



**KLASIFIKASI DAN ANALISIS SENTIMEN PADA DATA  
TWITTER MENGGUNAKAN ALGORITMA *NAÏVE BAYES*  
(STUDI KASUS: TIMNAS SEPAKBOLA INDONESIA SENIOR,  
U-23, DAN U-19)**

**SKRIPSI**

**REINO PRAJAMUKTI**

**1710511079**

**UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN JAKARTA  
FAKULTAS ILMU KOMPUTER  
PROGRAM STUDI INFORMATIKA  
2021**



**KLASIFIKASI DAN ANALISIS SENTIMEN PADA DATA  
TWITTER MENGGUNAKAN ALGORITMA *NAÏVE BAYES*  
(STUDI KASUS: TIMNAS SEPAKBOLA INDONESIA SENIOR,  
U-23, DAN U-19)**

**SKRIPSI**

**Diajukan Sebagai Salah Satu Syarat untuk Memperoleh Gelar  
Sarjana Komputer**

**Reino Prajamukti**

**1710511079**

**UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN JAKARTA  
FAKULTAS ILMU KOMPUTER  
PROGRAM STUDI INFORMATIKA  
2021**

## PERNYATAAN ORISINALITAS

### PERNYATAAN ORISINALITAS

Skripsi ini adalah karya sendiri dan semua sumber yang dikutip maupun yang dirujuk telah saya nyatakan dengan benar.

Nama : Reino Prajamukti

NIM : 1710511079

Tanggal : 26 Juli 2021

Bilamana dikemudian hari ditemukan tidaksesuaian dengan pernyataan saya ini, maka saya bersedia dituntut dan diproses sesuai dengan ketentuan yang berlaku.

Jakarta, 26 Juli 2021

Yang Menyatakan,



(Reino Prajamukti)

## **PERNYATAAN PERSETUJUAN PUBLIKASI SKRIPSI UNTUK KEPENTINGAN AKADEMIS**

---

Sebagai civitas akademik Universitas Pembangunan Nasional Veteran Jakarta, saya yang bertanda tangan dibawah ini:

Nama : Reino Prajamukti  
NIM : 1710511079  
Fakultas : Ilmu Komputer  
Program Studi : Informatika

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Pembangunan Nasional Veteran Jakarta Hak Bebas Royalti Non eksklusif (*Non-Exclusive Royalty Free Right*) atas karya ilmiah saya yang berjudul:

### **KLASIFIKASI DAN ANALISIS SENTIMEN PADA DATA TWITTER MENGUNAKAN ALGORITMA NAÏVE BAYES (STUDI KASUS: TIMNAS SEPAKBOLA INDONESIA SENIOR, U-23, DAN U-19)**

Beserta perangkat yang ada (jika diperlukan). Dengan Hak Bebas Royalti ini Universitas Pembangunan Nasional Veteran Jakarta berhak menyimpan, mengalih media/formatkan, mengelola dalam bentuk pengkalan data (*database*), merawat, dan mempublikasi Skripsi saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta. Demikian pernyataan ini saya buat dengan sebenarnya.

Dibuat di : Jakarta

Pada tanggal : 26 Juli 2021

Yang Menyatakan,



(Reino Prajamukti)

## PENGESAHAN

Skripsi diajukan oleh:

Nama : Reino Prajamukti

NIM : 1710512080

Program Studi : S-1 Informatika

Judul Skripsi : KLASIFIKASI DAN ANALISIS SENTIMEN PADA DATA  
TWITTER MENGGUNAKAN ALGORITMA *NAÏVE BAYES*  
(STUDI KASUS: TIMNAS SEPAKBOLA INDONESIA  
SENIOR, U-23, DAN U-19)

Telah berhasil dipertahankan di hadapan Tim Penguji dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh gelar Sarjana Komputer pada Program Studi S1 Informatika Fakultas Ilmu Komputer, Universitas Pembangunan Nasional Veteran Jakarta.



