

**PERBANDINGAN ANALISIS SENTIMEN MENGGUNAKAN
METODE *SUPPORT VECTOR MACHINE (SVM)* DAN
NAÏVE BAYES TERHADAP APLIKASI
KAI ACCESS**

SKRIPSI

Oleh:

MUHAMMAD RIKY SUDRAJAT

201710225185



**PROGRAM STUDI INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS BHAYANGKARA JAKARTA RAYA
2021**

LEMBAR PERSETUJUAN PEMBIMBING

Judul Skripsi : Perbandingan Analisis Sentimen Menggunakan
Metode *Support Vector Machine (SVM)* dan *Naïve
Bayes* Terhadap Aplikasi *KAI Access*

Nama Mahasiswa : Muhammad Ricky Sudrajat

Nomor Pokok Mahasiswa : 201710225185

Program Studi/Fakultas : Informatika / Ilmu Komputer

Tanggal Lulus Ujian Skripsi : 14 Juli 2021

Bekasi, 19 Juli 2021

MENYETUJUI

Pembimbing I

Pembimbing II



Prima Dina Atika, S.Kom., M.Kom.

Herlawati, S.Si., M.M., M.Kom.

NIDN. 0311037107

NIDN. 0311097302

LEMBAR PENGESAHAN

Judul Skripsi : Perbandingan Analisis Sentimen Menggunakan
Metode *Support Vector Machine (SVM)* dan *Naive
Bayes* Terhadap Aplikasi *KAI Access*

Nama Mahasiswa : Muhammad Riky Sudrajat

Nomor Pokok Mahasiswa : 201710225185

Program Studi / Fakultas : Informatika / Ilmu Komputer

Tanggal Lulus Ujian Skripsi : 14 Juli 2021

Bekasi, 19 Juli 2021

Mengesahkan,

Ketua Tim Penguji : Rakmat Purnomo, S.Pd., S.Kom., M.Kom
NIDN. 0322108201

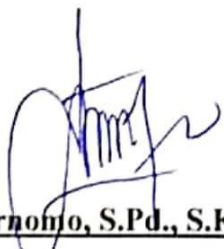
Penguji (II) : Rafika Sari, S.Si., M.Si.
NIDN. 0329098902


Penguji (III) : Prima Dina Atika, S.Kom., M.Kom.
NIDN. 0311037107

Mengetahui,

Ketua Program Studi
Informatika

Dekan
Fakultas Ilmu Komputer


Rakmat Purnomo, S.Pd., S.Kom., M.Kom.
NIDN. 0322108201


Herlawati, S.Si., M.M., M.Kom.
NIDN. 0311097302



LEMBAR PERNYATAAN BUKAN PLAGIASI

Yang bertanda tangan dibawah ini :

Nama : Muhammad Riky Sudrajat
NPM : 201710225185
Program Studi : Informatika
Fakultas : Ilmu Komputer
Judul Tugas Akhir : Perbandingan Analisis Sentimen Menggunakan
Metode *Support Vector Machines (SVM)* dan *Naïve Bayes*
Terhadap Aplikasi KAI Access

Dengan ini menyatakan bahwa hasil penulisan skripsi yang telah saya buat ini merupakan **hasil karya saya sendiri dan benar keasliannya**. Apabila dikemudian hari penulisan skripsi ini merupakan plagiat atau penjiplakan terhadap karya orang lain, maka saya bersedia mempertanggungjawabkan sekaligus bersedia menerima sanksi berdasarkan tata tertib di Universitas Bhayangkara Jakarta Raya.

Demikian pernyataan ini saya buat dalam keadaan sadar dan tidak dipaksakan dari pihak manapun.

Bekasi, 18 Juni 2021

Penulis



Muhammad Riky Sudrajat

ABSTRAK

Muhammad Riky Sudrajat. 201710225185. Perbandingan Analisis Sentimen Menggunakan Metode *Support Vector Machine (SVM)* dan *Naïve Bayes* Terhadap Aplikasi KAI Access.

