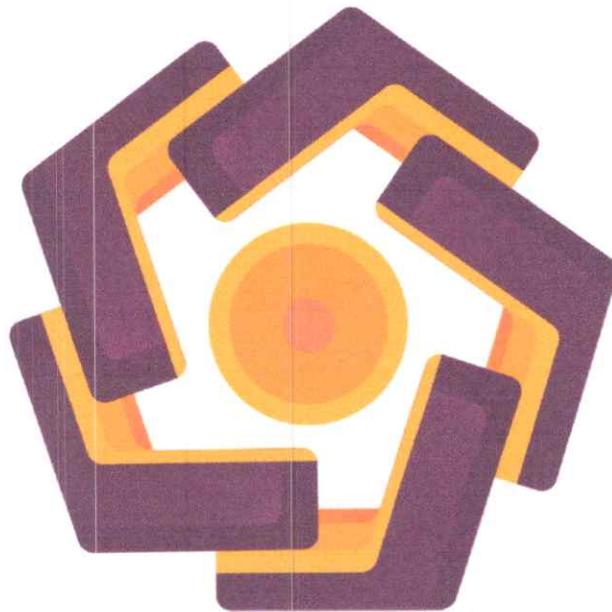


**KLASIFIKASI TEMA PENELITIAN MENGGUNAKAN ALGORITMA  
NAÏVE BAYES CLASSIFIER**

**SKRIPSI**

untuk memenuhi sebagian persyaratan  
mencapai gelar Sarjana  
pada Program Studi Informatika



disusun oleh  
**Refianto Yusuf Afandi**

**15.11.8867**

**PROGRAM SARJANA  
PROGRAM STUDI INFORMATIKA  
FAKULTAS ILMU KOMPUTER  
UNIVERSITAS AMIKOM YOGYAKARTA  
YOGYAKARTA  
2019**

**PERSETUJUAN**

**SKRIPSI**

**KLASIFIKASI TEMA PENELITIAN MENGGUNAKAN ALGORITMA  
NAÏVE BAYES CLASSIFIER**

yang dipersiapkan dan disusun oleh

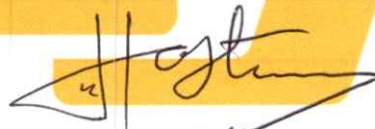
**Refianto Yusuf Afandi**

**15.11.8867**

telah disetujui oleh Dosen Pembimbing Skripsi

pada tanggal 7 Januari 2019

**Dosen Pembimbing,**



**Hastari Utama, M.Cs.**

**NIK. 190302230**

## PENGESAHAN

### SKRIPSI

#### KLASIFIKASI TEMA PENELITIAN MENGGUNAKAN ALGORITMA NAÏVE BAYES CLASSIFIER

yang dipersiapkan dan disusun oleh

**Refianto Yusuf Afandi**

15.11.8867

telah dipertahankan di depan Dewan Penguji  
pada tanggal 10 Februari 2019

#### Susunan Dewan Penguji

**Nama Penguji**

**Tanda Tangan**

Hastari Utama, M.C.  
NIK. 190302230

Ferry Wahyu Wibowo, S.Si, M.Cs.  
NIK. 190302235

Agus Fatkhurohman, M.Kom.  
NIK. 190302249



Skripsi ini telah diterima sebagai salah satu persyaratan  
untuk memperoleh gelar Sarjana Komputer  
Tanggal 15Februari 2018

**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 tanggungjawab saya pribadi.

Yogyakarta, 6 Febuari 2019



Refianto Yusuf Afandi

NIM. 15.11.8867

## MOTTO

Boleh jadi kamu membenci sesuatu, padahal ia amat baik bagi kamu. Dan boleh jadi kamu mencintai sesuatu, padahal ia amat buruk bagi kamu. Allah Maha mengetahui sedangkan kamu tidak mengetahui” (Al-Baqarah: 216)

Setiap lelaki adalah pemimpin dan setiap pemimpin bertanggung jawab atas umatnya.

Jika engka ingin mencari kesuskesan maka tempuhlah jalan kesuksesan, sesungguhnya kapal tidak berlayar di atas daratan.



