



**KOMPARASI ALGORITMA *SUPPORT VERTOR MACHINE*  
DAN *NAÏVE BAYES* UNTUK DETEKSI JUDUL BERITA  
*CLICKBAIT* INDONESIA**

**LAPORAN SKRIPSI**



**SARAH VINA FADHILAH**

**41519010086**

**PROGRAM STUDI TEKNIK INFORMATIKA**

**FAKULTAS ILMU KOMPUTER**

**UNIVERSITAS MERCU BUANA JAKARTA**

**2023**



**KOMPARASI ALGORITMA *SUPPORT VERTOR MACHINE*  
DAN *NAÏVE BAYES* UNTUK DETEKSI JUDUL BERITA  
*CLICKBAIT* INDONESIA**

**LAPORAN SKRIPSI**



**SARAH VINA FADHILAH**

**41519010086**

**Diajukan sebagai salah satu syarat untuk memperoleh gelar sarjana**

**PROGRAM STUDI TEKNIK INFORMATIKA**

**FAKULTAS ILMU KOMPUTER**

**UNIVERSITAS MERCU BUANA JAKARTA**

**2023**

## HALAMAN PERNYATAAN KARYA SENDIRI

### HALAMAN PERNYATAAN KARYA SENDIRI

Saya yang bertanda tangan di bawah ini:

Nama : Sarah Vina Fadhilah  
NTM : 41519010086  
Program Studi : Teknik Informatika  
Judul Laporan Skripsi : Komparansi Algoritma *Support Vector Machine* Dan *Naïve Bayes* Untuk Deteksi Judul Berita *Clickbait* Indonesia

Menyatakan bahwa Laporan Skripsi ini adalah hasil karya saya sendiri dan bukan plagiat, serta semua sumber baik yang dikutip maupun dirujuk telah saya nyatakan dengan benar. Apabila ternyata ditemukan di dalam Laporan Skripsi saya terdapat unsur plagiat, maka saya siap mendapatkan sanksi akademis yang berlaku di Universitas Mercu Buana.

Jakarta, 01 Agustus 2023



UNIVERSITA  
MERCU BUANA

METER  
TEMPEL  
SARAH VINA FADHILAH

## HALAMAN PENGESAHAN

### HALAMAN PENGESAHAN

Laporan Skripsi ini diajukan oleh:

Nama : Sarah Vina Fadhilah  
NIM : 41519010086  
Program Studi : Teknik Informatika  
Judul Laporan Skripsi : *Komparansi Algoritma Support Vector  
Machine Dan Naive Bayes Untuk Deteksi Judul Berita  
Clickbait Indonesia*

Telah berhasil dipertahankan pada sidang di hadapan Dewan Penguji dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh gelar Sarjana Strata I pada Program Studi Teknik Informatika, Fakultas Ilmu Komputer Universitas Mercu Buana.

Disahkan oleh:

Pembimbing : Dr. Afiyati, S.Si., MT  
NIDN : 114690444  
Ketua Penguji : Dr. Ruci Meiyanti, S.Kom., M. Kom  
NIDN : 0304056803  
Penguji I : Yunita Sartika Sari, S.Kom., M.Kom  
NIDN : 0309068903



Jakarta, 01 Agustus 2023

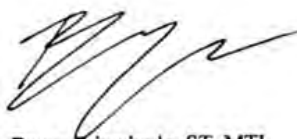
Mengetahui,

Dekan

Ketua Program Studi



Dr. Bambang Jokonowo, S.Si. MTI



Bagus Priambodo, ST, MTI

## KATA PENGANTAR

Puji syukur saya panjatkan kepada Tuhan Yang Maha Esa, karena atas berkat dan rahmat-Nya, saya dapat menyelesaikan Laporan Skripsi ini. Penulisan Laporan Skripsi ini dilakukan dalam rangka memenuhi salah satu syarat untuk mencapai gelar Sarjana Komputer pada Fakultas Ilmu Komputer Universitas Mercu Buana. Saya menyadari bahwa tanpa bantuan dan bimbingan dari berbagai pihak, dari masa perkuliahan sampai pada penyusunan skripsi ini, sangatlah sulit bagi saya untuk menyelesaikan Laporan Skripsi ini. Oleh karena itu saya mengucapkan terima kasih kepada:

