

**PERBANDINGAN TIGA MODEL KLASIFIKASI DALAM
ANALISIS SENTIMEN TIM REX REGUM QEON PADA MOBILE
LEGENDS BANG BANG PROFESSIONAL LEAGUE MUSIM KE
12 PADA TWITTER**

SKRIPSI

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



disusun oleh

FARRAS NUR HAIDAR RAMADHAN

19.11.3039

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2024

**PERBANDINGAN TIGA MODEL KLASIFIKASI DALAM
ANALISIS SENTIMEN TIM REX REGUM QEON PADA MOBILE
LEGENDS BANG BANG PROFESSIONAL LEAGUE MUSIM KE
12 PADA TWITTER**

SKRIPSI

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



disusun oleh

FARRAS NUR HAIDAR RAMADHAN

19.11.3039

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2024

HALAMAN PERSETUJUAN

HALAMAN PERSETUJUAN

SKRIPSI

PERBANDINGAN TIGA MODEL KLASIFIKASI DALAM ANALISIS SENTIMEN TIM REX REGUM QEON PADA MOBILE LEGENDS BANG BANG PROFESSIONAL LEAGUE MUSIM KE 12 PADA TWITTER

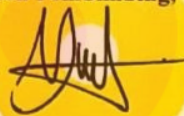
yang disusun dan diajukan oleh

Farras Nur Haidar Ramadhan

19.11.3039

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 18 Januari 2024

Dosen Pembimbing,



Yoga Pristyanto, S.Kom., M.Eng.
NIK. 190302412

HALAMAN PENGESAHAN

HALAMAN PENGESAHAN

SKRIPSI

PERBANDINGAN TIGA MODEL KLASIFIKASI DALAM ANALISIS SENTIMEN TIM REX REGUM QEON PADA MOBILE LEGENDS BANG BANG PROFESSIONAL LEAGUE MUSIM KE 12 PADA TWITTER

yang disusun dan diajukan oleh

Farras Nur Haidar Ramadhan

19.11.3039

Telah dipertahankan di depan Dewan Penguji
pada tanggal 18 Januari 2024

Susunan Dewan Penguji

Nama Penguji

Tanda Tangan

Supriatin, M.Kom
NIK. 190302239

Arifiyanto Hadinegoro, S.Kom, MT
NIK. 190302289

Rizqi Sukma Kharisma, M.Kom
NIK. 190302215



Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 18 Januari 2024

DEKAN FAKULTAS ILMU KOMPUTER



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

HALAMAN PERNYATAAN KEASLIAN SKRIPSI

HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama mahasiswa : Farras Nur Haidar Ramadhan
NIM : 19.11.3039

Menyatakan bahwa Skripsi dengan judul berikut:

Perbandingan Tiga Model Klasifikasi Dalam Analisis Sentimen Tim Rex Regum Qeon Pada Mobile Legends Bang Bang Professional League Musim Ke 12 Pada Twitter

Dosen Pembimbing : Yoga Pristyanto, S.Kom., M.Eng.

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, 18 Januari 2024

Yang Menyatakan,



Farras Nur Haidar Ramadhan

HALAMAN PERSEMBAHAN

Alhamdulillah, dengan mengucapkan rasa syukur dan juga segala puji Tuhan Yang Maha Esa, Allah Subhanahu Wa Ta'ala. Penelitian ini dapat terselesaikan dengan baik, maka dari itu, penulis ingin mempersembahkan keberhasilan penulis dalam penyusunan penelitian skripsi ini kepada:

1. Orang tua yang telah memberikan kesempatan penulis untuk melanjutkan ke jenjang perguruan tinggi, serta kakak penulis yang telah memberikan semangat untuk mendorong penulis agar menyelesaikan penyusunan penelitian skripsi ini.
2. Bapak Yoga Pristyanto S.Kom., M.Eng selaku dosen pembimbing penulis, dengan penuh kesabaran beliau selalu memberikan pengarahan kepada penulis dalam melakukan penulisan penelitian ini.
3. Semua teman-teman penulis yang telah memberikan motivasi dorongan untuk dapat menyelesaikan penyusunan penelitian ini.
4. Angkringan Tuan Marino yang telah memberikan tempat untuk *refreshing* disaat penulis sedang jenuh dan kebingungan dalam pengerjaan penelitian ini.

KATA PENGANTAR

Dengan menyebut nama Allah Subhanahu Wa Ta'ala, penulis memanjatkan puji dan syukur kepada-Nya. Atas rahmat dan nikmat yang dikaruniakan kepada penulis sehingga dapat menyelesaikan penyusunan penelitian skripsi yang berjudul “Perbandingan Tiga Model Klasifikasi Dalam Analisis Sentimen Tim Rex Regum Qeon Pada Mobile Legends: Bang Bang Professional League Musim ke 12 Pada Twitter”.