## PERSEMBAHAN

Saya mempersembahkan skripsi ini kepada semua pihak yang terlibat secara langsung maupun tidak langsung dalam proses pembuatan skripsi.

1. Allah subhanahu wa ta'ala yang telah memberi limpahan ramad dan nikmatnya sampai saat ini.
2. Kemudian kepada kedua orang tua saya yang telah mendo'akan dan mendukung saya.
3. Dosen pembimbing Bapak Hastari Utama, M.Cs yang telah memberi masukan dan membimbing saya samapai ahir proses pembuatan sekripsi.
4. Dosen-dosen Amikom yang telah memberi ilmu yang bermanfaat.
5. Terimakasih kepada teman-teman IKPPIM 15 Yogyakarta yang telah memberi dorongan untuk menyelesaikan tugas ahir ini.
6. Terimakasih kepada teman-teman amikom khususnya kelas Informatika-06 yang telah memberi masukan terkait dengan sekripsi.
7. Terimakasih kepada Bento Coffee yang telah memfasilitasi tempat untuk mengerjakan sekripsi.
8. Terimakasih kepada seluruh pihak yang telah membantu secara langsung dan tidak langsung.

Semoga Allah Subhanahu wata'ala memberikan balasan yang lebih kepada semua yang telah ikut membantu saya dan menyelesaikan skripsi ini. Demi perbaikan selanjutnya, saran dan kritik yang membangun akan diterima dengan senang hati dan rasa terima kasih. Semoga skripsi ini dapat bermanfaat bagi saya dan kita semua.

## KATA PENGANTAR

Puja dan puji syukur saya panjatkan kepada Tuhan yang maha Esa yang telah memberikan kita limpahan rahmat, nikmat, serta hidayahnya sehingga saya dapat menyelesaikan tugas ahir ini dengan judul klasifikasi tema penelitian menggunakan algoritma Naïve Bayes Clasiifier.

Skripsi ini saya buat guna menyelesaikan studi jenjang Strata Satu (S1) pada program studi Teknik Informatika fakultas Ilmu Komputer Universitas AMIKOM Yogyakarta. Selain itu juga merupakan suatu bukti bahwa mahasiswa telah menyelesaikan kuliah jenjang program strata 1 dan untuk memperoleh gelar Sarjana Komputer .

Semoga Allah Subhanahu wata'ala memberikan balasan yang lebih kepada semua yang telah ikut membantu saya dan menyelsaikan skripsi ini. Demi perbaikan selanjutnya, saran dan kritik yang membangun akan diterima dengan senang hati dan rasa terima kasih. Semoga skripsi ini dapat bermanfaat bagi saya dan kita semua.

Yogyakarta, 6 Febuari 2019

  
Refianto Yusuf Afandi  
NIM. 15.11.8867

## DAFTAR ISI

LEMBAR JUDUL .....	i
PERSETUJUAN .....	ii
PENGESAHAN .....	iii
PERNYATAAN.....	iv
MOTTO.....	v
PERSEMBAHAN .....	vi
KATAPENGANTAR .....	vii
DAFTAR ISI.....	viii
DAFTAR TABEL.....	xi
DAFTAR GAMBAR.....	xiii
INTISARI.....	xv
<i>ABSTRACT</i> .....	xvi
<b>BAB I PENDAHULUAN .....</b>	<b>1</b>
1.1. Latar Belakang Masalah .....	1
1.2. Rumusan Masalah.....	2
1.3. Batasan Masalah .....	2
1.4. Maksud Dan Tujuan Penelitian .....	2
1.5. Metode Penelitian .....	3
1. Metode Pengumpulan Data.....	3
2. Metode Analisis .....	3
1.6. Sistematika Penulisan .....	4
<b>BAB II LANDASAN TEORI .....</b>	<b>6</b>



