

SKRIPSI

**Implementasi *Cosine Similarity* dan *Naïve Bayes Classifier* Untuk
Mencari Dalil Pada Artikel Islami Berdasarkan *Indeks*
*Al-Qur'an***



LANA RAHIM

15650002

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI SUNAN KALIJAGA
YOGYAKARTA**

2019

**IMPLEMENTASI *COSINE SIMILARITY* DAN *NAÏVE BAYES*
CLASSIFIER UNTUK Mencari Dalil pada Artikel
Islami Berdasarkan *INDEKS AL-QUR'AN***

SKRIPSI

Untuk memenuhi sebagian persyaratan
mencapai derajat Sarjana S-1
Program Studi Teknik Informatika



Disusun Oleh
LANA RAHIM
15650002

Kepada

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI SUNAN KALIJAGA
YOGYAKARTA**

2019



KEMENTERIAN AGAMA
UNIVERSITAS ISLAM NEGERI SUNAN KALIJAGA
FAKULTAS SAINS DAN TEKNOLOGI

Jl. Marsda Adisucipto Telp. (0274) 540971 Fax. (0274) 519739 Yogyakarta 55281

PENGESAHAN TUGAS AKHIR

Nomor : B-1854/Un.02/DST/PP.00.9/05/2019

Tugas Akhir dengan judul : Implementasi Cosine Similarity dan Naive Bayes Classifier Untuk Mencari Dalil Pada Artikel Islami Berdasarkan Indeks AI -Qur'an

yang dipersiapkan dan disusun oleh:

Nama : LANA RAHIM
Nomor Induk Mahasiswa : 15650002
Telah diujikan pada : Senin, 06 Mei 2019
Nilai ujian Tugas Akhir : A

dinyatakan telah diterima oleh Fakultas Sains dan Teknologi UIN Sunan Kalijaga Yogyakarta

TIM UJIAN TUGAS AKHIR

Ketua Sidang

Muhammad Didik Rohmad Wahyudi, S.T., MT.
NIP. 19760812 200901 1 015

Penguji I

Dr. Agung Fatwanto, S.Si., M.Kom.
NIP. 19770103 200501 1 003

Penguji II

Agus Mulyanto, S.Si., M.Kom.
NIP. 19710823 199903 1 003

Yogyakarta, 06 Mei 2019
UIN Sunan Kalijaga
Fakultas Sains dan Teknologi
DEKAN



Dr. Muntono, M.Si.
NIP. 1969/212 200003 1 001



SURAT PERSETUJUAN SKRIPSI/TUGAS AKHIR

Hal : Persetujuan Skripsi

Lamp :

Kepada

Yth. Dekan Fakultas Sains dan Teknologi

UIN Sunan Kalijaga Yogyakarta

di Yogyakarta

Assalamu'alaikum wr. wb.

Setelah membaca, meneliti, memberikan petunjuk dan mengoreksi serta mengadakan perbaikan seperlunya, maka kami selaku pembimbing berpendapat bahwa skripsi Saudara:

Nama : Lana Rahim
NIM : 15650002
Judul Skripsi : " Impementasi *Cosine Semilirity* dan *Naïve Bayes Classifier*
Untuk Mencari Dalil Pada Artikel Islami Berdasarkan *Indeks Al-Qur'an*"

sudah dapat diajukan kembali kepada Program Studi Teknik Informatika Fakultas Sains dan Teknologi UIN Sunan Kalijaga Yogyakarta sebagai salah satu syarat untuk memperoleh gelar Sarjana Strata Satu dalam Program Studi Teknik Informatika

Dengan ini kami mengharap agar skripsi/tugas akhir Saudara tersebut di atas dapat segera dimunaqsyahkan. Atas perhatiannya kami ucapkan terima kasih.

Wassalamu'alaikum wr. wb.

