

**ANALISIS SENTIMEN OPINI PUBLIK TERHADAP KUALITAS  
PERMAINAN DAN STRATEGI TIM NASIONAL INDONESIA  
DI ROUND 3 KUALIFIKASI PIALA DUNIA 2026  
BERDASARKAN DATA DARI X MENGGUNAKAN  
METODE K-NEAREST NEIGHBORS (KNN)**

**SKRIPSI**

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana  
Program Studi S1-Informatika



disusun oleh

**RAIHAN MAULUDANA**

**21.11.4453**

Kepada

**FAKULTAS ILMU KOMPUTER  
UNIVERSITAS AMIKOM YOGYAKARTA  
YOGYAKARTA**

**2025**

**ANALISIS SENTIMEN OPINI PUBLIK TERHADAP KUALITAS  
PERMAINAN DAN STRATEGI TIM NASIONAL INDONESIA  
DI ROUND 3 KUALIFIKASI PIALA DUNIA 2026  
BERDASARKAN DATA DARI X MENGGUNAKAN  
METODE K-NEAREST NEIGHBORS (KNN)**

**SKRIPSI**

untuk memenuhi salah satu syarat mencapai derajat Sarjana  
Program Studi S1-Informatika



disusun oleh

**RAIHAN MAULUDANA**

**21.11.4453**

Kepada

**FAKULTAS ILMU KOMPUTER  
UNIVERSITAS AMIKOM YOGYAKARTA  
YOGYAKARTA**

**2025**

**HALAMAN PERSETUJUAN**

**SKRIPSI**

**ANALISIS SENTIMEN OPINI PUBLIK TERHADAP KUALITAS PERMAINAN  
DAN STRATEGI TIM NASIONAL INDONESIA DI ROUND 3 KUALIFIKASI  
PIALA DUNIA 2026 BERDASARKAN DATA DARI X MENGGUNAKAN  
METODE K-NEAREST NEIGHBORS (KNN)**

yang disusun dan diajukan oleh

**Raihan Mauludana**

**21.11.4453**

telah disetujui oleh Dosen Pembimbing Skripsi  
pada tanggal 08 Januari 2025

**Dosen Pembimbing,**

**Robert Marco, M.T.,Ph.D**

**NIK. 190302228**

## HALAMAN PENGESAHAN

### SKRIPSI

# ANALISIS SENTIMEN OPINI PUBLIK TERHADAP KUALITAS PERMAINAN DAN STRATEGI TIM NASIONAL INDONESIA DI ROUND 3 KUALIFIKASI PIALA DUNIA 2026 BERDASARKAN DATA DARI X MENGGUNAKAN METODE K-NEAREST NEIGHBORS (KNN)

yang disusun dan diajukan oleh

**Raihan Mauludana**

**21.11.4453**

Telah dipertahankan di depan Dewan Penguji  
pada tanggal 23 Januari 2025

#### Susunan Dewan Penguji

##### Nama Penguji

Yuli Astuti, S.Kom., M.Kom  
NIK. 190302146

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

Robert Marco, M.T., Ph.D  
NIK. 190302228

##### Tanda Tangan



Skripsi ini telah diterima sebagai salah satu persyaratan  
untuk memperoleh gelar Sarjana Komputer  
Tanggal 23 Januari 2025

#### DEKAN FAKULTAS ILMU KOMPUTER



Hanif Al Fatta, S.Kom., M.Kom., Ph.D.  
NIK. 190302096

## HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama mahasiswa : Raihan Mauludana  
NIM : 21.11.4453

Menyatakan bahwa Skripsi dengan judul berikut:

**Analisis Sentimen Opini Publik terhadap Kualitas Permainan dan Strategi Tim Nasional Indonesia di Round 3 Kualifikasi Piala Dunia 2026 Berdasarkan Data dari X Menggunakan Metode K-Nearest Neighbors (KNN)**