2.1. Tinjauan Pustaka .....	6
2.2. Dasar Teori .....	10
1. Text Mining .....	10
2. Tahapan Text Mining.....	11
3. Naïve Bayes Classifier .....	13
2.3. Perancangan UML .....	14
1. Use Case Diagram.....	14
2. Sequence Diagram .....	16
3. Activity Diagram .....	17
2.4. PHP .....	18
2.5. MySQL .....	18
2.6. XAMPP .....	19
2.7. Python .....	19
<b>BAB III ANALISI DAN PERANCANGAN PENELITIAN .....</b>	<b>20</b>
3.1. Identifikasi Masalah .....	20
3.2. Analisis Masalah.....	20
3.3. Hasil Analisa .....	20
3.4. Deskripsi Sistem .....	21
3.5. Perhitungan Manual Naïve Bayes Classifier .....	22
3.6. Analisa Kebutuhan .....	34
1. Kebutuhan Fungsional .....	34
2. Kebutuhan Non-Fungsional .....	35
3.7. Perancangan Perangkat Lunak.....	36
1. Use Case.....	36
2. Activity Diagram .....	42
3. Class Diagram .....	45
4. Sequence Diagram .....	46
5. User Interface .....	49
<b>BAB IV HASIL DAN PEMBAHASAN .....</b>	<b>51</b>

4.1. Create Database dan Tabel .....	51
4.2. Interface Pengguna .....	53
1. Admin .....	53
2. User .....	56
4.3. Pembahasan Source Code .....	58
3. Pre-Processing .....	58
4. TF-IDF .....	60
5. Naïve Bayes Calssifier .....	61
4.4. Pengujian Sistem .....	63
<b>BAB V PENUTUP</b> .....	<b>66</b>
5.1. Kesimpulan .....	66
5.2. Saran .....	66
<b>DAFTAR PUSTAKA</b> .....	<b>67</b>

## DAFTAR TABEL

Tabel 2.1 Perbedaan Penelitian .....	7
Tabel 2.2 Simbol Use Case .....	15
Tabel 2.3 Simbol Sequence Diagram.....	16
Tabel 2.4 Simbol Activity Diagram.....	17
Tabel 3.1 Data Training .....	22
Tabel 3.2 Case Fold.....	23
Tabel 3.3 Filtering.....	24
Tabel 3.4 Stop Word Removal .....	25
Tabel 3.5 Stemming .....	26
Tabel 3.6 Perhitungan TF-IDF.....	27
Tabel 3.7 Data Uji.....	31
Tabel 3.8 Perhitungan Naïve Bayes Classifier.....	33
Tabel 3.9 Hasil Perhitungan Naïve Bayes Classifier .....	33
Tabel 3.10 Definisi Fitur Utama .....	34
Tabel 3.11 Definisi Fitur Lanjut.....	34
Tabel 3.12 Kebutuhan Fungsional Aktor.....	35
Tabel 3.13 Kebutuhan Perangkat Keras.....	35
Tabel 3.14 Kebutuhan Perangkat Lunak .....	36
Tabel 3.15 Proses Klasifikasi .....	37
Tabel 3.16 Proses Perhitungan Naïve Bayes Classifier .....	38
Tabel 3.17 Proses Hasil Tema Penelitian .....	38

Tabel 3.18 Proses Input Data .....	39
Tabel 3.19 Proses Lihat Data .....	39
Tabel 3.20 Proses Edit Data .....	40
Tabel 3.21 Proses Delete Data .....	40
Tabel 3.22 Proses Pre-Processing .....	41
Tabel 3.23 Proses TF-IDF .....	42
Tabel 4.1 Perbandingan Algoritma .....	63
Tabel 4.2 Confusion Matrix Naïve Bayes .....	64
Tabel 4.3 Confusion Matrix Trees J48 .....	64
Tabel 4.4 Confusion Matrix LibSVM Linier .....	64
Tabel 4.5 Confusion Matrix LibSVM Polynominal .....	64
Tabel 4.6 Confusion Matrix LibSVM Radial .....	65