Yogyakarta, 30 April 2019

Pembimbing

M. Didik Rohmad Wahyudi, S.T., MT.

NIP. 19760812 200901 1 015

PERNYATAAN KEASLIAN SKRIPSI

Saya yang bertanda tangan dibawah ini :

Nama : Lana Rahim

NIM : 15650002

Jurusan : Teknik Informatika

Fakultas : Sains dan Teknologi

Menyatakan bahwa skripsi saya yang berjudul “*Impementasi Cosine Semilirity dan Naïve Bayes Classifier Untuk Mencari Dalil Pada Artikel Islami Berdasarkan Indeks Al-Qur’an*” merupakan hasil penelitian saya sendiri, tidak terdapat pada karya yang pernah di ajukan untuk memperoleh gelar kesarjana di suatu perguruan tinggi, dan bukan plagiasi karya orang lain kecuali yang secara tertulis diacu dalam naskah ini dan disebutkan dalam daftar pustaka.

Yogyakarta, 15 April 2019



Lana Rahim

NIM.15650002

KATA PENGANTAR

Puji syukur kehadirat Allah SWT yang telah melimpahkan rahmat dan hidayah-Nya sehingga penulis dapat menyelesaikan penelitian yang berjudul **Implementasi Cosine Similarity Dan Naivebayes Untuk Mencari Dalil Pada Artikel Islami Berdasarkan Indeks Al-Quran** sebagai salah satu syarat untuk mencapai gelar sarjana program studi Teknik Informatika UIN Sunan Kalijaga Yogyakarta. Sholawat serta salam selalu tercurahkan kepada junjungan kita Nabi Agung Muhammad SAW beserta seluruh keluarga dan sahabat beliau.

Penulis menyadari bahwa apa yang dilakukan dalam penyusunan laporan penelitian ini masih terlalu jauh dari kesempurnaan. Oleh karena itu, penulis mengharapkan kritik dan saran yang berguna dalam penyempurnaan analisa ini di masa yang akan datang. Semoga apa yang telah penulis lakukan dapat bermanfaat bagi pembaca.

Tidak lupa penulis juga mengucapkan terimakasih kepada pihak-pihak yang telah membantu dalam penyelesaian skripsi ini, baik secara langsung maupun tidak langsung. Ucapan terimakasih penulis sampaikan kepada:

1. Bapak Prof. Drs. K.H. Yudian Wahyudi, M.A., Ph.D., selaku Rektor UIN Sunan Kalijaga Yogyakarta.
2. Bapak Dr. Murtono, M.Si., selaku Dekan Fakultas Sains dan Teknologi UIN Sunan Kalijaga Yogyakarta.
3. Bapak Sumarsono, S.T., M.Kom., selaku Ketua Program Studi Teknik Informatika UIN Sunan Kalijaga Yogyakarta.

4. Bapak Muhammad Didik Rohmad Wahyudi, S.T., MT., selaku dosen pembimbing skripsi yang telah sabar dan meluangkan waktunya untuk memberikan motivasi, koreksi dan kritik saran kepada penulis sehingga skripsi ini dapat diselesaikan.
5. Bapak Dr. Bambang Sugiantoro, M.T., selaku Dosen Pembimbing Akademik.
6. Dr. Agung Fatwanto, S.Si., M.Kom., Agus Mulyanto, S.Si., M.Kom., Ph.D., Aulia Faqih Rifa'i, M.Kom., M. Taufiq NUruzzaman, S.T., Maria Ulfah Siregar, S.Kom. MIT., Ph.D., Nurochman, S.Kom., M.Kom., Rahmat Hidayat, S.Kom., M.Cs., Dr. Shofwatul 'Uyun, S.T., M.Kom., selaku dosen pengampu mata kuliah program studi Teknik Informatika UIN Sunan Kalijaga Yogyakarta yang telah banyak membantu penulis hingga penulis dapat menyusun tugas akhir.
7. Seluruh staf dan karyawan Fakultas Sains dan Teknologi UIN Sunan Kalijaga Yogyakarta.
8. Ucapan terimakasih yang terdalam untuk kedua orangtua, Ibu Sahamidah dan Bapak Rahman Hakim yang selalu memberikan doa, perhatian, kasih sayang dan semua support yang telah diberikan.
9. Nafi dan Umam yang telah senantiasa membantu penulis dalam proses awal berjalannya skripsi hingga selesai.
10. Seluruh Teman-Teman Teknik Informatika 2015 yang tidak dapat penulis sebutkan satu-satu.
11. Teman-teman KKN Angkatan 96 Kelompok 242.