Dosen Pembimbing : Robert Marco, M.T.,Ph.D

1. Karya tulis ini adalah benar-benar ASLI dan BELUM PERNAH diajukan untuk mendapatkan gelar akademik, baik di Universitas AMIKOM Yogyakarta maupun di Perguruan Tinggi lainnya.
2. Karya tulis ini merupakan gagasan, rumusan dan penelitian SAYA sendiri, tanpa bantuan pihak lain kecuali arahan dari Dosen Pembimbing.
3. Dalam karya tulis ini tidak terdapat karya atau pendapat orang lain, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan disebutkan dalam Daftar Pustaka pada karya tulis ini.
4. Perangkat lunak yang digunakan dalam penelitian ini sepenuhnya menjadi tanggung jawab SAYA, bukan tanggung jawab Universitas AMIKOM Yogyakarta.
5. Pernyataan ini SAYA buat dengan sesungguhnya, apabila di kemudian hari terdapat penyimpangan dan ketidakbenaran dalam pernyataan ini, maka SAYA bersedia menerima SANKSI AKADEMIK dengan pencabutan gelar yang sudah diperoleh, serta sanksi lainnya sesuai dengan norma yang berlaku di Perguruan Tinggi.

Yogyakarta, 23 Januari 2025

Yang Menyatakan,



Raihan Mauludana

## **HALAMAN PERSEMBAHAN**

Skripsi ini penulis persembahkan kepada kedua orang tua yang telah membesarkan, mendidik, mendukung dan mendoakan penulis. Kepada istri tercinta yang telah mendampingi, membantu dan mendukung penulis dengan kasih sayang. Serta kepada seluruh keluarga besar.





## KATA PENGANTAR

Puji dan syukur penulis panjatkan atas kehadiran Allah SWT, atas rahmat, dan karunia-NYA yang telah diberikan sehingga penulis dapat menyelesaikan skripsi yang berjudul “Analisis Sentimen Opini Publik terhadap Kualitas Permainan dan Strategi Tim Nasional Indonesia di Round 3 Kualifikasi Piala Dunia 2026 Berdasarkan Data dari X Menggunakan Metode K-Nearest Neighbors (KNN)”. Skripsi ini dapat diselesaikan atas bimbingan dan dukungan dari berbagai pihak, sehingga pada kesempatan kali ini dengan tulus penulis menyampaikan rasa hormat dan rasa terimakasih kepada :

