
PERBANDINGAN ALGORITMA SUPPORT VECTOR MACHINE DENGAN DECISION TREE PADA APLIKASI RUANG GURU

Indi Nurul Hassanah¹, Sutan Faisal², Amril Mutoi Siregar³

Universitas Buana Perjuangan Karawang

Jl. HS.Ronggo Waluyo, Puseurjaya, Telukjambe Timur, Karawang, Jawa Barat 41361

¹f19.indihassanah@mhs.ubpkarawang.ac.id, ²sutanfaisal@ubpkarawang.ac.id,

³amrilmutoi@ubpkarawang.ac.id

Abstract

Mobile Learning is electronic-based learning using a computer or computer-based. One of the most widely known Mobile Learning applications today is Ruang Guru. One way to determine the success of an application is to do a sentiment analysis of the application. The purpose of this study was to find the best accuracy model for classifying data in the SVM and Decision Tree algorithms. The data is taken from the comments column in the playstore on the Ruang Guru application as much as 1500 data. Then the data is labeled into 2 classes, namely positive and negative. After that, the data is divided into 70% training data and 30% testing data. The results of the comparison show that the best test model for sentiment classification cases is found in the SVM algorithm with an accuracy value of 84.2%, while the Decision Tree algorithm gets an accuracy value of 70%. So it can be concluded that the SVM algorithm has a better value for classification of review data in the Ruang Guru application compared to the Decision Tree algorithm.

Keywords: Ruang Guru, sentiment analysis, SVM, Decision Tree

Abstrak

Mobile Learning merupakan pembelajaran berbasis elektronik dengan menggunakan komputer atau berbasis komputer. Salah satu aplikasi Mobile Learning yang banyak dikenal saat ini adalah Ruang Guru. Salah satu cara untuk mengetahui keberhasilan suatu aplikasi adalah dengan melakukan analisis sentimen terhadap aplikasi tersebut. Tujuan penelitian ini untuk menemukan pemodelan akurasi terbaik terhadap pengklasifikasian data pada algoritma SVM dan Decision Tree. Data diambil dari kolom komentar di playstore pada aplikasi Ruang Guru sebanyak 1500 data. Kemudian data tersebut dilabelkan menjadi 2 kelas yaitu positif dan negatif. Setelah itu, data dibagi 2 menjadi data training sebanyak 70% dan data testing 30%. Hasil perbandingan menunjukkan model uji terbaik untuk kasus klasifikasi sentimen terdapat pada algoritma SVM dengan nilai akurasi sebesar 84.2% sedangkan pada algoritma Decision Tree mendapatkan nilai akurasi sebesar 70%. Sehingga dapat disimpulkan bahwa algoritma SVM memiliki nilai yang lebih baik untuk klasifikasi data ulasan pada aplikasi Ruang Guru dibandingkan algoritma Decision Tree.

Kata kunci: Ruang Guru, analisis sentimen, SVM, Decision Tree

1. PENDAHULUAN

Perkembangan teknologi dan informasi terus berkembang dari waktu ke waktu. Pengaruh perkembangan tersebut dialami oleh salah satu bidang yaitu pada bidang pendidikan. Dimana salah satu peran teknologi dalam bidang pendidikan yaitu *Mobile Learning*. Dengan adanya *Mobile Learning* maka proses belajar akan menjadi lebih mudah, seperti waktu yang tidak terikat, tidak terbatas ruang dan waktu, dan dapat diakses dari mana saja[1]. Cara mengajar seorang guru menjadi tolak ukur dalam menentukan prestasi belajar siswa. Tetapi, faktanya yang terlihat saat ini banyak siswa yang berprestasi namun mereka tidak dapat mengembangkannya karena kurang tertarik dengan cara gurunya mengajar.

Dengan menggunakan pembelajaran yang menarik dan menghadirkan hal-hal baru, siswa akan lebih tertarik dan menyukai pelajaran sehingga membuat siswa lebih tertantang untuk mencapai prestasinya. Hadirnya teknologi yang terus berkembang menjadi solusi bagi guru dan siswa untuk berkomunikasi dalam dunia pendidikan dan keuntungan lainnya yang dapat dinikmati oleh setiap siswa yang dimana faktor teknologi menjadi pengaruh prestasi belajar siswa. Dengan kehadiran teknologi seperti media *online* sebagai alat bantu pembelajaran karena mudah dalam mencari informasi dan materi-materi tambahan yang tidak didapati di sekolah.