12. Serta semua pihak yang tidak dapat penulis sebutkan satu persatu dan telah memberikan banyak campur tangan, doa, support sehingga penelitian ini dapat terselesaikan.

Yogyakarta, 29 April 2019

Penulis



HALAMAN PERSEMBAHAN

Dengan penuh rasa syukur, Karya sederhana ini saya persembahkan untuk :

”Kedua Orang Tua”



HALAMAN MOTTO

”Janganlah kamu berjalan dimuka bumi ini dengan penuh kesombongan”

(QS. Al-isra:237)



DAFTAR ISI

HALAMAN JUDUL	i
LEMBAR PENGESAHAN SKRIPSI/TUGAS AKHIR	ii
LEMBAR PERSETUJUAN SKRIPSI/TUGAS AKHIR	iii
HALAMAN PERNYATAAN KEASLIAN SKRIPSI	iv
KATA PENGANTAR	v
HALAMAN PERSEMBAHAN	viii
HALAMAN MOTTO	ix
DAFTAR ISI	x
DAFTAR TABEL	xiii
DAFTAR GAMBAR	xiv
DAFTAR LAMBANG	xv
INTISARI	xvi
ABSTRACT	xvii
I PENDAHULUAN	1
1.1. Latar Belakang	1
1.2. Rumusan Masalah	3
1.3. Batasan Masalah	3
1.4. Tujuan Penelitian	4
1.5. Manfaat Penelitian	4
1.6. Keaslian Penelitian	5
1.7. Sistematika Penulisan	5
II TINJAUAN PUSTAKA DAN LANDASAN TEORI	7

2.1. Tinjauan Pustaka	7
2.2. Landasan Teori	13
2.2.1. <i>Text Mining</i>	13
2.2.2. <i>Preprocessing</i>	14
2.2.3. <i>Naïve Bayes Classifier</i>	20
2.2.4. <i>TF/IDF</i>	23
2.2.5. Klasifikasi	25
2.2.6. Python	26
2.2.7. Al-Qur'an	26
2.2.8. <i>Cosine Similarity</i>	27
III METODE PENELITIAN	30
3.1. Metode Penelitian	30
3.2. Tahapan-tahapan Penelitian	30
3.2.1. Studi Pustaka	31
3.2.2. Pengumpulan Data	31
3.2.3. <i>Preprocessing</i>	31
3.2.4. Analisa Hasil	32
3.2.5. Implementasi	32
IV HASIL DAN PEMBAHASAN	33
4.1. Pengolahan Data	34
4.1.1. <i>Non Summarization</i>	35
4.1.2. <i>Text Processing</i>	35
4.1.3. <i>Text Summarization</i>	38

4.1.4. <i>Text Processing</i>	39
4.2. Analisa dan Hasil	41
4.2.1. Analisa	41
4.3. Pengujian	50
4.3.1. Pengujian Sistem	50
V KESIMPULAN DAN SARAN	52
5.1. Kesimpulan	52
5.2. Saran	52
DAFTAR PUSTAKA	53
LAMPIRAN	56
CURICULUM VITAE	80



DAFTAR TABEL

2.1	<i>Tinjaun Pustaka</i>	9
2.2	Contoh Pemecahan Kalimat	15
2.3	Contoh <i>Tokenizing</i>	16
2.4	Contoh <i>Stopword</i>	17
4.1	Contoh Dokumen Uji	34
4.2	Hasil <i>Case Folding</i> Dokumen Uji Non-Summarization	35
4.3	Hasil <i>Cleansing</i> Dokumen Uji Non-Summarization	36
4.4	Hasil <i>Stemming</i> Dokumen Uji Non-Summarization	37
4.5	Hasil <i>Summarization</i> Dokumen Uji	38
4.6	Hasil <i>Case Folding</i> Dokumen Uji	39
4.7	Hasil <i>Cleansing</i> Dokumen Uji	40
4.8	Hasil <i>Stemming</i> Dokumen Uji	40
4.9	Contoh Hasil <i>Preprocessing</i> Data Artikel beserta label	41
4.10	Data <i>Frequency</i>	42
4.11	Data <i>Frequency</i> Lanjut	43
4.12	<i>Term Frequency</i> Dokumen 1	44
4.13	<i>Term Frequency</i> Dokumen 2	45
4.14	<i>Term Frequency</i> Dokumen 2 Lanjut	46
4.15	<i>Perhitungan TF-IDF</i>	48
4.16	<i>Cosine Similarity</i>	49