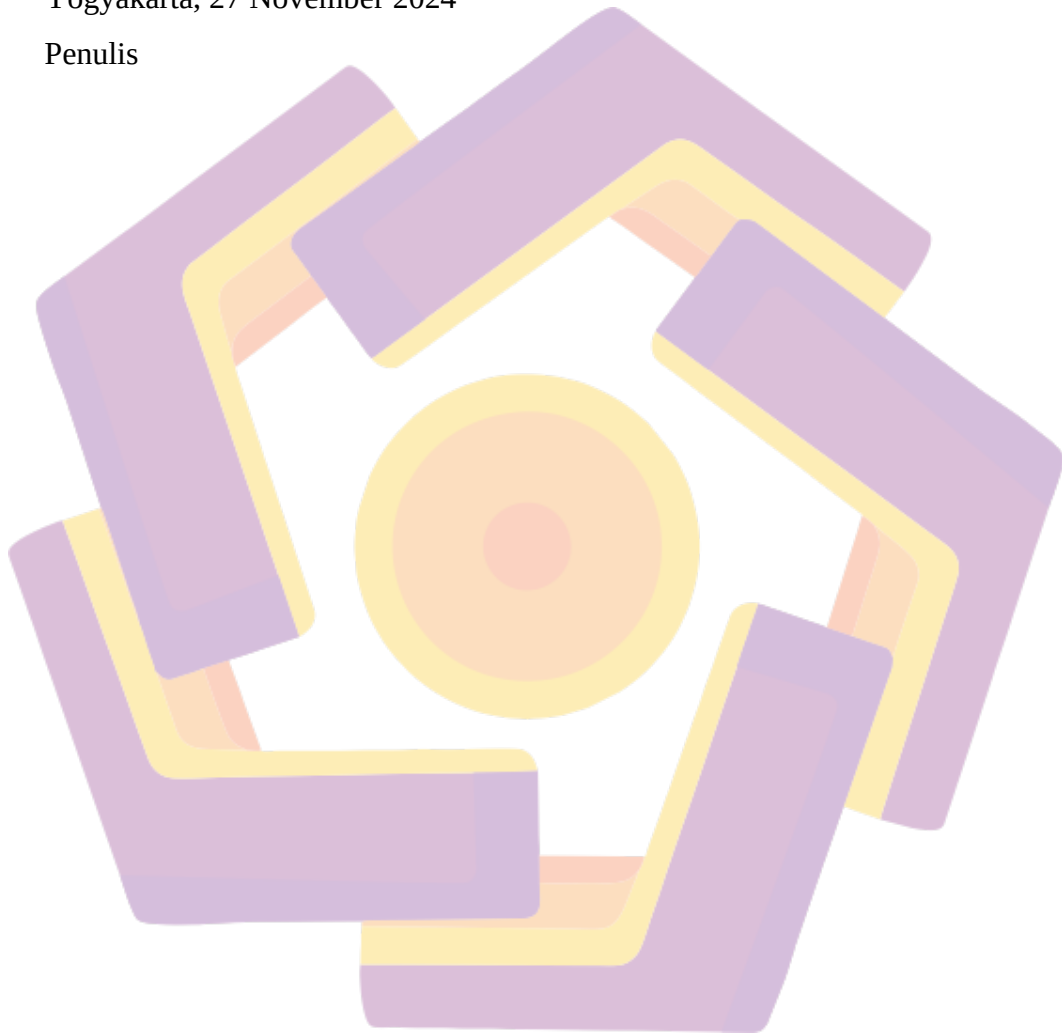
1. Prof. Dr. M. Suyanto, M.M. selaku Rektor Universitas Amikom Yogyakarta.
2. Bapak Hanif Al Fatta, S.Kom., M.Kom., Ph.D. selaku Dekan Fakultas Ilmu Komputer Universitas Amikom Yogyakarta.
3. Ibu Windha Mega Pradnya Duhita, M.Kom. selaku Ketua Program Studi Informatika Universitas Amikom Yogyakarta.
4. Bapak Robert Marco, M.T., Ph.D. selaku pembimbing yang telah bimbingan, arahan, dan motivasi kepada penulis selama proses penyusunan skripsi ini.
5. Ibu Yuli Astuti, S.Kom., M.Kom. dan Bapak Dr. Ferry Wahyu Wibowo, S.Si., M.Cs. selaku dosen penguji yang telah memberikan masukan, serta saran dalam penyusunan skripsi ini.
6. Udi Suprianto, S.Tp. dan Fitriyah, S.Tp. selaku orang tua yang selalu memberikan doa dan dukungan serta sebagai penyemangat penulis untuk menyelesaikan skripsi ini.
7. Anjali Putri Permata Sari, S.Kep. selaku istri tercinta yang selalu memberikan doa, dukungan dan menemani penulis untuk menyelesaikan skripsi ini.
8. Kakak-kakak dan saudara kembar penulis yang selalu mendukung dan mendoakan penulis menyelesaikan skripsi ini.

9. Serta semua teman-teman penulis dan semua pihak yang terlibat dalam penyusunan skripsi ini yang tidak bisa disebutkan satu persatu.

Semoga dengan selesainya penelitian ini mampu memberikan manfaat bagi semua orang dan menambah wawasan ilmu pengetahuan mengenai penelitian ini.