## DAFTAR GAMBAR

Gambar 3.1 Proses Perhitungan .....	21
Gambar 3.2 Usecase .....	36
Gambar 3.3 Proses Activity Diagram Klasifikasi data Uji .....	42
Gambar 3.4 Proses Activity Diagram Input Data Training.....	43
Gambar 3.5 Proses Activity Diagram Lihat Data .....	43
Gambar 3.6 Proses Activity Diagram Edit Data Training .....	44
Gambar 3.7 Proses Activity Diagram Delete Data Training.....	45
Gambar 3.8 Class Diagram .....	46
Gambar 3.9 Proses Sequence Diagram Klasifikasi Data Uji.....	46
Gambar 3.10 Proses Sequence Diagram Input Data Training .....	47
Gambar 3.11 Proses Sequence Diagram Lihat Data Training .....	47
Gambar 3.12 Proses Sequence Diagram Edit Data Training .....	48
Gambar 3.13 Proses Sequence Diagram Delete Data Training .....	48
Gambar 3.14 User Interface Submit Data Uji.....	49
Gambar 3.15 User Interface Hasil Klasifikasi .....	49
Gambar 3.16 User Interface Admin.....	50
Gambar 4.1 Buat Tabel Multimedia.....	51
Gambar 4.2 Setting Id Multimedia.....	51
Gambar 4.3 Buat Tabel HS .....	51
Gambar 4.4 Setting Id HS .....	52



Gambar 4.5 Buat Tabel AI .....	52
Gambar 4.6 Setting Id AI .....	52
Gambar 4.7 Buat Tabel Network .....	52
Gambar 4.8 Setting Id Network .....	52
Gambar 4.9 Buat Tabel TF-IDF .....	53
Gambar 4.10 Buat Tabel Val .....	53
Gambar 4.11 Tampilan Dashboard Admin .....	54
Gambar 4.12 Tampilan Menu Abstract Admin.....	54
Gambar 4.13 Tampilan Data Abstract Admin .....	55
Gambar 4.14 Tampilan Edit Abstract Admin .....	55
Gambar 4.15 Tampilan Hapus Abstract Admin.....	55
Gambar 4.16 Tampilan Menu TF-IDF.....	56
Gambar 4.17 Tampilan Menu Input Abstract Admin .....	56
Gambar 4.18 Tampilan About User .....	57
Gambar 4.19 Tampilan Hitung Abstract User .....	57
Gambar 4.20 Tampilan Hasil Perhitungan Abstract User.....	58
Gambar 4.21 Source Code Case Fold .....	58
Gambar 4.22 Source Code Filtering.....	59
Gambar 4.23 Source Code Stop Word .....	59
Gambar 4.24 Source Code Stemming .....	60
Gambar 4.25 Source Code TF-IDF .....	61
Gambar 4.26 Source Code Naïve Bayes Classifier.....	63

## INTISARI

Saat ini diperkirakan sekitar 80% teks dokumen berbentuk tak teratur, tingginya ketidak teraturan teks di picu berbagai aktifitas seperti jurnal online, aktifitas akademi, berita online dsb. Dokumen yang tak teratur mengakibatkan sulitnya memilah dokumen yang sesuai, sehingga di perlukan sebuah sistem untuk menata dokumen tersebut.

Naïve bayes adalah algoritma yang dapat mengklasifikasikan dokumen berdasarkan temanya, dasar perhitungan Naïve Bayes adalah probabilitas yang di hitung berdasarkan frekuensi kemunculan kata, sehingga cocok di kombinasikan dengan TF-IDF.

Hasil yang di harapkan adalah dengan di buatnya aplikasi berbasis Naïve Bayes dapat membantu menata ulang document yang awalnya tak teratur menjadi lebih teratur, agar bisa mempermudah peneliti sesudahnya untuk mencari jurnal sesuai dengan kategorinya.

**Kata Kunci :** Algoritma Naïve Bayes Classifier, klasifikasi teks, kecerdasan buatan

## ABSTRACT

*Currently it is estimated about 80% of the text document, irregular-shaped high trigger text manifest irregularity of a variety of activities such as online journals, the Academy's activities, online news, etc. Documents that are not regularly result in difficulty of sorting out the appropriate documents, and so on need a system to organize the document.*

*Naïve Bayes algorithm which is able to classify documents based on the theme, the basic calculation of Naïve Bayes probability is calculated based on the frequency of occurrence of the word, so it fits in the combine with TF-IDF.*

*Expect results is with in her Naïve Bayes based applications can help reorganized the document initially fruitless became more regularly, in order to facilitate researchers afterward to find the journal in accordance with the category.*

**Keywords:** *Naïve Bayes Classifier Algorithms, text classification, artificial intelligence*