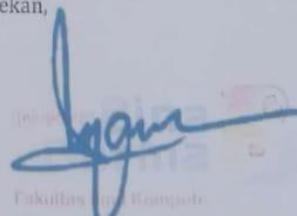
Telah diterima sebagai salah satu syarat untuk memperoleh gelar
Sarjana Komputer pada Program Studi Sistem Informasi

Pembimbing



Ilman Zuhri Yadi, M.M., M.Kom.

Palembang, 30 September 2020
Fakultas Ilmu Komputer
Universitas Bina Darma
Dekan,


Fakultas Ilmu Komputer
Universitas Bina Darma

Dedy Syamsuar, S.Kom., M.I.T., Ph.D.

HALAMAN PERSETUJUAN

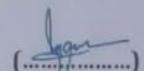
Skripsi Berjudul "Klasifikasi Ayat Al-Quran Terjemahan Menggunakan Metode Support Vector Machine dan K-Nearest Neighbors" Oleh "Aan Saputra", telah dipertahankan di depan komisi penguji pada hari Rabu tanggal 30 September 2020.

Komisi Penguji

1. Ketua : Ilman Zuhri Yadi, M.M., M.Kom.

(

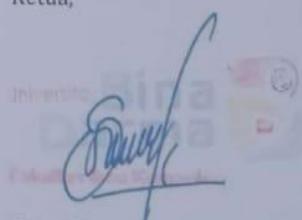
2. Anggota : Dedy Syamsuar, S.Kom., M.I.T., Ph.D.

(

3. Anggota : Dr. Edi Surya Negara, M.Kom.

(

Mengetahui,
Program Studi Sistem Informasi
Fakultas Ilmu Komputer
Universitas Bina Darma
Ketua,


Dr. Edi Surya Negara, M.Kom.

HALAMAN PERNYATAAN

Saya Bertanda tangan de bawah ini :

Nama : Aan Saputra
NIM : 161410277

Dengan ini menyatakan bahwa:

1. Karya tulis ini adalah hasil karya asli dan belum pernah diajukan untuk mendapatkan gelar akademik baik di universitas Bina Darma ataupun di universitas yang lain.
2. Karya tulis ini murni dari gagasan, rumusan dan penelitian saya sendiri dengan arahan dari pembimbing.
3. Dalam karya ini tidak ada karya ataupu pendapat orang lain, kecuali tulisan yang dikutip dengan mencantumkan nama dan tahun pengarang serta menambahkannya kedalam daftar pustaka.
4. Saya bersedia karya ini dicek keasliannya dengan turnitin dan dipublikasikan ke internet sehingga dapat diakses langsung oleh publik.
5. Surat pernyataan ini saya tulis dengan sebenar-benarnya dan seandainya terbukti melakukan ketidakbenaran dalam pernyataan ini, maka saya siap mendapatkan sanksi sesuai dengan peraturan perundang-undangan yang berlaku.

Demikian surat pernyataan ini saya buat agar dapat dipergunakan sebagai mana mestinya.

Palembang, 30 September 2020

Yang membuat pernyataan,



161410277

MOTTO & PERSEMBAHAN

MOTTO :

- ❖ *Buatlah orang terinspirasi melalui kita.*
- ❖ *Jadikanlah Kebenaran sebagai Kebiasaan, jangan jadikan Kebiasaan sebagai Kebenaran.*
- ❖ *Milikilah hati yang pandai bersyukur, Lisan yang selalu berusaha berkata baik dan tindakan yang selalu melangkah kepada kebaikan...*

Kupersembahkan kepada:

- ✓ Allah SWT yang senantiasa memberikan nikmat kesehatan, kekuatan serta kesabaran. Semoga keberhasilan ini menjadi satu langkah awal untuk masa depanku, dalam meraih cita cita.
- ✓ Kedua orang tuaku
- ✓ Keluargaku
- ✓ Dosen pembimbingku
- ✓ Kakak dan adikku
- ✓ Sahabat-sahabatku
- ✓ Teman seperjuangan
- ✓ Teman Santri
- ✓ Orang yang Sudah Berkontribusi
- ✓ Almamaterku