DAFTAR GAMBAR

3.1 Tahapan Penelitian	30
4.1 Tahapan Penelitian	33
4.2 <i>Source Code Case Folding Non-Summarization</i>	36
4.3 <i>Source Code Cleansing Non-Summarization</i>	37
4.4 <i>Source Code Steaming Non-Summarization</i>	38
4.5 <i>Source Code Summarize</i>	38
4.6 <i>Source Code Case Folding</i>	39
4.7 <i>Source Code Cleansing</i>	40
4.8 <i>Source Code Steaming</i>	41



DAFTAR LAMBANG

$x \in A$: x anggota A
$A \subseteq X$: A himpunan bagian (<i>subset</i>) atau sama dengan X
\mathbb{N}	: himpunan semua asli
\mathbb{Z}	: himpunan semua bilangan bulat
\mathbb{Z}^+	: himpunan semua bilangan bulat positif
\mathbb{R}	:himpunan semua bilangan real
C_r^n	: r -kombinasi dari n unsur yang berbeda
■	: akhir suatu bukti
□	: akhir suatu contoh
→	: menuju
$\sum_{i=1}^n a_i$: penjumlahan $a_1 + a_2 + \cdots + a_n$
$\prod_{i=1}^n a_i$: perkalian $a_1 \cdot a_2 \cdots a_n$
$p \Rightarrow q$: jika p maka q
\Leftrightarrow	: jika dan hanya jika
$x \leftarrow a$: nilai a dimasukkan ke x

INTISARI

Implementasi *Cosine Similarity* dan *Naïve Bayes Classifier* Untuk Mencari Dalil Pada Artikel Islami Berdasarkan *Indeks Al-Qur'an*

Oleh

LANA RAHIM

15650002

Indeks Al-Qur'an merupakan sebuah piranti navigasi yang digunakan untuk menelusuri keluasan dan keragaman tema-tema dalam Al-Qur'an. Indeks Al-Qur'an digunakan sebagai rujukan untuk mendapatkan ayat-ayat yang terdapat dalam Al-Qur'an. Namun pada proses mencari rujukan masih dilakukan secara manual sehingga proses pencariannya memakan waktu lama.

Membangun sebuah sistem cerdas yang dapat melakukan pengklasifikasian dan mencari dalil dari sebuah dokumen sesuai dengan indeks *Al-Qur'an* diharapkan dapat membantu masyarakat, baik muslim maupun non-muslim untuk memahami dan mengerti tentang ayat dalam *Al-Qur'an*. Dengan metode *cosine similarity* dan *Naive Bayes* dapat digunakan untuk melakukan pengklasifikasi dan mencari ayat yang relevan sesuai dengan Al-Qur'an.

Penelitian ini dilakukan dengan tahap *preprocessing* untuk artikel islami, pembobotan dengan *TF-IDF*, dan perhitungan dengan *Cosine Similarity*. Hasil uji coba yang dilakukan dengan menggunakan *confusion matrix* didapatkan: *accuracy* 38%, *precision* 9%, dan *recall* 8%.

Kata Kunci : *Cosine Similarity*, *Naïve Bayes*, *Al-Qur'an*, Artikel Islami

ABSTRACT

Implementation of Cosine Similarity and Naive Bayes Classifier for Seeking Proposals in Islamic Articles Based on the Qur'an Index

By

LANA RAHIM

15650002

The Al-Qur'an Index is a navigation tool that is used to explore the breadth and diversity of themes in the Qur'an. The Al-Qur'an Index is used as a reference to get the verses contained in the Qur'an. But in the process of searching for referrals it is still done manually so that the search process takes a long time.

Building an intelligent system that can classify and search for arguments from a document in accordance with the Al-Qur'an index is expected to help people, both Muslims and non-Muslims, to understand and understand the verses in the Qur'an. With the method of cosine similarity and Naive Bayes can be used to do classifiers and look for relevant verses according to the Qur'an.