Media teknologi yang sering digunakan pada saat ini adalah aplikasi *Mobile Learning* yang ada di *PlayStore*, aplikasi tersebut memiliki tingkat penggunaan dan jumlah unduhan yang tinggi. Salah satu aplikasi *Mobile Learning* yang ada di *PlayStore* dan sedang tren di kalangan masyarakat pada saat ini adalah Ruang Guru. Ruang Guru adalah aplikasi yang mengembangkan berbagai layanan belajar yang berbasis teknologi seperti kelas virtual, les privat, video belajar dan konten-konten lainnya yang bersifat pendidikan. Aplikasi Ruang Guru ini adalah produk dari PT. Ruang Raya Indonesia yang resmi berdiri pada tahun 2014, merupakan teknologi terbesar dan terlengkap di Indonesia yang berfokus pada layanan berbasis pendidikan[2]. Tujuan aplikasi Ruang Guru ini untuk membantu siswa, guru maupun orang tua dalam mengakses pembelajaran dimanapun mereka berada, sehingga mereka dapat menjalankan aktivitasnya menjadi lebih efektif dan efisien.

Suatu aplikasi selalu memiliki kekurangan dan kelebihan masing-masing, dimana hal tersebut dapat menimbulkan berbagai respon dari pengguna aplikasi seperti kepuasan dan kekecewaan terhadap aplikasi tersebut. Pada kolom ulasan komentar menjadi salah satu tempat untuk melontarkan kepuasan dan kekecewaan masyarakat terhadap aplikasi tersebut[3]. Banyaknya aplikasi pembelajaran online, sehingga mengharuskan Ruang Guru untuk terus meningkat dan memperbaiki kualitas performanya agar dapat mencapai kepuasan pengguna dan aplikasinya dapat bermanfaat untuk pengguna [4].

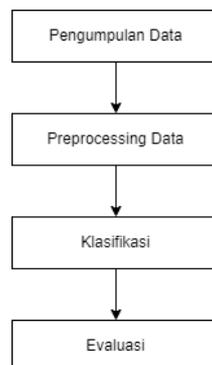
Penelitian terdahulu terkait sentimen dilakukan oleh [5] mengenai Analisis Sentimen Aplikasi Pembelajaran *Online* di *Playstore* Pada Masa Pandemi Covid-19 Menggunakan Algoritma *Support Vector Machine* (SVM) dengan pengumpulan data berjumlah 1500 ulasan yang mendapatkan hasil dari proses analisis data menunjukkan tingkat keakurasian sebesar 83,33%. Penelitian selanjutnya dilakukan oleh [6] yang mendapatkan hasil pada perhitungan akurasi menunjukkan bahwa metode *Support Vector Machine* lebih unggul dengan nilai 89,70% tanpa *K-Fold Cross Validation* dan 88,76% dengan *K-Fold Cross Validation* sedangkan pada perhitungan

waktu proses metode *K-Nearest Neighbor* lebih unggul dengan waktu proses 0.0160s tanpa *K-Fold Cross Validation* dan 0.1505s dengan *K-Fold Cross Validation*. Penelitian selanjutnya yaitu Komparasi Algoritma C4.5, *Naïve Bayes* Dan *Neural Network* Untuk Klasifikasi Tanah yang dilakukan oleh [7] Hasil dari penelitian ini didapatkan algoritma terbaik yaitu Algoritma C4.5. Algoritma C4.5 dalam klasifikasi dua kelas mencapai akurasi 98,45% dan AUC 0,981. Dalam klasifikasi tiga kelas C4.5 juga mencapai akurasi tertinggi (93,21%), demikian juga pada klasifikasi tujuh kelas (83,40%).