Yogyakarta, 27 November 2024

Penulis





## DAFTAR ISI

|                                           |      |
|-------------------------------------------|------|
| HALAMAN JUDUL .....                       | i    |
| HALAMAN PERSETUJUAN.....                  | ii   |
| HALAMAN PENGESAHAN .....                  | iii  |
| HALAMAN PERNYATAAN KEASLIAN SKRIPSI ..... | iv   |
| HALAMAN PERSEMBAHAN .....                 | v    |
| KATA PENGANTAR .....                      | vi   |
| DAFTAR ISI.....                           | viii |
| DAFTAR TABEL.....                         | xi   |
| DAFTAR GAMBAR.....                        | xii  |
| DAFTAR LAMPIRAN.....                      | xiii |
| DAFTAR LAMBANG DAN SINGKATAN .....        | xiv  |
| DAFTAR ISTILAH .....                      | xv   |
| INTISARI .....                            | xvi  |
| <i>ABSTRACT</i> .....                     | xvii |
| BAB I PENDAHULUAN.....                    | 1    |
| 1.1 Latar Belakang .....                  | 1    |
| 1.2 Rumusan Masalah .....                 | 3    |
| 1.3 Batasan Masalah.....                  | 3    |
| 1.4 Tujuan Penelitian.....                | 3    |
| 1.5 Manfaat Penelitian.....               | 4    |
| 1.5.1 Manfaat Teoritis.....               | 4    |
| 1.5.2 Manfaat Praktis .....               | 4    |
| 1.6 Sistematika Penelitian .....          | 4    |
| BAB II TINJAUAN PUSTAKA .....             | 6    |
| 2.1 Studi Literatur .....                 | 6    |
| 2.2 Dasar Teori .....                     | 15   |

|                                 |                                       |    |
|---------------------------------|---------------------------------------|----|
| 2.2.1                           | Analisis Sentimen .....               | 15 |
| 2.2.2                           | Opini Publik .....                    | 15 |
| 2.2.3                           | Aplikasi X (Sebelumnya Twitter) ..... | 15 |
| 2.2.4                           | Timnas Indonesia .....                | 16 |
| 2.2.5                           | Kualifikasi Piala Dunia 2026 .....    | 17 |
| 2.2.6                           | <i>Data Mining</i> .....              | 17 |
| 2.2.7                           | TF-IDF .....                          | 18 |
| 2.2.8                           | SMOTE .....                           | 18 |
| 2.2.9                           | K-Nearest Neighbor .....              | 19 |
| 2.2.10                          | <i>GridSearchCV</i> .....             | 19 |
| 2.2.11                          | Evaluasi Performa .....               | 20 |
| BAB III METODE PENELITIAN ..... |                                       | 22 |
| 3.1                             | Objek Penelitian .....                | 22 |
| 3.2                             | Alur Penelitian .....                 | 22 |
| 3.2.1                           | Pengambilan Dataset .....             | 24 |
| 3.2.2                           | Labeling .....                        | 25 |
| 3.2.3                           | Preprocessing Data .....              | 25 |
| 3.2.4                           | TF-IDF .....                          | 27 |
| 3.2.5                           | Split Data .....                      | 27 |
| 3.2.6                           | SMOTE .....                           | 27 |
| 3.2.7                           | K-Nearest Neighbor (KNN) .....        | 28 |
| 3.2.8                           | <i>GridSearchCV</i> .....             | 28 |
| 3.2.9                           | Evaluasi Performa .....               | 28 |
| 3.3                             | Alat dan Bahan .....                  | 29 |
| 3.3.1                           | Bahan Penelitian .....                | 29 |

|                                   |                                       |    |
|-----------------------------------|---------------------------------------|----|
| 3.3.2                             | Perangkat Keras .....                 | 30 |
| 3.3.3                             | Perangkat Lunak .....                 | 30 |
| 3.3.4                             | Tools dan Library .....               | 30 |
| BAB IV HASIL DAN PEMBAHASAN ..... |                                       | 31 |
| 4.1                               | Pengambilan Dataset .....             | 31 |
| 4.2                               | Labeling .....                        | 32 |
| 4.3                               | Preprocessing .....                   | 33 |
| 4.2.1                             | Cleaning .....                        | 33 |
| 4.2.2                             | Case Folding .....                    | 34 |
| 4.2.3                             | Normalisasi .....                     | 35 |
| 4.2.4                             | Stopword Removal .....                | 36 |
| 4.2.5                             | Tokenisasi .....                      | 38 |
| 4.2.6                             | Stemming .....                        | 38 |
| 4.4                               | TF-IDF .....                          | 39 |
| 4.5                               | Split Data .....                      | 42 |
| 4.6                               | SMOTE .....                           | 42 |
| 4.7                               | <i>K-Nearest Neighbor (KNN)</i> ..... | 43 |
| 4.8                               | Evaluasi Performa .....               | 43 |
| 4.8.1                             | Parameter Default .....               | 43 |
| 4.8.2                             | <i>GridSearchCV</i> .....             | 46 |
| BAB V PENUTUP .....               |                                       | 51 |
| 5.1                               | Kesimpulan .....                      | 51 |
| 5.2                               | Saran .....                           | 51 |
| REFERENSI .....                   |                                       | 52 |
| LAMPIRAN .....                    |                                       | 55 |