This research was conducted with the preprocessing stage for Islamic articles, weighting with TF-IDF, and calculations with Cosine Similarity. The results of the tests carried out using confusion matrix were obtained: 38% accuracy, 9% precision, and 8% recall.

Keywords: Cosine Similarity, Naive Bayes, Qur'an, Islamic Articles

BAB I

PENDULUAN

1.1. Latar Belakang

Al-Quran adalah sebuah dokumen dan sebuah perisai untuk umat manusia khususnya untuk umat beragama Islam (Muslim). Al Qur'an sendiri menamakan dirinya petunjuk bagi umat manusia (*Huda li al-Nas*). Al Qur'an wahyu dari Allah swt, bukan dari hawa nafsu perkataan Nabi Muhammad saw. Al Qur'an memuat aturan-aturan kehidupan manusia di dunia. Al Qur'an merupakan petunjuk bagi orang-orang yang beriman dan bertaqwa. Di dalam al Qur'an terdapat rahmat yang besar dan pelajaran bagi orang-orang yang beriman. Al Qur'an merupakan petunjuk yang dapat mengeluarkan manusia dari kegelapan menuju jalan yang terang.

Al Quran terdiri dari 114 surat, masing-masing surat terdiri dari beberapa ayat. Jumlah ayat dalam Al Quran mencapai 6236 ayat. *Al Quran* dibagi lagi menjadi 30 bagian yang sama banyak yang disebut dengan juz. Setiap juz dibagi menjadi beberapa ruku. Selain itu Al-Qur'an juga terbagi menjadi 14 indeks, dimana setiap indeks menjelaskan inti dari masing-masing ayat yang terdapat dalam Al-Qur'an. Surat-surat dalam *Al Quran* memiliki jumlah ruku yang berbeda-beda tergantung banyaknya ayat dalam surat dan panjang pendeknya masing-masing ayat. Ayat dalam kitab suci *Al-Quran* merupakan objek yang menarik untuk dikaji dan diteliti lebih lanjut.

Indeks Al-Qur'an adalah sebuah piranti navigasi yang digunakan untuk menelusuri keluasan dan keragaman tema-tema dalam Al-Qu'an. Dikatakan sebagai

piranti navigasi karena, digunakan sebagai rujukan untuk mendapatkan ayat-ayat yang terdapat dalam Al-Qur'an. Indeks Al-Qur'an ini sangat dibutuhkan oleh orang-orang, baik itu ulama, mubalig, pemikir, penafsir, mahasiswa, dan semua orang untuk dijadikan sebuah rujukan dalam mencari ayat-ayat Al-Qur'an.

Namun proses mencari rujukan ini masih dilakukan secara manual, sehingga dibutuhkan waktu yang lama untuk mendapatkan ayat-ayat Al-Qur'an. Oleh karena itu untuk mempermudah atau mempercepat dalam pencarian ayat-ayat Al-Qur'an maka diperlukan sebuah sistem cerdas. Sistem cerdas merupakan sistem yang dapat mengadopsi kecerdasan manusia melalui proses *learning* yang dilakukan. Dengan adanya sistem cerdas ini diharapkan dapat membantu orang-orang dalam mencari rujukan pada ayat-ayat Al-Qur'an secara cepat.

Dalam sistem cerdas terdapat beberapa konsep salah satunya yakni *text mining*. *Text mining* merupakan sebuah proses penambangan data dari sebuah data untuk diambil informasinya. Dalam *data mining* terdapat berbagai algoritma yang bisa digunakan seperti, *naive bayes classifier*, *k-means clustering*, *cosine similarity*, *decision tree* dan *apriory*.

Peneliti tertarik untuk melakukan penelitian ini, ketertarikan peneliti berdasarkan beberapa alasan, yaitu : Pertama, indeks Al-Quran dengan berbagai ragam dan bentuk penyajiannya berusaha memberikan informasi sebaik mungkin terkait dengan keberadaan ayat-ayat yang dibutuhkan dalam rangka mendukung usaha membumikan Al-Quran di tengah-tengah masyarakat. Kedua, indeks Al-Quran dijadikan rujukan oleh para pengkaji Al-Quran yang dikembangkan sesuai istilah-istilah kontemporer di bidang keagamaan yang terus berkembang setiap saat. Ke-

tiga, membantu masyarakat menyajikan isi artikel yang diperlukan dengan mudah, cepat, dan tepat sesuai dengan indeks ayat-ayat Al-Quran.