Dr. Ermatita, M.Kom.

Ketua Penguji



Nurul Chamidah, S.Kom., M.Kom.

Penguji



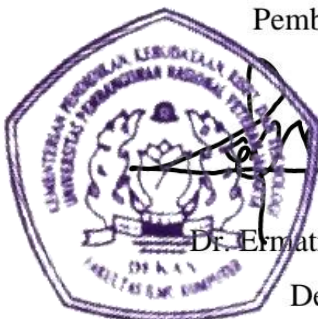
Jayanta, S.Kom., M.Si.

Pembimbing I



Mayanda Mega S., S.Kom, M.Kom.

Pembimbing II



Dr. Ermatita, M.Kom.

Dekan



Yuni Widiastiwi, S.Kom., M.Si.

Ketua Program Studi

Ditetapkan di : Jakarta

Tanggal Persetujuan : 26 Juli 2021



## **KATA PENGANTAR**

Segala puji dan syukur penulis panjatkan kehadiran Tuhan Yang Mahakuasa atas limpahan rahmat dan karunia-Nya sehingga penulis dapat menyelesaikan Proposal Tugas Akhir. Proposal tugas akhir ini disusun sebagai syarat Tugas Akhir Program Studi Informatika Fakultas Ilmu Komputer Universitas Pembangunan Nasional Veteran Jakarta.

Dalam penyelesaian karya tulis ilmiah ini tidak lepas dari bantuan banyak pihak yang telah memberikan masukan kepada penulis. Untuk itu penulis mengucapkan terima kasih kepada:

1. Dr. Ermatita, M.Kom., selaku dekan Fakultas Ilmu Komputer
2. Yuni Widiastiwi, S.Kom, M.Si., selaku Ketua Program Studi Sarjana Jurusan S1 Informatika.
3. Jayanta, S.Kom, M.Si., selaku dosen pembimbing 1 dari pihak jurusan.
4. Mayanda Mega Santoni, S.Kom, M.Kom., selaku dosen pembimbing 2 dari pihak jurusan.
5. Orang tua yang telah memberikan dukungan baik secara moril maupun materil.
6. Seluruh pihak yang terlibat dalam kelancaran pembuatan makalah karya ilmiah ini dan yang belum disebutkan di atas, penulis ucapkan terima kasih.

Penulis menyadari bahwa masih banyak kekurangan dari laporan ini, baik dari materi maupun teknik penyajiannya. Oleh karena itu, kritik dan saran yang membangun penulis harapkan.

Jakarta, 1 Juli 2021

Penulis

**KLASIFIKASI DAN ANALISIS SENTIMEN PADA DATA TWITTER  
MENGUNAKAN ALGORITMA NAÏVE BAYES  
(STUDI KASUS: TIMNAS SEPAKBOLA INDONESIA SENIOR, U-23,  
DAN U-19)**

**Reino Prajamukti**

**ABSTRAK**

Media sosial tidak hanya berfungsi sebagai media komunikasi yang efektif, namun juga dapat menjadi wadah untuk menampung opini masyarakat. Salah satu media sosial yang banyak dipakai untuk mewadahi opini-opini tersebut adalah Twitter. Pengguna Twitter di Indonesia sering mengutarakan pendapatnya pada semua bidang termasuk olahraga, khususnya Timnas Sepakbola Indonesia. Berbagai macam sentimen terhadap Timnas Sepakbola Indonesia ini dapat dilihat di Twitter. Berlatarkan kondisi tersebut, maka diperlukan penelitian tentang opini masyarakat tentang performa Timnas Sepakbola Indonesia. Salah satu cara yang dilakukan adalah dengan melakukan analisis sentimen terhadap Timnas Sepakbola Indonesia di jejaring sosial Twitter menggunakan metode klasifikasi dan algoritma *Naïve Bayes* untuk mengklasifikasikan tweet positif atau negatif yang masyarakat berikan tentang Timnas Sepakbola Indonesia. Hasil dari analisis sentimen menggunakan algoritma *Naïve Bayes* pada klasifikasi *tweet* tentang opini terhadap Timnas Sepakbola Indonesia berjalan dengan baik dengan nilai akurasi 83%, kemudian nilai *precision* nilai positif sebesar 86%, nilai *precision* nilai negatif sebesar 81%, nilai *recall* sebesar 78%, dan nilai *specificity* penelitian ini 87,5% dengan menggunakan metode *confusion matrix* berdasarkan data yang diambil pada Januari sampai Mei 2021.