Google Play Store merupakan salah satu *platform* di *android* untuk mengunduh suatu aplikasi, Google Play juga menyediakan fitur untuk masyarakat dapat memberikan komentar / ulasan terhadap aplikasi yang di unduh nya. Ulasan mengenai aplikasi tersebut berupa sebuah persepsi, baik itu positif maupun negatif, ulasan dari salah satu aplikasi yang ada di google play yaitu aplikasi KAI *access* dapat dijadikan sebagai bahan penelitian untuk mencari sebuah informasi. Teknik yang dapat digunakan untuk penelitian ini adalah analisis sentimen, metode klasifikasi yang akan digunakan untuk analisis sentimen ini adalah *support vector machine* serta *Naive Bayes* sebagai pembanding untuk mencari akurasi yang lebih baik dari kedua algoritma tersebut, Metode *Support Vector Machine (SVM)* merupakan metode learning machine (pembelajaran mesin) satu kelas dengan metode jaringan syaraf tiruan yang dapat mengenali pola dari masukan atau contoh yang diberikan dan juga termasuk ke dalam supervised learning. Sedangkan *Naïve Bayes* merupakan metode pengklasifikasian peluang sederhana dengan asumsi antar variabel penjelas saling bebas (*independen*). Alasan penulis memilih kedua metode ini karena kedua metode mempunyai tingkat akurasi yang tinggi pada setiap pengklasifikasiannya. Penelitian ini dapat membantu para *developer* untuk mengetahui kekurangan maupun kelebihan yang harus di tingkatkan terhadap aplikasinya. Hasil penelitian dengan klasifikasi *Support Vector Machine (SVM)* diperoleh tingkat akurasi sebesar 93% sedangkan dengan menggunakan metode *Naïve Bayes* yaitu sebesar 89%. Sehingga metode *Support Vector Machine* memberikan tingkat akurasi yang lebih tinggi daripada metode *Naïve Bayes*.

Kata Kunci: Google Play, KAI Access, *Support Vector Machine*, *Naïve Bayes*

ABSTRACT

Muhammad Ricky Sudrajat. 201710225185. *Comparison of Sentiment Analysis Using Support Vector Machine (SVM) and Nave Bayes Methods Against KAI Access Applications.*

Google Play Store is one of the platforms on Android to download an application, Google Play also provides a feature for the public to be able to provide comments/reviews of the downloaded application. Reviews of the application are in the form of perception, both positive and negative, a review of one of the applications on google play, namely the KAI access application, can be used as research material to find information. The technique that can be used for this research is sentiment analysis, the classification method that will be used for this sentiment analysis is the support vector machine and Naïve Bayes as a comparison to find better accuracy from the two algorithms, the Support Vector Machine (SVM) method is a learning method machine (machine learning) a class with an artificial neural network method that can recognize patterns from the input or examples given and is also included in supervised learning. Meanwhile, Naïve Bayes is a simple probability classification method with the assumption that the explanatory variables are independent. The reason the author chooses these two methods is that both methods have a high level of accuracy in each classification. This research can help developers to find out the advantages and disadvantages that must be improved on the application. The results of the study using the Support Vector Machine (SVM) classification obtained an accuracy rate of 93% while using the Naïve Bayes method, it was 89%. So the Support Vector Machine method provides a higher level of accuracy than the Naïve Bayes method.

Keywords: *Google Play, KAI Access, Support Vector Machine, Naïve Bayes*

LEMBAR PERNYATAAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIK

Sebagai sivitas akademik Universitas Bhayangkara Jakarta Raya, saya yang bertanda tangan di bawah ini :

Nama : Muhammad Riky Sudrajat
NPM : 2017102252185
Program Studi : Informatika
Fakultas : Ilmu Komputer
Jenis Karya : Skripsi

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

PERBANDINGAN ANALISIS SENTIMEN MENGGUNAKAN METODE *SUPPORT VECTOR MACHINES (SVM)* DAN *NAÏVE BAYES* TERHADAP APLIKASI KAI ACCESS

beserta perangkat yang ada (bila diperlukan). Dengan hak bebas royalti non-eksklusif ini, Universitas Bhayangkara Jakarta Raya berhak menyimpan, mengalihmediakan, mengelolanya dalam bentuk pangkalan data (*database*), mendistribusikannya dan mempublikasikannya di Internet atau media lain untuk kepentingan akademis tanpa perlu meminta ijin dari saya selama tetap mencantumkan nama saya sebagai penulis dan sebagai pemilik hak cipta.

Segala bentuk tuntutan hukum yang timbul atas pelanggaran hak cipta dalam karya ilmiah ini menjadi tanggung jawab saya pribadi

Demikian pernyataan ini saya buat dengan sebenarnya.

Dibuat di : Bekasi
Pada tanggal : 18 Juni 2021
Yang Menyatakan



Muhammad Riky Sudrajat

KATA PENGANTAR

Puji syukur Kehadirat Allah yang telah melimpahkan Rahmat, Hidayah serta Karunia-Nya sehingga Penelitian ini dengan judul “Perbandingan Analisis Sentimen Menggunakan Metode Support Vector Machine (SVM) dan Naïve Bayes Terhadap Aplikasi KAI Access” ini dapat terselesaikan. Shalawat dan salam semoga senantiasa tercurahkan kepada Nabi Muhammad SAW beserta keluarga, sahabat serta para pengikut beliau hingga akhir zaman yang syafaatnya dinantikan di akhirat kelak.