Peneliti menggunakan Metode Naive Bayes Classifier dan Cosine Similarity karena memiliki beberapa kelebihan, antara lain sederhana, cepat dan berakurasi tinggi. metode Naive Bayes Classifier untuk klasifikasi atau kategorisasi teks menggunakan atribut kata yang muncul dalam suatu dokumen sebagai dasar klasifikasinya. Penelitian Rish dalam jurnal Hamzah (2012) menunjukkan bahwa meskipun asumsi independensi antar kata dalam dokumen tidak sepenuhnya dapat dipenuhi, tetapi kinerja NBC dalam klasifikasi relatif sangat bagus.

Dengan demikian, peneliti tertarik untuk melakukan penelitian dengan judul **Implementasi *Cosine Similarity* Dan *Naive Bayes Classifier* Untuk Mencari Dalil Pada Artikel Islami Berdasarkan Indeks Al-Quran.**

1.2. Rumusan Masalah

Berdasarkan latar belakang tersebut dirumuskan masalah yaitu belum adanya implementasi metode *cosine similarity* dan *naive bayes classifier* untuk mencari dalil pada artikel islami berdasarkan indeks al-qur'an.

1.3. Batasan Masalah

Untuk lebih focus pada penelitian ini penulis membatasi pembahasan yang akan dibahas pada pembuatan aplikasi ini:

1. Data uji yang digunakan yaitu berupa artikel islami yang didapat dari website `republika.co.id`.

2. Data uji yang digunakan hanya dalam bahasa indonesia, jika ada bahasa arab maka dihilangkan dan digunakan artinya.
3. Training Set yang digunakan hanya terjemahan Al-Qur'annya saja.

1.4. Tujuan Penelitian

Berdasarkan latar belakang serta rumusan masalah di atas, maka tujuan yang hendak dicapai dalam penelitian ini sebagai berikut:

1. Mencari dalil menggunakan metode *cosine similarity* dan *Nave Bayes Classifier* pada artikel islami berdasarkan *indeks Al-Qur'an*.
2. Mengetahui akurasi metode *cosine similarity* dan *naive bayes classifier* dalam melakukan pencarian dalil.

1.5. Manfaat Penelitian

Adapun manfaat dari penelitian ini adalah:

1. Menambah wawasan dalam penerapan *text mining* dalam kehidupan sehari-hari.
2. Melihat kecocokan algoritma *naive bayes* untuk melakukan pencarian dalil dalam artikel islami.
3. Mempermudah dan mempercepat untuk mendapatkan rujukan pada ayat-ayat Al-Qur'an sesuai dengan indeks.
4. Mencari dalil artikel islami berdasarkan indeks *Al-Qur'an* secara otomatis.

1.6. Keaslian Penelitian

Penelitian mengenai klasifikasi pada *indeks Al-Qur'an* menggunakan metode *cosine similarity* dan *Naïve Bayes Classifier*, sudah pernah dilakukan oleh peneliti sebelumnya. Namun, berdasarkan referensi dan tinjauan pustaka, penelitian yang diajukan sebagai Tugas Akhir S1 pada Program Studi Teknik Informatika Fakultas Sains dan Teknologi UIN Sunan Kalijaga Yogyakarta mengenai implementasi *cosine similarity* dan *naive bayes classifier* untuk mencari dalil pada artikel islami berdasarkan *indeks Al-Qur'an* belum pernah ditemukan oleh peneliti lain.

1.7. Sistematika Penulisan

Untuk memberikan gambaran dan kerangka yang jelas mengenai pokok bahasan setiap bab dalam penelitian ini, maka diperlukan sistematika penulisan. Berikut ini adalah gambaran sistematika penulisan pada masing-masing bab :

BAB I PENDAHULUAN

Bab satu pendahuluan, berisikan latar belakang masalah, rumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian, keaslian penelitian, dan sistematika penulisan skripsi.

BAB II TINJAUAN PUSTAKA DAN LANDASAN TEORI

Bab dua tinjauan pustaka dan landasan teori, menjelaskan tentang tinjauan pustaka dan landasan teori yang berhubungan dengan topik yang akan di bahas dalam penelitian ini.