Diantara metode data mining SVM dan *Decision Tree* dapat digunakan untuk mengklasifikasi sekumpulan data dengan mengidentifikasi pola dari input data. Penulis menggunakan metode SVM ini karena memiliki kelebihan menemukan *hyperplane* terbaik yang memisahkan dua buah *class* pada input *space* dan mencari titik maksimalnya. Selain itu kelebihan lain dari metode SVM adalah pada proses *learning* yang cepat. Sedangkan pada metode *Decision Tree* memiliki salah satu kelebihan yaitu daerah pengambilan keputusan yang sebelumnya kompleks dan sangat global, dapat diubah menjadi lebih *simple* dan spesifik. Maka dari itu, dalam penelitian ini akan dilihat perbandingan dari hasil klasifikasi SVM dan *Decision Tree* mana yang memiliki performa serta tingkat akurasi terbaik dalam menganalisis sentimen pada aplikasi Ruang Guru.

2. METODOLOGI PENELITIAN

Metode penelitian yang akan dilakukan terdiri dari beberapa tahap yaitu pengumpulan data, *Preprocessing*, klasifikasi menggunakan algoritma *Support Vector Machine* dan *Decision Tree* dan yang terakhir evaluasi. Untuk lebih jelasnya bisa dilihat pada gambar berikut :



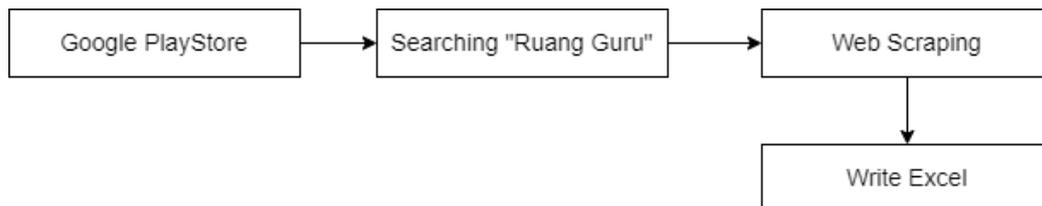
Gambar 1. Metode Penelitian

2.1 Pengumpulan Data

Tahap ini meliputi pengumpulan data yang terdiri dari langkah pertama untuk mendapatkan aplikasi Ruang Guru berdasarkan jumlah unduhan dan rating pada *Google Playstore* yaitu dengan parameter *query* "Ruang Guru". Setelah mendapatkan aplikasi Ruang Guru, langkah selanjutnya yaitu mendapatkan data opini dalam bentuk ulasan komentar yang diberikan pengguna terhadap aplikasi Ruang Guru yang ada di *Google Playstore*. Data berasal dari *Google Playstore* dengan

menggunakan teknik *Web Scraping* dengan tools *Google Colab* untuk mendapatkan teks ulasan, pemberi ulasan dan waktu ulasan tersebut diberikan. Setelah itu, data dilabelkan manual menggunakan aplikasi *Excel* untuk mendapatkan dua kelas sentimen yaitu Positif dan Negatif. Hasil *Scraping* mendapatkan sebanyak 1500 data ulasan.

Berikut adalah gambar tahapan untuk pengumpulan data :



Gambar 2. Pengumpulan Data

2.2. *Preprocessing* Data

Preprocessing Data adalah proses perubahan data mentah menjadi data yang bisa di proses untuk pelatihan dan pengujian pada saat pemodelan data. Selain itu *preprocessing* data dilakukan untuk melakukan pembersihan data agar data tersebut siap untuk pemodelan data. Berikut tahap *Preprocessing* Data :