Penelitian ini merupakan salah satu syarat guna memperoleh gelar Sarjana di Program Studi Informatika, Fakultas Ilmu Komputer, Universitas Bhayangkara Jakarta Raya. Selama penyelesaian Penelitian ini, penulis menyadari banyak pihak yang telah memberikan dorongan, bantuan serta bimbingan hingga Penelitian ini dapat terselesaikan. Oleh karena itu, pada kesempatan ini penulis ingin menyampaikan ucapan terima kasih terlebih kepada Kedua orang tua tercinta, Bapak Ajat dan Ibu Nurhayati dan juga kepada:

1. Bapak Irjen Pol (Purn) Dr. bambang Karsono, S.H., M.M. selaku Rektor Universitas Bhayangkara Jakarta Raya.
2. Ibu Herlawati, S.Si., M.M., M.Kom. selaku Dekan Fakultas Ilmu Komputer Univesitas Bhayangkara Jakarta Raya, dan Pembimbing II penyusunan skripsi ini.
3. Bapak Rakhmat Purnomo, S.Pd., S.Kom., M.Kom. selaku Ketua Program Studi Informatika Universitas Bhayangkara Jakarta Raya.
4. Ibu Prima Dina Atika, S.Kom., M.Kom. selaku Pembimbing I yang selalu memberikan arahan serta ilmunya.
5. Keluarga saya beserta adik saya Fadli Januar yang selalu memberikan semangat, dukungan dan doa setiap saat.
6. Untuk kekasih tercinta, Salsabiila Tsaabitah P.H. yang telah banyak membantu dan memberikan dukungan hingga selesainya skripsi ini.
7. Teman-teman grup “KERKOM” yang telah banyak memberikan semangat dan dukungan moril.
8. Segenap civitas akademika Prodi Informatika, terutama seluruh dosen, terima kasih atas segenap ilmu bimbingannya.

9. Semua pihak yang ikut membantu dalam menyelesaikan skripsi ini baik berupa materiil maupun moril.

Penulis menyadari bahwa Penelitian ini masih jauh dari kesempurnaan. Oleh karena itu, segala kritik dan saran yang bersifat membangun selalu penulis harapkan. Semoga Penelitian ini dapat memberikan manfaat. Akhir kata, semoga Allah senantiasa melimpahkan Rahmat dan Hidayah-Nya kepada kita semua.

Bekasi, 18 Juni 2021



Muhammad Ricky Sudrajat



DAFTAR ISI

	Halaman
JUDUL	i
LEMBAR PERSETUJUAN PEMBIMBING	ii
LEMBAR PENGESAHAN	iii
LEMBAR PERNYATAAN BUKAN PLAGIASI	iv
ABSTRAK	v
ABSTRACT	vi
LEMBAR PERNYATAAN PUBLIKASI	vii
KATA PENGANTAR	viii
DAFTAR ISI	x
DAFTAR TABEL	xiii
DAFTAR GAMBAR	xiv
DAFTAR LAMPIRAN	xvi
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Identifikasi Masalah	4
1.3 Rumusan Masalah	5
1.4 Batasan Masalah	5
1.5 Tujuan Penelitian	5
1.6 Manfaat Penelitian	6
1.7 Tempat dan Waktu Penelitian	6
1.8 Metode Penelitian	6
1.9 Sistematika Penulisan	7
BAB II LANDASAN TEORI	9
2.1 Tinjauan Pustaka	9
2.1.1 Penelitian Terdahulu	9
2.1.2 Tabel <i>Related Research / State of The Art</i>	11
2.2 Landasan Teori	14
2.2.1 PT Kereta Api Indonesia	14
2.2.2 KAI Access	14
2.2.3 <i>Google Play</i>	15
2.2.4 <i>Online Review</i>	16

2.2.5 Analisis Sentimen	16
2.2.6 <i>Text Mining</i>	16
2.2.7 <i>Text Preprocessing</i>	17
2.2.7.1 <i>Stop Word Removal (Filtering)</i>	17
2.2.7.2 <i>Stemming</i>	17
2.2.7.3 <i>Punctuation Marks</i>	17
2.2.8 Pembobotan TF-IDF	17
2.2.9 <i>Python</i>	18
2.2.10 <i>Google Colaboratory</i>	19
2.2.11 <i>Support Vector Machines</i>	19
2.2.12 <i>Naïve Bayes</i>	22
2.2.13 Evaluasi Sistem Klasifikasi	23
2.2.13.1 <i>Train/Text Split</i>	24
2.2.13.2 <i>Precision</i>	25
2.2.13.3 <i>Recall</i>	25
2.2.13.4 <i>Accuracy</i>	25
2.2.14 Metode <i>Waterfall</i>	25
2.2.15 <i>World Wide Web (WWW)</i>	27
2.2.16 <i>Hypertext Preprocessor (PHP)</i>	27
2.2.17 <i>Java Script</i>	27
2.2.18 <i>Hyperlink Text Markup Language (HTML)</i>	28
2.2.19 <i>MySQL</i>	28
2.2.20 <i>PhpMyAdmin</i>	28
2.2.21 <i>Sublime Text 3</i>	28
2.2.22 <i>CodeIgniter</i>	29
BAB III METODOLOGI PENELITIAN	30
3.1 Obyek Penelitian	30
3.2 Variabel dan Defisini Operasional Variabel	30
3.3 Jenis dan Sumber Data	30
3.4 Metode Analisis Data	32
3.5 Kerangka Pemikiran	32
3.6 Alur Penelitian	34
3.7 Perancangan Antar Muka	35
3.8 Analisis Kebutuhan Sistem	39