**Kata Kunci:** *Twitter, Analisis Sentimen, Timnas Sepakbola Indonesia, Algoritma Naïve Bayes*

## DAFTAR ISI

|                                            |           |
|--------------------------------------------|-----------|
| KATA PENGANTAR.....                        | V         |
| ABSTRAK .....                              | VI        |
| DAFTAR ISI.....                            | VII       |
| DAFTAR GAMBAR.....                         | IX        |
| DAFTAR TABEL.....                          | X         |
| DAFTAR SIMBOL .....                        | XI        |
| <b>BAB I PENDAHULUAN.....</b>              | <b>1</b>  |
| 1.1. Latar Belakang.....                   | 1         |
| 1.2. Rumusan Masalah.....                  | 2         |
| 1.3. Ruang Lingkup.....                    | 2         |
| 1.4. Tujuan .....                          | 2         |
| 1.5. Manfaat .....                         | 2         |
| 1.6. Luaran yang Diharapkan .....          | 3         |
| 1.7. Sistematika Penulisan .....           | 3         |
| <b>BAB II TINJAUAN PUSTAKA .....</b>       | <b>4</b>  |
| 2.1. Twitter .....                         | 4         |
| 2.2. Analisis Sentimen .....               | 4         |
| 2.3. <i>Text Mining</i> .....              | 5         |
| 2.4. Metode Klasifikasi.....               | 5         |
| 2.4.1 Algoritma <i>Naïve Bayes</i> .....   | 5         |
| 2.5. Penelitian Terdahulu .....            | 7         |
| <b>BAB III METODOLOGI PENELITIAN .....</b> | <b>8</b>  |
| 3.1. Kerangka Pikir .....                  | 8         |
| 3.1.1. Identifikasi masalah .....          | 8         |
| 3.1.2. Studi Literatur .....               | 9         |
| 3.1.3. Akuisisi Data.....                  | 9         |
| 3.1.4. Pelabelan Data.....                 | 10        |
| 3.1.5. Praproses Data.....                 | 10        |
| 3.1.6. Pembobotan TF IDF.....              | 11        |
| 3.1.7. Klasifikasi .....                   | 12        |
| 3.1.8. Hasil Klasifikasi .....             | 13        |
| 3.1.9. Evaluasi .....                      | 13        |
| 3.2. Perangkat Penelitian.....             | 14        |
| 3.3. Waktu Kegiatan .....                  | 14        |
| <b>BAB IV HASIL DAN PEMBAHASAN .....</b>   | <b>15</b> |
| 4.1. Data .....                            | 15        |
| 4.1.1. Pelabelan Data.....                 | 16        |
| 4.2. Pra Proses Data .....                 | 17        |
| 4.2.1. <i>Transform Cases</i> .....        | 18        |
| 4.2.2. <i>Cleaning</i> .....               | 18        |