BAB III METODE PENELITIAN

Bab tiga metode penelitian, membahas mengenai metode yang dilakukan dan atau yang digunakan dalam menyelesaikan penelitian ini.

BAB IV HASIL DAN PEMBAHASAN

Bab empat hasil dan pembahasan, menjelaskan hasil dari penelitian yang dicapai. Dari proses pengambilan data hingga hasil yang diperoleh dari penelitian ini.

BAB V KESIMPULAN DAN SARAN

Bab lima penutup dan kesimpulan, menjelaskan kesimpulan yang dihasilkan serta saran yang akan diberikan berdasarkan hasil yang telah dicapai sehingga dapat digunakan sebagai bahan pertimbangan bagi pihak-pihak yang berkepentingan serta untuk perkembangan penelitian-penelitian selanjutnya.

BAB V

KESIMPULAN DAN SARAN

5.1. Kesimpulan

Berdasarkan tujuan penelitian yang telah dikemukakan, maka kesimpulan dalam penelitian ini adalah sebagai berikut:

1. Bahwa metode *Cosine Similarity* dan *Naive Bayes* berhasil diterapkan dalam sistem dengan baik dimana sistem dapat memberi *output* berupa ayat dan makna sesuai dengan yang diharapkan.
2. Hasil pengujian yang dilakukan sebanyak 4 kali mendapatkan hasil terbesar pada data uji sebanyak 60% dan data latih 40%, dengan nilai *accuracy* 38%, *precision* 9% dan *recall* 8%.

5.2. Saran

Pada penelitian ini tentunya tidak terlepas dari kekurangan dan kelemahan. Oleh karena itu, peneliti menyarankan adanya pengembangan penelitian berikutnya :

1. Diharapkan pada penelitian berikutnya dapat membandingkan dengan metode yang berbeda
2. Penelitian selanjutnya diharapkan dapat membuat sistem pencarian dalil berdasarkan indeks secara otomatis *AL-Qur'an*
3. Diharapkan pada penelitian selanjutnya dapat membaca ayat *Al-Qur'an*

DAFTAR PUSTAKA

- A., Widiastuti, N., S., Santosa, & C., Supriyanto. (2014). "Algoritma Klasifikasi Data Mining *Naïve Bayes Classifier Particle SWARM Optimization* Untuk Deteksi Penyakit Jantung. *Jurnal Pseudocode*, Vol. 1 No. 1, pp 11-14.
- Alfiani Aisha M. Saptono R. Anggrainingsih R. (2015). "Sistem Klasifikasi Feedback Pelanggan Dan Rekomendasi Solusi Atas Keluhan Di Upt Puskom Uns Dengan Algoritma Nave Bayes Classifier Dan Cosine Similarity ". *Jurnal ITSMART* vol 4. No 1.
- Aputro, Michael & Tyas. (2017). "Klasifikasi Lagu Daerah Indonesia Berdasarkan Lirik Menggunakan Metode TF-IDF dan *Naïve Baye*". *Jurnal Teknologi Informatika dan Terapan*, Vol. 04, No. 01
- B., Dwi Pramita, Saptono, Ristu, & A, Rini. (2018) "Academica Articles Classification Using *Naïve Bayes Classifier (NBC) Method*". *ITSMART: Jurnal Ilmiah Teknologi dan Informasi*, Vol. 7, No.2.
- Devi S, & dkk. (2016). "Klasifikasi Terjemahan Ayat Al-Qur'an Tentang Ilmu Sains Menggunakan Algoritma Decision Tree Berbasis Mobile ".
- Faizin, AS Wahid. (2018). "Implementasi *K-Means Clustering* Pada Terjemah *Al-Qur'an* Berdasarkan Keterkaitan Topik". *Teknik Informatika UIN Sunan Kalijaga Yogyakarta*.
- Ferdinandus, Subari. (2015). "Sistem Information Retrieval Layanan Kesehatan Untuk Berobat dengan Metode Vector Space Model berbasis WebGis". *Jurnal. Teknik Informatika. STIKI Malang*.
- Herwijayanti, Bening , Ratnawati, Dian Eka , dan Muflikhah Lailil. (2018). "Klasifikasi Berita *Online* dengan menggunakan Pembobotan *TF-IDF* dan *Cosine Si-*

