

Analisis Sentimen *Review* Masyarakat Terhadap Pelayanan Rumah Sakit Umum Daerah Dr.Moewardi Surakarta Pada Situs Google Maps Menggunakan Metode *Support Vector Machine*, *Random Forest*, Dan *Term Frequency – Inverse Document Frequency*

Abstrak

Rumah Sakit Umum Daerah (RSUD) Dr.Moewardi Surakarta merupakan sebuah institusi pelayanan Kesehatan di bawah naungan pemerintah daerah. Pelayanan yang di berikan oleh RSUD dituntut untuk senantiasa melakukan perubahan, agar pelayanan dapat sesuai dengan standar yang telah ditentukan sehingga mampu memberikan pelayanan kesehatan yang berkualitas bagi semua orang. Oleh karena itu, penelitian ini dilakukan untuk menganalisis sentimen opini masyarakat terhadap pelayanan RSUD Dr.Moewardi Surakarta pada situs Google Maps guna melihat performa pelayanan rumah sakit. Analisis ini menggunakan metode *Support Vector Machine* dan *Random Forest* dengan pembobotan kata menggunakan *Term Frequency-Inverse Document Frequency*. Hasil pengujian menggunakan Cross Validation dengan data latih sebanyak 714 dan data uji sebanyak 179 menjelaskan jika metode SVM lebih efektif untuk melakukan klasifikasi pada data ini dengan besaran nilai akurasi 92%, recall sebesar 100%, precision sebesar 92%, dan f-measure sebesar 96%. Nilai tersebut didapat pada saat menggunakan parameter $C=1$ dan $Itermax=1000$.

Kata kunci: analisis sentimen, *support vector machine*, TF-IDF, RSUD

Pendahuluan

Keberadaan rumah sakit negeri sebagai Lembaga pelayanan di bidang kesehatan menjadi tempat bergantung masyarakat. Seluruh layanan di rumah sakit baik dari Unit Gawat Darurat hingga poli khusus penyakit dalam menjadi tempat acuan untuk mengendalikan kesehatan dan kesejahteraan seluruh masyarakat Indonesia. Rumah Sakit negeri terkhusus Rumah Sakit Umum Daerah (RSUD) yang berada di bawah naungan pemerintah diharuskan untuk selalu berubah dari setiap waktu guna memenuhi kebutuhan serta kepentingan masyarakat. Sebagai sebuah instansi yang menjadi acuan utama pada tingkat daerah, RSUD dituntut untuk memenuhi standar

yang ditetapkan untuk selalu memberikan pelayanan terbaik bagi masyarakat. Pelayanan yang berkualitas kepada publik terutama pasien menjadi prioritas bagi keberlangsungan RSUD. Hal ini diatur pada UU No. 36 dan No. 44 tahun 2009 mengenai kesehatan dan Rumah sakit, secara substansial dijelaskan apabila pelayanan kesehatan rumah sakit harus dilaksanakan dengan rasa tanggung jawab, aman, bermutu, adil, dan dilaksanakan dengan standar mutu pelayanan pemerintah.

Salah satu RSUD yang sering menjadi rujukan di daerah Jawa Tengah ialah RSUD Dr. Moewardi Surakarta. RSUD dr. Moewardi merupakan Rumah Sakit Umum Daerah di bawah naungan pemerintah bertaraf nasional yang menyediakan beberapa pelayanan seperti UGD, operasi, rawat inap, dan lain – lain. RSUD Dr. Moewardi juga menawarkan lebih dari 40 pelayanan poli spesialis yang dapat dimanfaatkan oleh masyarakat. Sebagai RSUD, rumah sakit dr. Moewardi juga dituntut untuk memenuhi standar mutu pelayanan yang telah ditetapkan pemerintah untuk memberikan pelayanan terbaik bagi masyarakat.