- a. *Case Folding*, adalah salah satu bentuk text *preprocessing* yang paling sederhana dan efektif. Berikut macam-macam *case folding* :
 - 1) *Remove HTML* : untuk menghapus *HTML* yang terdapat dalam teks.
 - 2) *Annotation Removal* : untuk menghilangkan tanda @ (*Annotation*) yang terdapat dalam teks.
 - 3) *Remove Number* : untuk menghapus angka yang terdapat dalam teks.
 - 4) *Transform Case* : Untuk mengubah huruf besar atau kapital menjadi huruf kecil dan sebaliknya.
- b. *Tokenization*, adalah proses untuk melakukan pemecahan sekumpulan kalimat menjadi potongan karakter atau kata-kata sesuai kebutuhan yang sering disebut token, sehingga menjadi kata yang memiliki arti tertentu [8].
- c. *Stemming*, adalah proses untuk menghilangkan imbuhan yang terdapat pada masing-masing kata sehingga menjadi kata dasar, selain itu bertujuan untuk membersihkan suatu kata dengan pengejaan yang kurang tepat.
- d. *Filtering*, adalah tahap untuk mengambil kata-kata penting dari hasil token dengan menggunakan algoritma *stoplist* (membuang kata kurang penting) atau *wordlist* (menyimpan kata penting). Berikut macam-macam *filtering* :
 - 1) *Filter Stopwords* : untuk menghilangkan kata-kata yang tidak mempunyai arti yang biasanya merupakan kata sambung.
 - 2) *Filter Tokens (By Length)* : untuk menghilangkan kata dengan panjang huruf tertentu.
- e. *TF-IDF*, adalah proses untuk mengetahui frekuensi dari istilah tertentu yang relatif terhadap sebuah kata dalam kumpulan dokumen dan melihat seberapa umum atau tidak umum sebuah kata yang ada diantara *corpus* (sekumpulan teks yang terstruktur) atau memberikan bobot pada setiap kata yang ada menggunakan *TF-IDF* [9]. Rumus *TF-IDF* sebagai berikut :

$$TF(t) = f_{t,d} / \sum_t d$$

Dimana $f_{t,d}$ adalah frekuensi sebuah kata dalam dokumen, $\sum_t d$ adalah total kata yang terdapat didalam dokumen. Kemudian hitung IDF (*Inverse Document Frequency*) dari dokumentasi menggunakan persamaan berikut :

$$IDF(t) = \log(|D| / f_{t,D})$$

$|D|$ adalah jumlah dokumen dalam koleksi, sedangkan $f_{t,D}$ adalah jumlah dokumen dimana t ada di D . Setelah nilai TF dan ID didapatkan, langkah selanjutnya adalah menghitung TF-IDF menggunakan persamaan berikut :

$$TF-IDF = TF(t) * IDF(t)$$

2.3. Klasifikasi

Tahapan klasifikasi merupakan proses perhitungan yang menerapkan algoritma data mining untuk mengolah dataset ulasan dari aplikasi Ruang Guru pada *Playstore*. Data yang telah diambil dari *Google Playstore*, kemudian melalui tahap *preprocessing* dimana data tersebut akan dibagi menjadi dua bagian yaitu data *training* dan data *testing*. Algoritma yang digunakan yaitu SVM dan *Decision Tree* dibantu dengan pembobotan TF-IDF.

2.3.1. Support Vector Machine

Klasifikasi SVM adalah metode yang memaksimalkan batas *hyperplane* untuk memisahkan suatu set data dengan cara menghitung jarak margin dengan data terdekat dari masing-masing kelas. SVM standar mengambil kumpulan data input, dan memprediksi, untuk setiap input yang diberikan, mungkin input adalah anggota salah satu kelas dari dua kelas yang ada, yang membuat SVM sebagai a pengklasifikasi nonprobabilitas linier biner. Karena SVM adalah classifier, maka diberikan pelatihan set, masing-masing ditandai sebagai milik salah satu dari dua kategori, konstruksi algoritma pelatihan SVM model yang memprediksi apakah data baru termasuk dalam kategori atau lainnya[10]. Kelebihan dari metode ini yaitu mempunyai kemampuan generalisasi yang tinggi, mampu menghasilkan model klasifikasi yang baik meskipun dilatih dengan himpunan data yang relatif sedikit hanya dengan pengaturan parameter yang sederhana. Pada penelitian ini penulis menggunakan rumus *linear*. Berikut merupakan persamaan beberapa kernel yang dapat dilihat pada tabel dibawah ini :