- milarity*". Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer, Vol.2, No. 1, pp 306-312.
- Jananto A, (2013). "Algoritma Naive Bayes untuk Mencari Perkiraan Waktu Studi Mahasiswa ". Jurnal Teknologi Informasi DINAMIK Volume 18, No.
- Maarif, Abdul Aziz. (2015) "Penerapan Algoritma *TF-IDF* Untuk Pencarian Karya Ilmiah". Universitas Dian Nuswantoro Semarang.
- Manning, Christopher D., Prabhakar Raghavan, and Hinrich Schütze. 2008. "Introduction to Information Retrieval ". United Kingdom: Cambridge University Press.
- Mustaqhfi, M., Abidin, Z., dan Kusumawati, R. (2011). "Peringkasan Teks Otomatis Berita Berbahasa Indonesia Menggunakan *Metode Maximum Marginal Relevance*". Jurusan Teknik Informatika, Fakultas Sains dan Teknologi, UIN MALIKI.
- M. Syukri Mustafa, Muh Rizky Ramadhan, Angelina P. Thenata (2017). "Implementasi Data Mining untuk Evaluasi Kinerja Akademik Mahasiswa Menggunakan Algoritma Naive Bayes Classifier". STMIK Dipenagara, Makassar.
- Nurdiana O, Jumadi, & Nursantika D, (2016). "Perbandingan Metode *Cosine Similarity* Dengan Metode *Jaccard Similarity* Pada Aplikasi Pencarian Terjemah Al-Qur'an Dalam Bahasa Indonesia
- P. Rizki Wijayatun dan N., Yusuf Sulistyono. (2016). "Prediksi *Rating* Film Menggunakan *Metode Nave Bayes Classifier*". Jurnal Teknik Elektro Vol. 8, No. 2.
- R., Feldman & J, Sanger. (2007). "The text mining handbook: advanced approaches in analyzing unstructured data". Cambridge University Press : New York.
- Rozas, I. R. dan Sarno, R. (2011). "Sistem Pemilihan Kontrol Keamanan Informasi Berbasis ISO 27001 ". Seminar Nasional Pascasarjana XI. IT. Surabaya.

- Sevsa, Bama A. (2018). "Analisis Sentimen Komentar Pada Indeks Kinerja Dosen Fakultas Sains dan Teknologi UIN Sunan Kalijaga Yogyakarta menggunakan *Naive Bayes Classifier*". Yogyakarta: UIN Sunan Kalijaga.
- Sugiyanto, Surarso B, Sugiharto A, (2014). "Analisa Performa Metode Cosine Similarity dan Jacard Pada Pengujian Kesamaan Dokumen ". Vol.5 No 10, Universitas Diponegoro.
- Suyanto, 2018. "*Machine Learning* Tingkat Dasar dan Lanjut). Bandung: Informatika.
- Suryana, Asep (2018). "Implementasi *Cosine Simlirity* Dan *Naive Bayes* Pada Pengklasifikasian Hadits Shahih Bukhari Berdasarkan Anjuran Dan Larangan". Bandung, Uin Sunan Gunung Djati.
- Vinodhini, G., and Chandrasekaran, M. (2012). "*Sentiment Analysis and Opinion Mining: A Survey*. International Journal Of Advanced Research in Computer Science and Software Engineering, Vol.2, NO. 2.
- Wahid, Devid Haryalesmana dan SN, Azhari. (2016) "Peringkasan Sentimen Es-traktif di *Twitter* Menggunakan *Hybrid TF-IDF* dan *Cosine Similarity*". IJCCS, Vol. 10, No. 2, pp 207-208.
- W. Rizki Tri, P. Dhidik, & S., Eko. (2017) "Penerapan Algoritma Cosine Similarity dan Pembobotan TF-IDF pada Sistem Klasifikasi Dokumen Skripsi". Jurnal Teknik Elektro, Vol. 9, No. 1.
- Zaman, B. & Winarko, E. (2011). "Analisa Fitur Kalimat Peringkas Teks Otomatis pada Bahasa Indonesia". IJCCS, Vol.5 No.2.