|                       |                                 |           |
|-----------------------|---------------------------------|-----------|
| 4.2.3.                | <i>Stemming</i> .....           | 19        |
| 4.2.4.                | Tokenisasi .....                | 19        |
| 4.2.5.                | <i>Stopword Removal</i> .....   | 19        |
| 4.2.6.                | Normalisasi .....               | 20        |
| <b>4.3.</b>           | <b>Pembobotan TF IDF</b> .....  | <b>20</b> |
| <b>4.4.</b>           | <b>Proses Klasifikasi</b> ..... | <b>22</b> |
| 4.4.1.                | Data Latih.....                 | 23        |
| 4.4.2.                | Data Uji .....                  | 24        |
| <b>4.5.</b>           | <b>Hasil Klasifikasi</b> .....  | <b>27</b> |
| <b>4.6.</b>           | <b>Evaluasi</b> .....           | <b>28</b> |
| <b>BAB V</b>          | <b>PENUTUP</b> .....            | <b>31</b> |
| 5.1.                  | Kesimpulan .....                | 31        |
| 5.2.                  | Saran .....                     | 31        |
| <b>DAFTAR PUSTAKA</b> | .....                           | <b>32</b> |
| <b>RIWAYAT HIDUP</b>  | .....                           | <b>34</b> |
| <b>LAMPIRAN</b>       | .....                           | <b>35</b> |

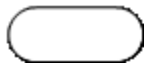
## **DAFTAR GAMBAR**

|                                   |    |
|-----------------------------------|----|
| Gambar 1 Kerangka Berpikir .....  | 8  |
| Gambar 2 Hasil Crawling Data..... | 15 |

## DAFTAR TABEL

|                                                    |    |
|----------------------------------------------------|----|
| Tabel 1 Confusion Matrix .....                     | 13 |
| Tabel 2 Waktu Kegiatan.....                        | 14 |
| Tabel 3 Contoh Pelabelan Data Tweet.....           | 16 |
| Tabel 4 Hasil Pelabelan Data .....                 | 16 |
| Tabel 5 Validasi dan Realibilitas Pengukuran ..... | 17 |
| Tabel 6 Sampel Data Sebelum Pra Proses .....       | 18 |
| Tabel 7 Hasil Transform Cases.....                 | 18 |
| Tabel 8 Hasil Cleaning.....                        | 18 |
| Tabel 9 Hasil Stemming.....                        | 19 |
| Tabel 10 Hasil Tokenisasi.....                     | 19 |
| Tabel 11 Hasil Stopword Removal .....              | 20 |
| Tabel 12 Hasil Normalisasi.....                    | 20 |
| Tabel 13 Sampel Data Latih.....                    | 20 |
| Tabel 14 Hasil Pembobotan TF-IDF .....             | 21 |
| Tabel 15 Probabilitas Sampel Data Latih.....       | 24 |
| Tabel 16 Sampel Data Uji.....                      | 25 |
| Tabel 17 Probabilitas Sampel Data Uji.....         | 26 |
| Tabel 18 Pembagian Data.....                       | 27 |
| Tabel 19 Evaluasi .....                            | 28 |
| Tabel 20 Variabel Evaluasi .....                   | 29 |

## DAFTAR SIMBOL

| Simbol Flowchart |                                                                                     |                                  |                                                                                                                                                                 |
|------------------|-------------------------------------------------------------------------------------|----------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------|
| NO               | GAMBAR                                                                              | NAMA                             | KETERANGAN                                                                                                                                                      |
| 1                |    | <i>Terminal<br/>(start, end)</i> | Menggambarkan bagaimana kegiatan dimulai atau kegiatan berakhir.                                                                                                |
| 2                |    | <i>Flow<br/>Direction</i>        | Menggambarkan hubungan antar simbol yang menyatakan suatu jalannya proses dalam sistem.                                                                         |
| 3                |    | <i>Process</i>                   | Menggambarkan deskripsi dari proses yang dijalankan.                                                                                                            |
| 4                |  | <i>Document</i>                  | Menggambarkan bahwa masukkan ( <i>input</i> ) berasal dari sebuah data dokumen yang dapat berupa kertas atau keluaran ( <i>output</i> ) yang dicetak ke kertas. |
| 5                |  | <i>Predefine<br/>process</i>     | Menggambarkan pelaksanaan dari sebuah proses atau disebut sebagai subprogram.                                                                                   |