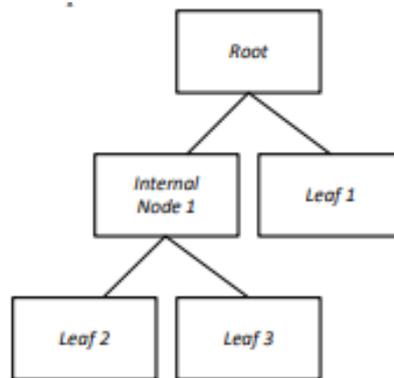
Tabel 1. Persamaan Kernel

Kernel	Persamaan
<i>Linear</i>	$K(x_i, x) = x_i^T x$
<i>Polynomial</i>	$K(x, x_i) = 1 + \sum (x * x_i)^d$

RBF	$K(x, x_i) = \exp(-\gamma \sum (x - x_i)^2)$
Sigmoid	$K(x, x_i) = \tanh(y(x_i^T x) + r)$

2.3.2. Decision Tree

Decision Tree adalah algoritma klasifikasi data yang menggunakan pohon keputusan [11]. Salah satu teknik klasifikasi terhadap objek atau *record*. Teknik ini terdiri dari kumpulan *decision node*, dan dihubungkan oleh cabang, bergerak ke bawah dari *root node* sampai berakhir di *leaf node*[12].



Gambar 3. Arsitektur Pohon Keputusan

Arsitektur pohon keputusan dibuat sedemikian rupa agar menyerupai pohon asli, dimana terdapat beberapa bagian yaitu :

- Root Node* : *Node* ini terletak pada bagian paling atas dari pohon keputusan.
- Internal Node* : *Node* ini merupakan percabangan dimana membutuhkan satu input dan mengeluarkan maksimal dua *output*.
- Leaf Node* : *Node* ini merupakan *node* yang terletak pada ujung pohon. *Node* ini hanya memiliki satu *input* dan tidak memiliki *output*.

Algoritma *Decision Tree* memiliki 2 input yaitu sample latih dan sample. Sample ini nantinya akan dipakai dalam membangun *tree* yang telah diproses keahliannya, kemudian sample digunakan untuk klasifikasi data. Berikut merupakan rumus dari *decision tree*.

$$Entropy(A) = \sum_i^n -p_i \log_2 p_i$$

2.4. Evaluasi

Pada tahap ini bertujuan untuk mengukur hasil pemodelan dengan membandingkan hasil pemodelan dari pemodelan yang berupa *accuracy*, *presicion*, *recall*, spesifisitas dan *F1-Score* dengan menggunakan algoritma SVM dan *Decision Tree*. Pengukuran akurasi ini menggunakan *confusion matrix*, berikut pengukuran evaluasi dengan menggunakan *confusion matrix* :

Tabel 2. Rumus *Confusion Matrix*

Data Aktual	Data Prediksi		
	TRUE	FALSE	TOTAL
TRUE	TP	FN	P
FALSE	FP	TN	N
TOTAL	P'	N'	P + N

Keterangan :

TP (*True Positive*) = Data positif diklasifikasikan secara benar

TN (*True Negative*) = Data negatif diklasifikasikan secara benar

FP (*False Positive*) = Data negatif diklasifikasikan menjadi positif

FN (*False Negative*) = Data positif diklasifikasikan menjadi negatif

3. HASIL DAN PEMBAHASAN

3.1. Pengumpulan Data

Data yang diambil pada *Google Playstore* merupakan ulasan pengguna pada aplikasi Ruang Guru. Pengambilan data menggunakan teknik *Scraping* dengan memanfaatkan *Google Colab*. Pada penelitian ini *Google Colab* digunakan untuk mengambil data ulasan dan mengolah data menggunakan Bahasa pemrograman *python* dengan jumlah data yang didapatkan sebanyak 1500 data ulasan terbaru yang didapatkan. Berikut hasil *Scraping* data pada Gambar 4.