ABSTRACT

Al-Quran is the holy book of Muslims that fosters aspects of life both individually and socially. For Muslims, they should make the Koran a new life. At least the basic things that must be studied in the Al-Quran are prohibitions, commands and information contained in the Al-Quran. However, nowadays it is still rare to group Al-Quran into this classification, even though there are many classification methods that can be applied to translated Al-Quran Verses. Several methods can be used such as Support vector machine and k-nearest neighbor. In this study, the authors used these two methods to present a comparison of the level of classification in the translated Al-Quran Verse. In doing text classification, there are often words that are not important. So that affects the classification. Therefore, to solve this problem, the author uses stemming, tokenizing, and stopword processes. The results of the comparison of these two methods show that the KNN method ($k = 3,6,9$) gives better results than the linear SVM method.

Kata Kunci : Al-Quran, Classification, Support vector machine, k-nearest neighbors, *stemming*, *tokenizing*, dan *stopword*

ABSTRAK

Al-Quran merupakan kitab suci umat islam yang mengatur seluruh aspek kehidupan baik secara individu maupun bermasyarakat. Bagi umat islam, sudah seharusnya mereka mejadikan Al-Quran sebagai pedoman hidup. Minimal hal mendasar yang harus dipelajari dalam Al-Quran yaitu Larangan, perintah dan informasi yang terdapat di dalam Al-Quran. Namun, saat ini masih jarang pengelompokkan Al-Quran didalam klasifikasi tersebut, padahal Banyak metode pengklasifikasian yang dapat diterapkan pada Ayat Al-Quran terjemahan. Beberapa metode yang bisa digunakan seperti *Support vector machine* dan *k-nearest neighbors*. Pada penelitian ini, penulis menggunakan kedua metode tersebut untuk menampilkan perbandingan tingkat akurasi klasifikasi pada Ayat Al-Quran terjemahan. Didalam melakukan klasifikasi teks, sering terdapat kata-kata yang tidak penting. Sehingga mempengaruhi proses klasifikasi. Oleh karena itu, untuk mengatasi permasalahan tersebut, penulis menggunakan proses *stemming*, *tokenizing*, dan *stopword*. Hasil perbandingan dari kedua metode ini menunjukkan bahwa Metode KNN ($k=3,6,9$) memberikan hasil yang lebih baik daripada metode SVM *linear*.

Kata Kunci : Al-Quran, klasifikasi, *Support vector machine*, *k-nearest neighbors*, *stemming*, *tokenizing*, dan *stopword*

KATA PENGANTAR

Puji syukur saya panjatkan kehadirat Allah SWT yang telah memberi rahmat dan karunia-Nya, tidak lupa shalawat dan salam saya curah limpahkan kepada Nabi besar Muhammad SAW beserta para keluarga dan sahabatnya, karena atas berkat dan rahmat-Nya, saya dapat menyelesaikan skripsi ini.

Penulisan skripsi ini dilakukan dalam rangka memenuhi salah satu syarat untuk mencapai gelar Sarjana pada Program Studi Sistem Informasi Universitas Bina Darma. Saya menyadari bahwa skripsi ini bukanlah tujuan akhir dari belajar, karena belajar adalah sesuatu yang tidak terbatas. Hal ini disebabkan keterbatasan saya dari segi ilmu pengetahuan, bahan, data, dana dan waktu sehingga kritik dan saran yang sangat saya harapkan demi kesempurnaan skripsi ini.