Dalam penyusunan penelitian skripsi ini tak lupa untuk mengucapkan terima kasih kepada pihak yang telah membantu penulis dalam menyusun penelitian ini, diantaranya penulis berterima kasih kepada:

1. Bapak Prof. Dr. M. Suyanto, MM. selaku rektor Universitas AMIKOM Yogyakarta
2. Bapak Yoga Pristyanto, S.Kom, M.Eng. sebagai dosen pembimbing yang telah memberikan arahan kepada penulis dalam penyusunan penelitian skripsi.
3. Bapak, Ibu dosen Universitas AMIKOM Yogyakarta yang telah memberikan ilmu selama perkuliahan.
4. Staf Resource Center yang telah memberikan akses untuk meminjam buku dan menyediakan fasilitas yang sangat membantu penulis dalam penyusunan penelitian skripsi ini.

Penelitian ini tidak sepenuhnya sempurna, masih terdapat kekurangan dalam penyusunan skripsi ini. Tetapi penulis mengharapkan skripsi ini dapat bermanfaat kepada peneliti selanjutnya yang ingin meneliti dengan topik yang berhubungan dengan penelitian penulis. Akhir kata penulis mengucapkan terima kasih.

Yogyakarta, 18 Maret 2024

Penulis

DAFTAR ISI

HALAMAN PERSETUJUAN	ii
HALAMAN PENGESAHAN.....	iii
HALAMAN PERNYATAAN KEASLIAN SKRIPSI	iv
HALAMAN PERSEMBAHAN	v
KATA PENGANTAR	vi
DAFTAR ISI.....	vii
DAFTAR TABEL.....	x
DAFTAR GAMBAR	xi
INTISARI.....	xii
ABSTRACT.....	xiii
BAB I.....	1
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 Sistematika Penulisan	4
BAB II.....	5
TINJAUAN PUSTAKA	5
2.1 Studi Literatur.....	5
2.2 Dasar Teori.....	11
2.2.1 Analisis Sentimen	11
2.2.2 Text Mining.....	11
2.2.3 Preprocessing	12
2.2.4 Naïve Bayes	12
2.2.5 Support Vector Machine	13
2.2.6 K-Nearest Neighbor	14
2.2.7 Rex Regum Qeon.....	15

2.2.8	Twitter	15
2.2.9	Confusion Matrix	15
2.2.10	Term Frequency–Invers Document Frequency.....	17
2.2.11	Stratified K-Fold Cross Validation	17
2.2.12	Klasifikasi	18
BAB III		19
METODOLOGI PENELITIAN.....		19
3.1	Alat dan Bahan Penelitian.....	19
3.1.1	Alat Penelitian.....	19
3.1.2	Bahan Penelitian.....	20
3.2	Alur Penelitian	21
3.3	Preprocessing	23
3.3.1	Case Folding	23
3.3.2	Cleaning	23
3.3.3	Tokenizing.....	24
3.3.4	Stopword Removal.....	24
3.3.5	Stemming	25
3.4	Pelabelan Data.....	25
3.5	Pembobotan Kata TF-IDF.....	26
3.6	Pembagian Data Dengan Stratified K-Fold Cross Validation.....	26
3.7	Klasifikasi	26
3.8	Evaluasi Dengan Confusion Matrix.....	27
3.9	Perbandingan Hasil Klasifikasi.....	27
3.10	Kesimpulan	27
3.11	Dokumentasi	27
BAB IV		28
HASIL DAN PEMBAHASAN.....		28
4.1	Pengambilan dan Pengumpulan Data	28
4.2	Preprocessing Teks.....	30
4.2.1	Case Folding	30
4.2.2	Cleaning	31

4.2.3	Tokenizing.....	33
4.2.4	Stopword Removal.....	33
4.2.5	Stemming	34
4.3	Data Labeling	36
4.4	Pembobotan TF-IDF	36
4.5	Pembagian Data	37
4.7	Klasifikasi	38
4.7.1	Metode Naïve Bayes	38
4.7.2	Metode Support Vector Machine	39
4.7.3	Metode K-Nearest Neighbor	39
4.8	Evaluasi.....	42
4.8.1	Metode Naïve Bayes	42
4.8.2	Metode Support Vector Machine	46
4.8.3	Metode K-Nearest Neighbor	49
4.9	Perbandingan.....	52
BAB V.....		54
PENUTUP.....		54
5.1	Kesimpulan	54
5.2	Saran.....	54
REFERENSI		55

DAFTAR TABEL

Tabel 2.1 Keaslian Penelitian.....	8
Tabel 2.2 Confusion Matrix	16
Tabel 3.1 Spesifikasi Perangkat Keras	19
Tabel 3.2 Perangkat Lunak.....	19
Tabel 3.3 Hasil Pengumpulan Data dari <i>Crawling</i> Data.....	20
Tabel 3.4 Case Folding.....	23
Tabel 3.5 Cleaning.....	24
Tabel 3.6 Tokenizing	24
Tabel 3.7 Stopword Removal.....	25
Tabel 3.8 Stemming.....	25
Tabel 3.9 Daftar 5 Kata dengan Bobot Tertinggi	26
Tabel 4.1 Matriks metode Naïve Bayes	43
Tabel 4.2 Matriks metode Support Vector Machine.....	46
Tabel 4.3 Matriks metode K-Nearest Neighbor	49
Tabel 4.4 Hasil kinerja setiap metode klasifikasi.....	52