1. Prof. Dr. Ir. Andi Adriansyah, M. Eng selaku Rektor Universitas Mercu Buana.
2. Dr. Bambang Jokonowo, S.Si., MTI selaku Dekan Fakultas Ilmu Komputer.
3. Bagus Priambodo, ST., MTI selaku Ketua Program Studi Teknik Informatika.
4. Dr. Afiyati, S.Si., MTI selaku Dosen Pembimbing yang telah menyediakan waktu, tenaga, dan pikiran untuk mengarahkan saya dalam penyusunan skripsi ini.
5. Kedua Orangtua yang selalu memberikan do'a dan dukungan selama perkuliahan.
6. Oktalia Kumala Sari selaku rekan seperjuangan yang selalu mendukung dan membantu saya selama masa perkuliahan.
7. Serta pihak-pihak yang telah berjasa dan memotivasi saya, yang namanya tidak dapat disebutkan satu persatu namun tanpa mengurangi rasa terima kasih sedikit pun dari saya.

Akhir kata, saya berharap Tuhan Yang Maha Esa berkenan membalas segala kebaikan semua pihak yang telah membantu. Semoga Laporan Skripsi ini membawa manfaat bagi pengembangan ilmu.

Jakarta, 30 Juni 2023

Sarah Vina Fadhilah

## HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR UNTUK KEPENTINGAN AKADEMIS

### HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR UNTUK KEPENTINGAN AKADEMIS

Sebagai sivitas akademik Universitas Mercu Buana, saya yang bertanda tangan di bawah ini:

Nama : Sarah Vina Fadhillah  
NIM : 41519010086  
Program Studi : Teknik Informatika  
Judul Laporan Skripsi : Komparansi Algoritma *Support Vector Machine* Dan *Naive Bayes* Untuk Deteksi Judul Berita *Clickbait* Indonesia

Demi pengembangan ilmu pengetahuan, dengan ini memberikan izin dan menyetujui untuk memberikan kepada Universitas Mercu Buana **Hak Bebas Royalti Non-Eksklusif (Non-exclusive Royalti-Free Right)** atas karya ilmiah saya yang berjudul di atas beserta perangkat yang ada (jika diperlukan).

Dengan Hak Bebas Royalti Non-Eksklusif ini Universitas Mercu Buana berhak menyimpan, mengalih media/format-kan, mengelola dalam bentuk pangkalan data (*database*), merawat, dan mempublikasikan Laporan Magang/Skripsi//Tesis/Disertasi saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

Jakarta, 01 Agustus 2023

Yang menyatakan,



(Sarah Vina Fadhillah)

## ABSTRAK

Nama : Sarah Vina Fadhilah  
NIM : 41519010086  
Program Studi : Teknik Informatika  
Judul Laporan Skripsi : Komparasi Algoritma *Support Vector Machine* Dan *Naïve Bayes* Untuk Deteksi Judul Berita *Clickbait* Indonesia  
Pembimbing : Dr. Afiyati, SSi., MT

Banyak aktivitas yang tidak dapat lepas dari yang namanya internet, termasuk membaca berita *online*. Di era sekarang ini, berita *online* dapat dikatakan telah menjadi gaya hidup masyarakat dan mengubah cara masyarakat dalam memperoleh informasi terkini. Semakin maraknya berita *online*, para pembuat berita menjadi saling bersaing dalam hal meningkatkan peringkat masing-masing *platform*. Hal itu memunculkan timbulnya judul berita *clickbait* yaitu judul berita yang dilebih-lebihkan dan tak sesuai dengan isi beritanya. Maka tujuan dari penelitian ini adalah untuk mendeteksi judul berita *clickbait* agar masyarakat tidak terhasut oleh judul berita yang dilebih-lebihkan serta dapat terhindar dan tidak terpengaruh dari paparan berita bohong. Adapun metode yang dilakukan dalam penelitian ini yaitu menggunakan algoritma *Support Vector Machine (SVM)* dan algoritma *Naïve Bayes* untuk perbandingan. Hasil dari penelitian ini memperoleh akurasi 71%, presisi 72%, *recall* 58%, dan *f1 score* 64% dari penggunaan model *SVM* dan akurasi 72%, presisi 73%, *recall* 59%, dan *f1 score* 65% dari penggunaan model *Naïve Bayes*.