Saya juga menyadari sepenuhnya bahwa skripsi ini masih jauh dari kesempurnaan baik isi, bahasa serta penulisannya, tanpa bantuan dan bimbingan dari berbagai pihak, dari masa perkuliahan sampai pada penyusunan skripsi, sangatlah sulit bagi saya untuk menyelesaikan skripsi ini. Oleh karena itu, saya mengucapkan terima kasih kepada :

1. Dr. Sunda Ariana, M.Pd., M. M selaku Rektor Universitas Bina Darma Palembang.
2. Dedy Syamsuar, PH. D selaku Dekan Fakultas Ilmu Komputer Universitas Bina Darma Palembang.
3. Dr. Edi Surya Negara, M.Kom selaku Ketua Program Studi Sistem Informasi.
4. Deni Erlansyah, M.M., M.Kom Selaku pembimbing akademik yang selalu memberikan saya arahan selama proses perkuliahan.
5. Ilman Zuhri Yadi, M. M., M.Kom selaku Dosen Pembimbing Tugas Akhir yang telah memberikan motivasi serta bimbingan yang luar biasa dalam penulisan skripsi ini.
6. Yesi Novaria Kunang, ST., M.Kom selaku Dosen Pembimbing Tugas Akhir

yang telah memberikan masukan-masukan terbaiknya untuk kelancaran penulisan skripsi ini.

7. Kedua orang yang paling berharga bagi saya yaitu orang tua saya bak (SARI PUDIN) & mak (YATI) Terima kasih atas semua cinta yang telah bak dan mak berikan kepada saya, terimakasih untuk perjuangan dan rasa sakit yang telah kalian lalui sehingga bisa membuat saya sampai dititik ini.
8. Saudaraku yang terbaik kakak Fitrah Fengki Arista dan adikku Cindi Caniaga yang telah memberikan doa dan dukungannya hingga selesai.
9. Gusman Suryadi, S.E selaku pembina Pesantren Mahasiswa Palembang yang selalu mendukung dan mendo'akan kebaikan untuk kelancaran penulisan skripsi ini.
10. Teman-teman "Pesantren Mahasiswa Palembang Mushtofaynal Akhyar" yang selalu mendukungku
11. Teman kelas SI J yang telah berjuang bersama-sama selama delapan semester.
12. Teman-teman terdekat dan seperjuangan yang selalu memberikan saran, bantuan dan semangat.
13. Dan seluruh rekan-rekan yang telah membantu penulis dalam penulisan tugas akhir ini.

DAFTAR ISI

HALAMAN PENGESAHAN	ii
HALAMAN PERSETUJUAN	iii
HALAMAN PERNYATAAN	iii
MOTTO & PERSEMBAHAN	v
MOTTO :	v
Kupersembahkan kepada:	v
ABSTRACT	viii
ABSTRAK	vi
KATA PENGANTAR.....	viii
DAFTAR ISI	x
DAFTAR GAMBAR.....	xii
DAFTAR TABEL	xiii
BAB I	1
1.1. Latar Belakang	1
1.2. Rumusan Masalah.....	2
1.3. Tujuan Penelitian.....	2
1.4. Batasan Masalah.....	3
1.5. Manfaat Penelitian.....	3
1.6. Metodologi Penelitian	4
1.6.1 Waktu Penelitian.....	4
1.6.2 Metode Pengumpulan data	4
1.6.3 Metode Penelitian.....	4
1.7. Sistematika Penulisan	7
BAB II.....	9
2.1 Penelitian Terdahulu	9
2.2 Teori Umum	12
2.2.1 Al-Quran	12
2.2.2 <i>Text mining</i>	13
2.2.3 Klasifikasi.....	14
2.2.4 <i>Pre Processing Text</i>	14