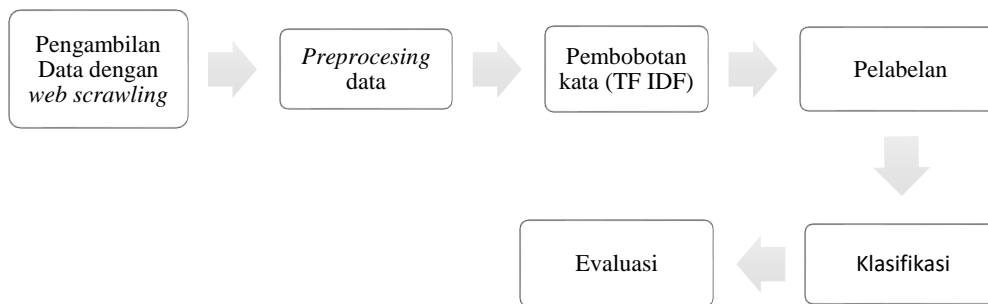
Analisis sentimen merupakan suatu analisis yang digunakan untuk mengetahui pandangan publik terhadap suatu hal baik fenomena, manusia, maupun tempat. Analisis sentimen menitikberatkan pada pandangan emosi atas komentar teks yang bertujuan untuk menganalisis emosi dan gambaran suasana pada masyarakat terhadap suatu permasalahan dengan tujuan untuk mengelompokkan polaritas suatu opini positif ataupun negatif (Aruan, dkk., 2022).

Metode yang akan digunakan dalam penelitian ini ialah SVM (*Support Vector Machine*) dan Random Forest. Penggunaan metode SVM dilatarbelakangi oleh ketinggian akurasi yang didapatkan berdasarkan pengelompokan data sesuatu atribut label, sehingga opini dapat terbagi menjadi beberapa kategori (Anjasmos, Istiadi, & Marisa, 2020). *Random Forest* digunakan sebagai salah satu metode pembandingan karena memiliki tingkat akurasi yang tinggi dan cocok digunakan pada klasifikasi data. Sebelum dilakukan klasifikasi diperlukan suatu pembobotan dalam setiap kata pada dokumen karena data yang dianalisis merupakan suatu data teks. Metode pembobotan kata yang akan digunakan merupakan bagian dari *Term Frequency-Inverse Document*

Frequency (TF-IDF). TF-IDF ialah metode yang berfungsi untuk melakukan pembobotan pada setiap kata hasil ekstraksi.

Penelitian ini dilakukan untuk menganalisa review pelayanan RSUD Dr. Moewardi Surakarta pada situs Google Maps dengan menggunakan metode *Support Vector Machine*, *Random Forest*, dan TF-IDF. Penggunaan dua metode klasifikasi pada penelitian ini bertujuan untuk mencari model terbaik untuk melakukan klasifikasi pada teks review secara lebih baik dan akurat. Hasil klasifikasi ini nantinya dapat menjadi acuan bagi RSUD Dr. Moewardi untuk meningkatkan kualitas pelayanan, menjaga popularitas di masyarakat, dan memperbaiki kekurangan serta mengevaluasi kembali kinerja ke arah yang lebih baik.

Metodologi Penelitian



Gambar 1. Alur Penelitian

A. Pengambilan Data

Data yang digunakan pada penelitian ini merupakan data primer yang diperoleh langsung dari sumber asli tanpa perantara. Data ini dikumpulkan secara khusus oleh peneliti untuk menyelesaikan permasalahan penelitian (Johnson dan Gupta, 2012). Data yang diambil merupakan 893 data review dengan kategori yang paling relevan dengan pelayanan RSUD dr. Moewardi Surakarta. Proses pengambilan data dari situs google maps menggunakan metode *Scrawling* (Penambangan Teks). Penambangan teks merupakan penggalian data berupa kata-kata yang diperoleh dari data tidak terstruktur pada kalimat-kalimat dalam suatu dokumen (Maulana dan Pratiwi, 2019). Penambangan teks dapat dilakukan dengan cara crawling data. Crawling adalah teknik yang digunakan untuk mengumpulkan data yang terdapat dalam suatu web.