username	score	at	content
Gracia Sak	5	01/11/2022 07:31	Ruangguru membuat aku lebih semangat untuk belajar Terima kasih ruangguru 🙏🙏🙏
wang senj	4	01/11/2022 05:33	Video sering error
Nedick Be	1	01/11/2022 09:55	Maaf ya ruang guru soal nya ada biaya nya jadi nya saya hapus deh aplikasi ini
Luh Cinta	2	01/11/2022 09:55	Apakah aplikasi ini membayar 🙏🙏
Ayi Barabi	5	01/11/2022 09:31	Ini aplikasinya bagus banget
Audina cit	5	01/11/2022 02:16	Pembelajaran materi lengkap, lumayan buat nyicil utbk Moga aja bisa diterima UGM
Nada Mall	5	01/11/2022 00:50	Banyak bintang
Hayfaa Az	5	31/10/2022 23:01	Wahh terimakasih ya ruangguru atas bantuannya, sekarang aku jadi paham pelajaran yg nilaiku rendah semua di rapor... 🙏🙏🙏🙏🙏
Zerafin Pi	3	31/10/2022 14:49	Aplikasi ini sangat bagus untuk siswa, namun saat sy bertanya kepada orang-orang yang memberi diskon untuk ruangguru, mereka tidak mau menjawab dan, ad
Reskiani K	4	31/10/2022 14:35	Kak kenapa fitur rangkuman tidak bisa muncul gambarnya.... Padahal aplikasinya udah fitur terbaru & saya sudah berlangganan
Bayu Indri	1	31/10/2022 14:25	Ada dari tim official katanya, chat malam2, "sibuk ya dek kok belum balas chatnya" at 22:20 wth, mau research ke anak-anak segitu larut? BEWARE untuk orang tu
rachelia vi	5	31/10/2022 14:08	Aku suka sama ruangguru ini karena aku dapat juara 1 di sekolah
Haura Ahc	1	31/10/2022 13:20	Jd lama loadingnya terutama klo liat screenshot sm rangkuman persub-bab.
Banugian	1	31/10/2022 13:17	Seperti biasa, selalu error dan pet selalu sekarat walau belajar lewat 10 menit
Kasmi Ksn	5	31/10/2022 13:10	Ksami
Nur Asyiki	5	31/10/2022 12:51	bagus banget apk nya, jadi bikin tambah semangat belajarr
Lukmanul	5	31/10/2022 12:30	Bagus sekali lengkap pembahasannya🙏🙏
itz ra	4	31/10/2022 12:01	Min, jadi aku kan jarang buka rung guru gitu kan, padahal aku udah berlangganan 1 tahun, tadi ruangguru aku buka lagi, tp kenapa malah disuruh ngisi ulang data
Reza Widj	2	31/10/2022 11:57	Ruangguru moment video 12:21/00:00
Atya Putri	1	31/10/2022 10:53	Memenuhi penyimpanan ja nih
Nvy Nasy	5	31/10/2022 08:45	Bagus, membantu saya mencari ilmu. Kadang aplikasinya loadingnya lama nih, padahal jaringannya sudah ok, gak ada kendala, dan biasanya loading lama cuman a
20. Satya f	3	31/10/2022 06:22	ruang guru sangat membantu sih ,tapi sy terus menonton vidionya kucing sy tidak menambah energi nya tolong di bantu

Gambar 4. Dataset Hasil Scraping

3.2. Preprocessing

Data yang sudah didapatkan pada tahap Scapping akan diolah menjadi data terstruktur atau biasa disebut *Preprocessing*. Berikut merupakan tahapan *Preprocessing* yang dilakukan :

a. Case Folding

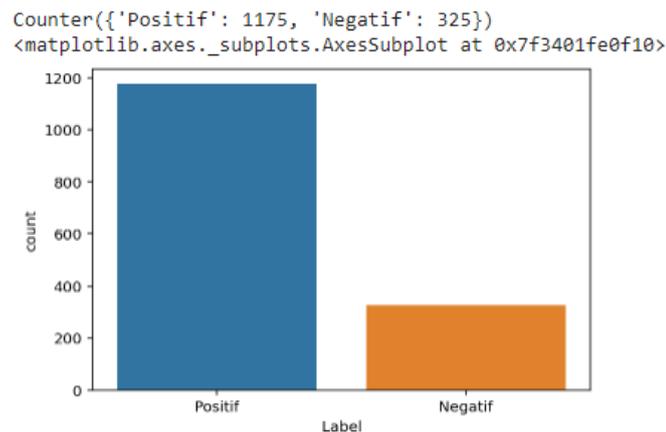
Merupakan proses untuk mengubah semua huruf dalam dokumen dari huruf besar menjadi huruf kecil ataupun sebaliknya, selain itu proses *case folding* juga dapat menghapus angka, tanda baca, dan karakter kosong yang ada di dalam dokumen. Berikut merupakan proses dari *case folding*.