## DAFTAR TABEL

|                                                                         |    |
|-------------------------------------------------------------------------|----|
| Tabel 2. 1 Keaslian Penelitian .....                                    | 9  |
| Tabel 3. 1 Dataset .....                                                | 24 |
| Tabel 3. 2 Parameter .....                                              | 28 |
| Tabel 4. 1 Dataset .....                                                | 31 |
| Tabel 4. 2 Tahap Cleaning .....                                         | 33 |
| Tabel 4. 3 Tahap Case Folding .....                                     | 34 |
| Tabel 4. 4 Tahap Normalisasi .....                                      | 36 |
| Tabel 4. 5 Tahap Stopword Removal.....                                  | 37 |
| Tabel 4. 6 Tahap Tokenisasi .....                                       | 38 |
| Tabel 4. 7 Tahap Stemming .....                                         | 39 |
| Tabel 4. 8 Tahap Labeling .....                                         | 32 |
| Tabel 4. 9 Sampel Sebelum TF-IDF .....                                  | 40 |
| Tabel 4. 10 Sampel Setelah TF-IDF .....                                 | 41 |
| Tabel 4. 11 Hasil Pengujian Parameter Default.....                      | 44 |
| Tabel 4. 12 Hasil Pengujian Gridsearch.....                             | 47 |
| Tabel 4. 13 Perbandingan Algoritma KNN Default dan KNN Gridsearch ..... | 48 |