**Kata Kunci :** *naïve bayes, SVM, judul berita, clickbait*

## **ABSTRACT**

*Name* : Sarah Vina Fadhillah  
*NIM* : 41519010086  
*Study Program* : Informatics Engineering  
*Title Thesis* : Comparing Support Vector Machine and Naïve Bayes Algorithm for Clickbait News Headline Detection Indonesia  
*Counsellor* : Dr. Afiyati, SSi., MT

*Many activities can't be separated from the internet, including reading online news. In this era, online news has become a way of life and changed the way people obtain the latest information. With the rise of online news, news makers have become competitive with each other in terms of increasing the ranking of each platform. This has given rise to the emergence of "clickbait" headlines, which are overstated headlines that do not match the content of the news. So, the purpose of this research is to detect clickbait headlines so that people are not instigated by overstated headlines and can avoid and not be affected by exposure to fake news. The methods used in this study are using the Support Vector Machine (SVM) algorithm and the Naïve Bayes algorithm for comparison. The results of this study showed 71% accuracy, 72% precision, 58% recall, and 64% f1 score from using the SVM model and 72% accuracy, 73% precision, 59% recall, and 65% f1 score from using the Naïve Bayes model.*

**Keywords** : naïve bayes, SVM, news headline, clickbait



## DAFTAR ISI

<b>HALAMAN JUDUL</b> .....	<b>i</b>
<b>HALAMAN PERNYATAAN KARYA SENDIRI</b> .....	<b>ii</b>
<b>HALAMAN PENGESAHAN</b> .....	<b>iii</b>
<b>KATA PENGANTAR</b> .....	<b>iv</b>
<b>HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR UNTUK KEPENTINGAN AKADEMIS</b> .....	<b>v</b>
<b>ABSTRAK</b> .....	<b>vi</b>
<b>ABSTRACT</b> .....	<b>vii</b>
<b>DAFTAR ISI</b> .....	<b>viii</b>
<b>DAFTAR TABEL</b> .....	<b>x</b>
<b>DAFTAR GAMBAR</b> .....	<b>xi</b>
<b>DAFTAR LAMPIRAN</b> .....	<b>xii</b>
<b>BAB 1 PENDAHULUAN</b> .....	<b>1</b>
1.1. Latar Belakang .....	1
1.2. Perumusan Masalah.....	2
1.3. Tujuan Penelitian.....	2
1.4. Manfaat Penelitian.....	3
1.5. Batasan Penelitian .....	3
<b>BAB II TINJAUAN PUSTAKA</b> .....	<b>4</b>
2.1. Penelitian Terdahulu .....	4
2.2. Teori Pendukung.....	13
2.2.1. Berita <i>Online</i> .....	13
2.2.2. <i>Clickbait</i> .....	14
2.2.3. <i>Support Vector Machine</i> .....	14
2.2.4. Naïve Bayes .....	15
2.2.5. NLTK (Natural Language Toolkit) dan Sastrawi .....	16
<b>BAB III METODE PENELITIAN</b> .....	<b>18</b>
3.1. Jenis Penelitian .....	18
3.2. Tahap Penelitian .....	18
3.2.1. Persiapan Perencanaan .....	19
3.2.2. Pemilihan <i>Dataset</i> .....	19