2.2.5 <i>tfidf</i>	14
2.2.6 Algoritma <i>Support vector machine</i>	15
2.2.7 Algoritma <i>k-nearest neighbors</i>	17
2.2.8 Pengukuran Performa.....	18
2.2.9 <i>Anaconda</i>	20
2.2.10 Bahasa Pemrograman <i>Python</i>	20
2.2.11 <i>Jupyter Notebook</i>	21
BAB III	22
3.1 Sumber Data	22
3.2 Metode Penelitian.....	22
3.3 Proses Kerja Penelitian.....	24
BAB IV.....	31
4.1 <i>Pre Processing Text</i>	31
4.1.1 <i>Stemming</i>	31
4.1.2 <i>Stopword</i>	32
4.1.3 <i>Tokenizing</i>	33
4.1.4 <i>Pembobotan Kata (Term Weighting TF-IDF)</i>	33
4.2 <i>Support Vector Machine</i> dalam <i>Text mining</i>	34
4.2.1 fungsi <i>kernel linear</i> dalam <i>confusion matrix</i>	34
4.2.2 Performensi klasifikasi.....	35
4.2.3 tingkat akurasi SVM.....	35
4.3 <i>K-Nearest Neighbors</i> dalam <i>text mining</i>	36
4.3.1 <i>confusion matrix</i> KNN.....	36
4.3.2 Hasil performansi klasifikasi.....	38
4.3.3 Hasil KNN menggunakan k=3,6,9 pada tingkat akurasi	40
4.4 Perbandingan antara SVM dan KNN.....	42
BAB V	43
5.1 Kesimpulan.....	43
5.2 Saran	43
DAFTAR PUSTAKA	44

DAFTAR GAMBAR

Gambar 1.1	Kerangka metode penelitian	6
Gambar 2.1	<i>Data Spiral(nonlinear)</i>	16
Gambar 2.2	Transformasi ruang fungsi	16
Gambar 3.1	diagram analisis penelitian	25
Gambar 3.2	Tampilan download <i>anaconda</i>	26
Gambar 3.3	Tampilan awal <i>anaconda</i>	26
Gambar 3.4	<i>Library Jupyter notebook</i>	27
Gambar 3.5	proses <i>stemming</i>	27
Gambar 3.6	proses <i>Stopword</i>	28
Gambar 3.7	Proses <i>tokenizing</i>	28
Gambar 3.8	hasil <i>pre processing text</i>	29
Gambar 3.9	proses pembobotan kata	29
Gambar 3.10	pembobotan kata pada <i>tf-idf</i>	30
Gambar 3.11	pembagian data <i>training</i> dan <i>testing</i>	30
Gambar 3.12	<i>coding</i> metode SVM	31
Gambar 3.13	<i>coding</i> metode KNN	31
Gambar 4.1	<i>rank term</i>	34
Gambar 4.2	<i>confusion matrix</i> SVM.....	35
Gambar 4.3	analisis akurasi SVM	36
Gambar 4.4	analisis <i>confusion matrix</i> k=3.....	37
Gambar 4.5	analisis <i>confusion matrix</i> k=6.....	38
Gambar 4.6	analisis <i>confusion matrix</i> k=9.....	38
Gambar 4.7	analisis tingkat akurasi k=3	41
Gambar 4.8	analisis tingkat akurasi k=6	41
Gambar 4.9	analisis tingkat akurasi k=9	52

DAFTAR TABEL

Tabel 2.1 Daftar Penelitian Sebelumnya.....	9
Tabel 2.2 Perbandingan penelitian terdahulu dan sekarang.....	11
Tabel 2.3 ketepatan <i>Confusion Matrix</i>	20
Tabel 3.1 Struktur Data penelitian	23
Tabel 4.1 Proses <i>Stemming</i> data.....	32
Tabel 4.2 Proses <i>Stopword</i> data	33
Tabel 4.3 Proses <i>Tokenizing</i> data.....	34
Tabel 4.4 <i>Confusion matrix</i>	35
Tabel 4.5 Performansi Klasifikasi.....	36
Tabel 4.6 Hasil akurasi SVM.....	36
Tabel 4.7 <i>Confusion matrix</i> k=3	37
Tabel 4.8 <i>Confusion matrix</i> k=6	38
Tabel 4.9 <i>Confusion matrix</i> k=9	39
Tabel 4.10 Performansi Klasifikasi k=3	39
Tabel 4.11 Performansi Klasifikasi k=6	40
Tabel 4.12 Performansi Klasifikasi k=9	40
Tabel 4.13 Tingkat akurasi k=3	41
Tabel 4.14 Tingkat akurasi k=6.....	42
Tabel 4.15 Tingkat akurasi k=9	42
Tabel 4.16 Rata-rata nilai k=3,69 pada KNN.....	42
Tabel 4.17 Perbandingan antara SVM dan KNN	43