BAB IV PERANCANGAN SISTEM DAN IMPLEMENTASI	40
4.1 <i>Preprocessing Data</i>	40
4.1.1 <i>Stop Word Removal</i>	41
4.1.2 <i>Stemming</i>	42
4.1.3 <i>Punctuation Marks</i>	43
4.2 Representasi Model	44
4.2.1 Perhitungan Skor Sentimen	44
4.2.2 Perhitungan Label Sentimen	45
4.3 Klasifikasi Sentimen	46
4.3.1 Data Latih dan Data Uji	46
4.3.2 Klasifikasi <i>Support Vector Machine</i> dan <i>Naïve Bayes</i>	47
4.4 Uji Coba Aplikasi	50
BAB V PENUTUP	56
5.1 Kesimpulan	56
5.2 Saran	56
DAFTAR PUSTAKA	57
LAMPIRAN	60

DAFTAR TABEL

	Halaman
Tabel 2.1 Related Research / State of The Art	11
Tabel 2.2 Confusion Matrix	24
Tabel 3.1 Variabel Definisi Operasional	30
Tabel 4.1 Contoh Stop Word Removal	42
Tabel 4.2 Contoh Stemming	43
Tabel 4.3 Contoh Punctuation Marks	44
Tabel 4.4 Perhitungan Skor Sentimen	45
Tabel 4.5 Hasil Label Kelas Sentimen	45
Tabel 4.6 Perbandingan Data Latih dan Data Uji	46
Tabel 4.7 Akurasi SVM	47
Tabel 4.8 Akurasi Naïve Bayes	47
Tabel 4.9 Perbandingan Akurasi antara Metode SVM dan Naïve Bayes	48
Tabel 4.10 Hasil Confusion Matrix	49

DAFTAR GAMBAR

	Halaman
Gambar 1.1 Capture Aplikasi KAI Access	2
Gambar 1.2 Grafik Pengguna Kereta Api Mengetahui Adanya KAI Access	3
Gambar 1.3 Perbandingan Penelitian Terdahulu	4
Gambar 2.1 Visualiasi SVM	19
Gambar 2.2 Train dan Testing Set	24
Gambar 2.2 Train dan Testing Set	26
Gambar 3.1 Script Library Google Play	31
Gambar 3.2 Script Crawling Data	31
Gambar 3.3 Kerangka Pemikiran	33
Gambar 3.4 Flowchart Penelitian	34
Gambar 3.5 Tampilan Login	35
Gambar 3.6 Tampilan Dashboard	35
Gambar 3.7 Tampilan proses Preprocessing	36
Gambar 3.8 Tampilan Data Uji	36
Gambar 3.9 Tampilan Klasifikasi SVM	37
Gambar 3.10 Tampilan Klasifikasi Naïve Bayes	37
Gambar 3.11 Tampilan Visualisasi SVM	38
Gambar 3.12 Tampilan Visualisasi Naïve Bayes	38
Gambar 3.13 Tampilan Profile	39
Gambar 4.1 Flowchart Preprocessing	40
Gambar 4.2 Flowchart Stop Word Removal	41
Gambar 4.3 Flowchart Stemming	42
Gambar 4.4 Flowchart Punctuation Marks	43
Gambar 4.5 Tampilan Login	51
Gambar 4.6 Tampilan Dashboard	51
Gambar 4.7 Tampilan Data Latih	52
Gambar 4.8 Tampilan Data Uji	52
Gambar 4.9 Tampilan Klasifikasi SVM	53
Gambar 4.10 Tampilan Klasifikasi Naïve Bayes	53
Gambar 4.11 Tampilan Visualsasi SVM	54

Gambar 4.12 Tampilan Visualisasi Naïve Bayes 54
Gambar 4.13 Tampilan Profile 55



DAFTAR LAMPIRAN

1. Hasil Penelitian Kuesioner
2. Plagiarisme
3. Biodata Mahasiswa
4. Kartu Bimbingan