(0, 2187)	0.18703180757684257
(0, 1711)	0.3070822237223931
(0, 1450)	0.21986143014760667
(0, 1285)	0.44667234315439747
(0, 1192)	0.23705608521293875
(0, 469)	0.16018453073711003
(0, 429)	0.15925248473108844
(0, 426)	0.24487822870694717
(0, 362)	0.2932032635879284
(0, 232)	0.6080418194766017
(1, 813)	1.0
(2, 2090)	0.4967798731153253
(2, 1949)	0.30325074384857287
(2, 1194)	0.34870711275880373
(2, 679)	0.629785401150418
(2, 521)	0.3781830800446189
(3, 790)	0.33346443377324886

Gambar 5. Hasil TF-IDF

3.3. Klasifikasi

Pada tahapan klasifikasi diawali menggunakan pembobotan TF-IDF dengan pengelompokkan label. Setelah pembobotan TF-IDF selesai maka data dibagi dua yaitu data *training* dan data *testing*. Klasifikasi dilakukan di *Google Colab* dengan Bahasa pemrograman *Python* menggunakan algoritma SVM dan *Decision Tree*. Dimana pada tahap klasifikasi ini dibagi menjadi 2 kelas, yaitu positif dan negatif. Dataset dibagi menjadi data *training* sebanyak 1050 dan data *testing* sebanyak 450 menggunakan ratio 0.3. Artinya, 70% data digunakan untuk proses *training* dan 30% data untuk proses *testing*. Berikut merupakan hasil kelas sentimen positif dan negatif.



Gambar 6. Hasil Klasifikasi Sentimen Positif dan Negatif

3.4. Evaluasi

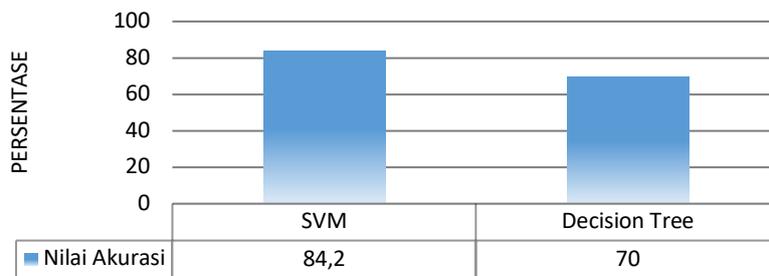
Pada tahap ini merupakan hasil pengujian dengan menggunakan algoritma SVM dan *Decision Tree*.

Tabel 7. Hasil *Confusion Matrix*

Metode	TP	FP	TN	FN
SVM	50	58	329	13
Decision tree	18	76	301	55

Tabel diatas menampilkan hasil proses *Confusion Matrix* dari setiap metode. Untuk grafik hasil perbandingan dari setiap metode dapat dilihat pada Gambar 7.

PERBANDINGAN NILAI AKURASI



Gambar 7. Perbandingan Nilai Akurasi

Berdasarkan Gambar diatas menjelaskan bahwa nilai perbandingan akurasi terbaik terdapat pada metode SVM dengan nilai 84.2%, sedangkan pada metode *Decision Tree* hanya mendapat nilai akurasi 70%. Untuk hasil Akurasi, *Precision* dan *Recall* dapat ditunjukkan pada Tabel 8.

Tabel 8. Hasil Perbandingan Seluruh Metode

No	Algoritma	Akurasi(%)	<i>Precision</i> (%)	<i>Recall</i> (%)
1	SVM	84.2	0.85	0.96
2	<i>Decision Tree</i>	70.0	0.80	0.85
	Selisih	14.2	0.05	0.11

Berdasarkan tabel tersebut dapat diketahui bahwa hasil evaluasi dengan menggunakan algoritma SVM setelah dibagi menjadi data *training* dan data *testing* mendapatkan nilai akurasi sebesar 84.2%, *precision* 85% dan *recall* 96%. Sedangkan pada algoritma *decision tree* mendapatkan nilai akurasi 70%, *precision* 80% dan *recall* 85%. Dan mendapatkan selisih nilai sebesar 14.2%.