## DAFTAR GAMBAR

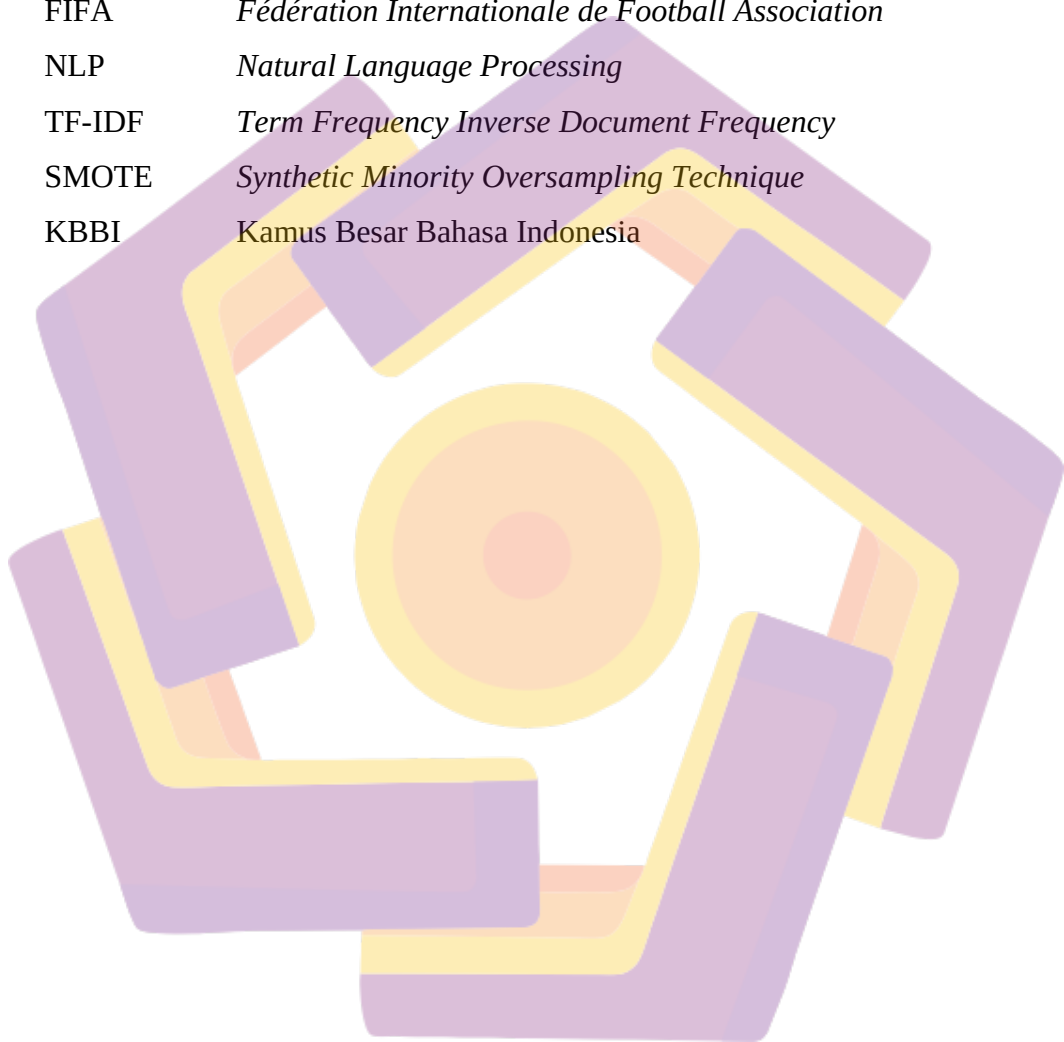
|                                                      |    |
|------------------------------------------------------|----|
| Gambar 2. 1 Logo Aplikasi X.....                     | 16 |
| Gambar 2. 2 Logo PSSI .....                          | 16 |
| Gambar 2. 3 Logo Piala Dunia.....                    | 17 |
| Gambar 2. 4 Confusion Matrix .....                   | 20 |
| Gambar 3. 1 Alur Penelitian .....                    | 23 |
| Gambar 4. 1 Visualisasi Dataset .....                | 32 |
| Gambar 4. 2 Visualisasi Tahap Cleaning .....         | 34 |
| Gambar 4. 3 Visualisasi Tahap Case Folding.....      | 35 |
| Gambar 4. 4 Visualisasi Tahap Normalisasi.....       | 36 |
| Gambar 4. 5 Visualisasi Tahap Stopword Removal ..... | 37 |
| Gambar 4. 6 Visualisasi Tahap Stemming.....          | 39 |
| Gambar 4. 7 Sentimen Sebelum SMOTE .....             | 42 |
| Gambar 4. 8 Sentimen Setelah SMOTE .....             | 43 |
| Gambar 4. 9 Confusion Matrix Parameter Default.....  | 44 |
| Gambar 4. 10 Confusion Matrix Gridsearch.....        | 46 |
| Gambar 4. 11 Visualisasi Sentimen Positif.....       | 49 |
| Gambar 4. 12 Visualisasi Sentimen Negatif .....      | 50 |
| Gambar 4. 13 Visualisasi Sentimen Netral .....       | 50 |

## DAFTAR LAMPIRAN

|                                                  |    |
|--------------------------------------------------|----|
| Lampiran 1 Kode Crawling .....                   | 55 |
| Lampiran 2 Kode Tahap Cleaning .....             | 55 |
| Lampiran 3 Kode Tahap Case Folding .....         | 55 |
| Lampiran 4 Kode Tahap Normalisasi .....          | 56 |
| Lampiran 5 Kode Tahap Stopword Removal .....     | 56 |
| Lampiran 6 Kode Tahap Tokenisasi .....           | 57 |
| Lampiran 7 Kode Tahap Stemming .....             | 57 |
| Lampiran 8 Kode Tahap Labeling .....             | 57 |
| Lampiran 9 Kode Tahap TF-IDF .....               | 58 |
| Lampiran 10 Kode Tahap Split Data .....          | 58 |
| Lampiran 11 Kode Tahap SMOTE .....               | 58 |
| Lampiran 12 Kode KNN Default .....               | 58 |
| Lampiran 13 Kode Evaluasi KNN Default .....      | 58 |
| Lampiran 14 Kode KNN GridSearchCV .....          | 59 |
| Lampiran 15 Kode Evaluasi KNN GridSearchCV ..... | 59 |