DAFTAR GAMBAR

Gambar 2.1 Pencarian <i>hyperplane</i>	14
Gambar 2.2 <i>hyperplane</i> maksimum	14
Gambar 3.1 Alur Penelitian.....	22
Gambar 4.1 Hasil observasi pada minggu pertama.....	28
Gambar 4.2 Hasil observasi pada minggu ke enam.	29
Gambar 4.3 Instalasi Node.js.	29
Gambar 4.4 Konfigurasi TweetHarvest.	30
Gambar 4.5 Case Folding.....	31
Gambar 4.6 Cleaning 1.....	31
Gambar 4.7 Cleaning 2.	32
Gambar 4.8 Tokenizing.....	33
Gambar 4.9 Stopword Removal.....	34
Gambar 4.10 Stemming	35
Gambar 4.11 Kelas Sentimen.....	36
Gambar 4.12 Script TF-IDF.....	37
Gambar 4.13 Inisialisasi Stratified K-Fold Cross Validation	37
Gambar 4.14 Script Pembagian Data Latih dan Data Uji	38
Gambar 4.15 Script Naïve Bayes	38
Gambar 4.16 Script Support Vector Machine.....	39
Gambar 4.17 Script Pencarian K Terbaik Menggunakan GridSearchCV 1.....	40
Gambar 4.18 Script Pencarian K Terbaik Menggunakan GridSearchCV 2.....	40
Gambar 4.19 Hasil K Terbaik	41
Gambar 4.20 Script K-Nearest Neighbor.....	41
Gambar 4.21 Script Evaluasi Menghitung Confusion Matrix Pada Model	42
Gambar 4.22 Hasil Setiap Iterasi Naïve Bayes	45
Gambar 4.23 Rata-rata Matrik Naïve Bayes	45
Gambar 4.24 Hasil Setiap Iterasi Support Vector Machine.....	48
Gambar 4.25 Rata-rata Matrik Support Vector Machine.....	48
Gambar 4.26 Hasil Setiap Iterasi K-Nearest Neighbor.....	51
Gambar 4.27 Rata-rata Matrik K-Nearest Neighbor.....	52
Gambar 4.28 Perbandingan Metode Klasifikasi	53

INTISARI

Twitter merupakan sosial media untuk menjalin atau berinteraksi dengan orang lain, twitter menjadi tempat untuk menyampaikan opini terhadap suatu peristiwa, seperti halnya tim RRQ yang memperoleh kemenangan dan kekalahan pada MPL Indonesia musim ke 12, para pendukung tim tersebut pun mengekspresikan tentang pertandingan yang dialami oleh tim RRQ, terdapat sentimen positif dan juga ada sentimen negatif atas hasil pertandingan. Hal tersebut pun dapat digunakan sebagai data untuk melakukan penelitian perbandingan kinerja pada metode dalam melakukan klasifikasi terhadap sentimen. Metode yang digunakan diantaranya Naïve Bayes, Support Vector Machine dan K-Nearest Neighbor. Tahapan untuk melakukan perbandingan kinerja metode klasifikasi yaitu melakukan pengumpulan data, *preprocessing*, pembobotan kata dengan menggunakan TF-IDF, pembagian data dengan menggunakan Stratified K-Fold Cross Validation sebanyak 10 iterasi, kemudian melakukan klasifikasi dengan metode yang digunakan. Hasil dari klasifikasi dengan 10 iterasi menunjukkan bahwa kinerja dari metode Naïve Bayes dengan nilai *overall* sebesar 81,54%, kemudian metode K-Nearest Neighbor memperoleh nilai 81,17% dan metode Support Vector Machine yang memperoleh nilai sebesar 80,94%. Dari hasil dan pembahasan penelitian tersebut bisa disimpulkan bahwa metode klasifikasi Naïve Bayes unggul dalam melakukan klasifikasi sentimen pada penelitian ini.

Kata kunci: RRQ, kinerja, Naïve Bayes, Support Vector Machine, K-Nearest Neighbor.

ABSTRACT

Twitter is a social media to establish or interact with other people, twitter is a place to express opinions on an event, such as the RRQ team that won and lost in the 12th season of MPL Indonesia, the team's supporters also expressed about the match experienced by the RRQ team, there are positive sentiments and there are also negative sentiments on the results of the match. This can also be used as data to conduct performance comparison research on methods in classifying sentiment. The methods used include Naïve Bayes, Support Vector Machine and K-Nearest Neighbor. Stages to compare the performance of classification methods are collecting data, preprocessing, weighting words using TF-IDF, dividing data using K-Fold Cross Validation for 10 iterations, then classifying with the method used. The results of the classification with 10 iterations show that the performance of the Naïve Bayes method with an overall value of 81.54%, then the K-Nearest Neighbor method obtained a value of 81.17% and the Support Vector Machine method which obtained a value of 80.94%. From the results and discussion of the research, it can be concluded that the Naïve Bayes classification method is superior in performing sentiment classification in this study.

Keyword: RRQ, performance, Naïve Bayes, Support Vector Machine, K-Nearest Neighbor.