4. SIMPULAN

Komparasi menggunakan algoritma SVM dan *Decision Tree* mendapatkan hasil terbaik yaitu pada algoritma SVM dikarenakan pada algoritma SVM melakukan perhitungan probabilitas pada data dengan pembobotan kata menggunakan TF-IDF mendapatkan hasil nilai akurasi sebesar 84.2%, sedangkan pada algoritma *Decision Tree* mendapatkan nilai akurasi sebesar 70%. Jumlah data yang digunakan untuk diklasifikasikan sebanyak 1500 data dari hasil *scrapping* pada *playstore*, kemudian data dibagi menjadi 2 yaitu data *training* sebanyak 1050 data dan data *testing* sebanyak 450 data. Sehingga dapat disimpulkan bahwa algoritma *Support Vector Machine* memiliki nilai yang lebih baik untuk klasifikasi data ulasan pada aplikasi Ruang Guru dibandingkan algoritma *Decision Tree*.

DAFTAR PUSTAKA

- [1] F. F. Irfani and S. A. Yogyakarta, "Analisis Sentimen Review Aplikasi Ruangguru Menggunakan Algoritma Support Vector Machine," no. February, 2020, doi: 10.26487/jbmi.v16i3.8607.
- [2] N. Manado and N. S. Londa, "Pengaruh komunikasi media".
- [3] A. P. Giovani, T. Haryanti, and L. Kurniawati, "Analisis Sentimen Aplikasi Ruang Guru Di Twitter Menggunakan Algoritma Klasifikasi," vol. 14, no. 2, pp. 116–124, 2020.
- [4] D. Normalisasi and K. Levenshtein, "Analisis sentimen data ulasan aplikasi ruangguru pada situs google play menggunakan algoritma naïve bayes classifier dengan normalisasi kata levenshtein distance," vol. 11, no. 1, pp. 257–266, 2022.
- [5] A. Erfina *et al.*, "Analisis Sentimen Aplikasi Pembelajaran Online Di Play Store Pada Masa Pandemi Covid-19 Menggunakan Algoritma Support Vector Machine," vol. 2020, no. Semasif, pp. 145–152, 2020.
- [6] M. Rangga, A. Nasution, and M. Hayaty, "Perbandingan Akurasi dan Waktu Proses Algoritma K-NN dan SVM dalam Analisis Sentimen Twitter," vol. 6, no. 2, pp. 226–235, 2019.
- [7] A. Mukminin and D. Riana, "Komparasi Algoritma C4 . 5 , Naïve Bayes Dan Neural Network Untuk Klasifikasi Tanah," vol. 4, no. 1, pp. 21–31, 2017.
- [8] A. S. Rahayu and A. Fauzi, "Komparasi Algoritma Naïve Bayes Dan Support Vector Machine (SVM) Pada Analisis Sentimen Spotify," vol. 4, pp. 349–354, 2022, doi: 10.30865/json.v4i2.5398.
- [9] I. P. Rahayu, A. Fauzi, and J. Indra, "Analisis Sentimen Terhadap Program Kampus Merdeka Menggunakan Naive Bayes Dan Support Vector Machine," vol. 4, pp. 296–301, 2022, doi: 10.30865/json.v4i2.5381.
- [10] A. M. Siregar, S. Faisal, A. Puspabhuana, and M. S. H. Simarankir, "Comparison Study Of Term Weighting Optimally With SVM In Sentiment Analysis," 2019, doi: 10.4108/eai.18-7-2019.2288508.
- [11] A. Tobby Wiratama Putra , Agung Triayudi, "Analisis Sentimen Pembelajaran Daring menggunakan Metode Naïve Bayes , KNN , dan Decision Tree," vol. 6,

- no. 1, 2022.
- [12] R. P. P. S. Abdul Najib, Damar Nurcahyono, "Klasifikasi Diagnosa Penyakit Diabetes Mellitus (DM)," no. Dm.