## DAFTAR LAMBANG DAN SINGKATAN

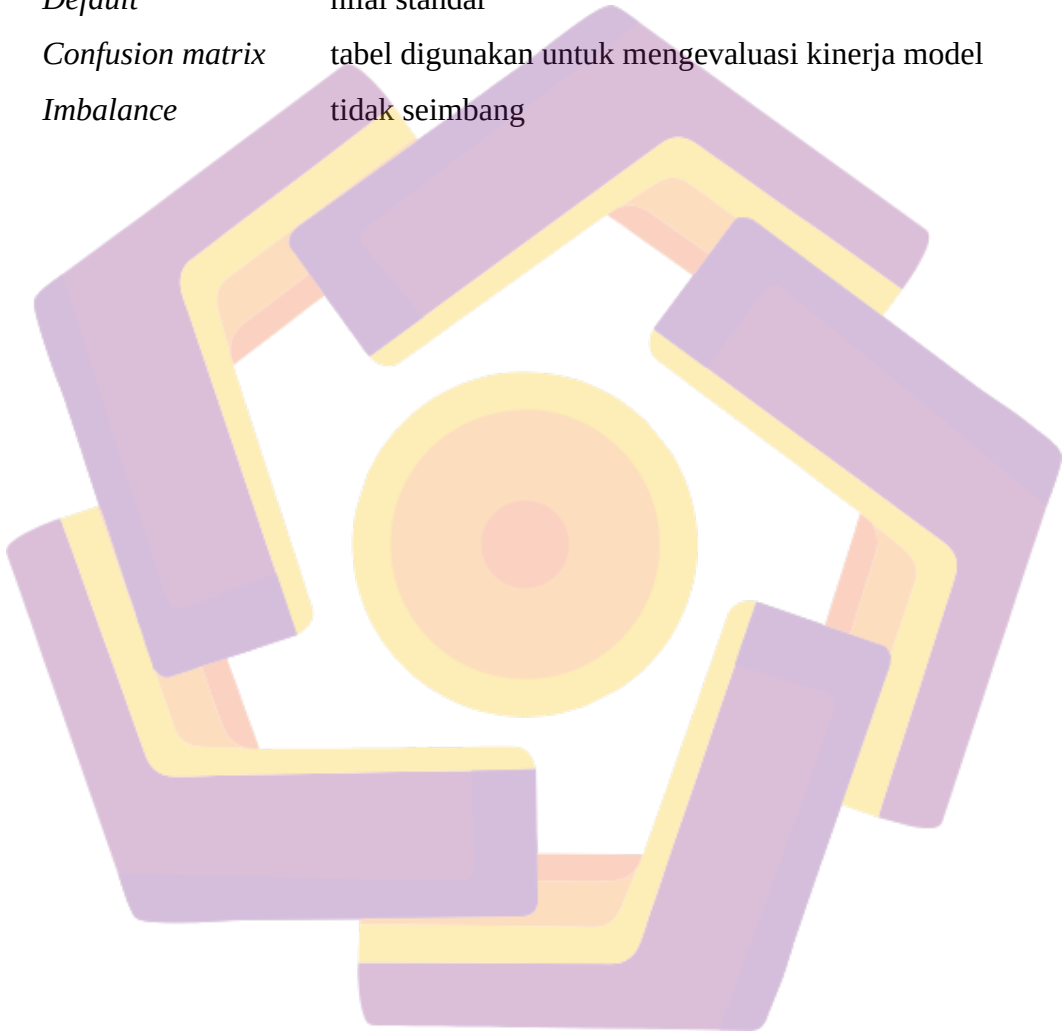
|          |                                                          |
|----------|----------------------------------------------------------|
| $\Sigma$ | Sigma Besar                                              |
| KNN      | K-Nearest Neighbor                                       |
| PSSI     | Persatuan Sepak Bola Seluruh Indonesia                   |
| FIFA     | <i>Fédération Internationale de Football Association</i> |
| NLP      | <i>Natural Language Processing</i>                       |
| TF-IDF   | <i>Term Frequency Inverse Document Frequency</i>         |
| SMOTE    | <i>Synthetic Minority Oversampling Technique</i>         |
| KBBI     | Kamus Besar Bahasa Indonesia                             |