3.2.3. Pre-processing.....	19
3.2.4. Split Data.....	20
3.2.5. Pemodelan.....	21
3.2.6. Evaluasi.....	21
<b>BAB IV HASIL DAN PEMBAHASAN.....</b>	<b>22</b>
4.1. Dataset.....	22
4.2. Pre-Processing.....	23
4.3. Pembuatan Model.....	26
4.4. Visualisasi Data.....	28
4.5. Pengujian.....	30
4.6. Analisis Hasil.....	32
<b>BAB V KESIMPULAN DAN SARAN.....</b>	<b>35</b>
5.1. Kesimpulan.....	35
5.2. Saran.....	35
<b>DAFTAR PUSTAKA.....</b>	<b>36</b>
<b>LAMPIRAN.....</b>	<b>40</b>
Lampiran 1. Asistensi Bimbingan.....	40
Lampiran 2. Lampiran Persetujuan.....	41
Lampiran 3. Lampiran Luaran Tugas Akhir.....	42
Lampiran 4. Bukti Submit Jurnal.....	43
Lampiran 5. Naskah Jurnal yang disubmit.....	44
Lampiran 6. <i>Curriculum Vitae (CV)</i> .....	52
Lampiran 7. Lampiran Surat Pernyataan HKI (jika belum <i>published</i> HKI) ....	54
Lampiran 8. Sertifikat BNSP.....	55
Lampiran 9. <i>Plagiarisme Check</i> .....	57

## DAFTAR TABEL

Tabel 2. 1 Penelitian terdahulu.....	4
Tabel 4. 1 Contoh kata dalam kamus slangwords .....	24
Tabel 4. 2 <i>Confusion Matrix SVM</i> .....	32
Tabel 4. 3 <i>Confusion Matrix Naïve Bayes</i> .....	32



## DAFTAR GAMBAR

Gambar 3. 1 Tahap Penelitian .....	18
Gambar 4. 1 Contoh <i>dataset</i> yang digunakan .....	22
Gambar 4. 2 Diagram Perbandingan <i>Clickbait</i> dan <i>Non-Clickbait</i> .....	23
Gambar 4. 3 Sebelum Tahap <i>Data Cleaning</i> .....	24
Gambar 4. 4 Sesudah Tahap <i>Data Cleaning</i> .....	24
Gambar 4. 5 Grafik perbandingan sebelum dan sesudah normalisasi .....	25
Gambar 4. 6 Contoh hasil tahap <i>case folding</i> .....	25
Gambar 4. 7 <i>Source code stopwords removal</i> .....	26
Gambar 4. 8 Tokenisasi judul berita.....	26
Gambar 4. 9 Pembuatan model <i>SVM</i> .....	27
Gambar 4. 10 Contoh prediksi <i>SVM</i> .....	27
Gambar 4. 11 Pembuatan model <i>Naïve Bayes</i> .....	28
Gambar 4. 12 Contoh prediksi <i>Naïve Bayes</i> .....	28
Gambar 4. 13 <i>Wordcloud</i> data <i>clickbait</i> .....	29
Gambar 4. 14 <i>Wordcloud</i> data <i>non-clickbait</i> .....	29
Gambar 4. 15 Grafik frekuensi kata <i>clickbait</i> .....	30
Gambar 4. 16 Grafik frekuensi kata <i>non-clickbait</i> .....	30
Gambar 4. 17 <i>Confusion matrix SVM</i> .....	31
Gambar 4. 18 <i>Confusion matrix Naïve Bayes</i> .....	31
Gambar 4. 19 Uji coba deteksi judul berita .....	34

## DAFTAR LAMPIRAN

Lampiran 1. Asistensi Bimbingan .....	40
Lampiran 2. Lampiran Persetujuan .....	41
Lampiran 3. Lampiran Luaran Tugas Akhir.....	42
Lampiran 4. Bukti Submit Jurnal .....	43
Lampiran 5. Naskah Jurnal yang disubmit.....	44
Lampiran 6. <i>Curriculum Vitae (CV)</i> .....	52
Lampiran 7. Lampiran Surat Pernyataan HKI (jika belum <i>published</i> HKI) .....	54
Lampiran 8. Sertifikat BNSP .....	55
Lampiran 9. <i>Plagiarisme Check</i> .....	57