## DAFTAR ISTILAH

|                         |                                                  |
|-------------------------|--------------------------------------------------|
| <i>Crawling</i>         | teknik pengambilan data                          |
| <i>Data mining</i>      | proses menganalisis data                         |
| <i>Vektor</i>           | besaran yang mempunyai arah                      |
| <i>Default</i>          | nilai standar                                    |
| <i>Confusion matrix</i> | tabel digunakan untuk mengevaluasi kinerja model |
| <i>Imbalance</i>        | tidak seimbang                                   |



## INTISARI

Sepak bola merupakan olahraga paling populer di Indonesia, dengan penggemar yang tersebar di seluruh negeri. Media sosial seperti aplikasi X (sebelumnya Twitter) menjadi platform utama bagi masyarakat untuk menyampaikan opini mereka terhadap performa Tim Nasional Indonesia, terutama dalam babak ketiga kualifikasi Piala Dunia 2026. Banyaknya opini publik ini memberikan peluang untuk menganalisis sentimen yang dapat mencerminkan dukungan, kritik, atau saran terhadap strategi dan permainan Timnas. Penelitian ini bertujuan untuk menganalisis sentimen opini publik terhadap kualitas permainan dan strategi Tim Nasional Indonesia berdasarkan data dari aplikasi X. Data dikumpulkan menggunakan teknik *crawling* dengan kata kunci spesifik, kemudian diproses melalui tahapan *preprocessing* seperti *cleaning*, *case folding*, normalisasi, dan *stopword removal*. Metode K-Nearest Neighbors (KNN) digunakan sebagai algoritma klasifikasi utama, dengan penerapan TF-IDF untuk pembobotan fitur dan SMOTE untuk menyeimbangkan data. Evaluasi performa model dilakukan dengan dua pendekatan: KNN dengan parameter *default* dan KNN yang dioptimalkan melalui *GridSearch*. Model KNN dengan parameter *default* menghasilkan akurasi sebesar 53%, sedangkan model KNN dengan *GridSearch* mencapai akurasi terbaik sebesar 90%. Hasil analisis menunjukkan bahwa opini publik terbagi menjadi sentimen positif, negatif, dan netral.

**Kata kunci:** Analisis Sentimen, Tim Nasional Indonesia, Kualifikasi Piala Dunia, KNN, Aplikasi X.

## ABSTRACT

Football is the most popular sport in Indonesia, with fans spread across the country. Social media platforms like X (formerly Twitter) serve as a primary medium for the public to express their opinions on the performance of the Indonesian National Team, especially during the third round of the 2026 World Cup Qualifiers. The abundance of public opinions provides an opportunity to analyze sentiments that reflect support, criticism, or suggestions regarding the team's strategies and gameplay. This study aims to analyze public sentiment regarding the quality of gameplay and strategies of the Indonesian National Team based on data from the X platform. Data was collected using crawling techniques with specific keywords and processed through preprocessing stages, including cleaning, case folding, normalization, and stopword removal. The K-Nearest Neighbors (KNN) method was employed as the primary classification algorithm, with the application of TF-IDF for feature weighting and SMOTE for balancing the dataset. The model's performance was evaluated using two approaches: KNN with default parameters and KNN optimized through GridSearch. The KNN model with default parameters achieved an accuracy of 53%, while the GridSearch-optimized KNN model achieved the best accuracy of 90%. The results indicate that public opinions are divided into positive, negative, and neutral sentiments.

**Keyword:** Sentiment Analysis, Indonesian National Team, World Cup Qualifiers, KNN, X application.