B. Preprocessing

Preprocessing data dilakukan sebelum menganalisis suatu data. Sebelum melakukan analisis sentimen perlu adanya preprocessing text yang terdiri dari case folding, punctuation removal, stopwords removal, tokenization, dan stemming (Mining, 2020). Case holding adalah proses mengubah semua huruf menjadi bentuk yang sama, misal huruf kecil. Punctuation removal merupakan proses menghilangkan tanda baca. Stopwords removal untuk menghapus kata-kata yang tidak penting. Tokenization berupa pembagian kalimat menjadi beberapa bagian yang disebut token. Sedangkan stemming berupa proses mendapatkan kata dasar dengan menghilangkan imbuhan.

C. Pembobotan

TF-IDF (*Term Frequency-Inverse Document Frequency*) merupakan metode yang digunakan dalam pembobotan sebuah kata di dalam sistem pencarian informasi. Sistem kerja TF-IDF ialah menghitung nilai dari masing – masing kata di dalam teks menggunakan frekuensi kemunculan kata. Bobot pada *term* menunjukkan sejauh apa sebuah *term* mewakili makna dari satu dokumen. Semakin besar bobotnya maka akan semakin besar pengaruhnya dalam mewakili dokumen dan begitu sebaliknya (Karmayasa & Mahendra, 2010).

D. Pelabelan

Selisih skor jumlah kata positif dengan skor jumlah kata negatif dalam setiap kalimat ulasan adalah cara menentukan kelas sentimen. Jika Kalimat yang memiliki skor 1 → kelas positif sedangkan jika kalimat yang memiliki skor 0 → kelas negatif, Setelah pelabelan data maka selanjutnya data terbagi dua yaitu data latih dan data uji yang dibagi secara manual dengan cara mengacak keseluruhan data dari tiap masing-masing kelas.

E. Klasifikasi dengan *Support Vector Machine* (SVM) dan *Random Forest*

Support Vector Machine (SVM) adalah algoritma supervised learning yang digunakan untuk klasifikasi, regresi, dan deteksi anomali atau outlier. Dalam menerapkan model SVM ini membutuhkan suatu parameter. Di mana keakuratan

model sangat bergantung pada parameter tersebut, dan karenanya nilai optimal akan didapat dari teknik grid-search yang sesuai (Fitri dkk, 2019).

Random Forest merupakan algoritma ensemble learning yang pengklasifikasiannya berdasarkan hasil yang diperoleh dari sekumpulan decision tree (Nayak, 2016). Banyaknya pohon dapat mengurangi varians dalam model keseluruhan dan mengendalikan overfitting. Serta semakin banyak pohon di random forest semakin tinggi hasil akurasi yang diberikan (Fauzi, 2018).

F. Evaluasi

Jika data memiliki karakteristik balance data maka metrik evaluasi eror yang cocok adalah akurasi dan untuk imbalance data maka metrik evaluasi eror yang cocok adalah F1-score. Evaluasi eror ini juga menggunakan confusion matrix sebagai berikut:

Klasifikasi	Positif	Negatif
Positif	<i>True Positif</i> (TP)	<i>False Negatif</i> (FN)
Negatif	<i>False Positif</i> (FP)	<i>True Negatif</i> (TN)

Tingkat Akurasi dapat dihitung dengan persamaan :

$$Akurasi = \frac{TP + TN}{TP + TN + FP + FN}$$

Presisi adalah jumlah prediksi yang dibuat yang benar-benar relevan dari semua prediksi berdasarkan kelas positif. Semakin tinggi presisi maka akan mengembalikan hasil yang lebih relevan. Presisi dapat dihitung melalui persamaan berikut:

$$Presisi = \frac{TP}{TP + FP}$$

Berbeda dengan presisi, recall didefinisikan sebagai jumlah dari kelas positif yang diprediksi dengan benar yang dihitung melalui persamaan berikut:

$$Recall = \frac{TP}{TP + FN}$$

Selain itu, dalam mengevaluasi model juga bisa menggunakan atau Fmeasure yang bisa dihitung menggunakan persamaan berikut :

$$f_1 - score = 2 \times \frac{Recall * Presisi}{Recall + Presisi}$$

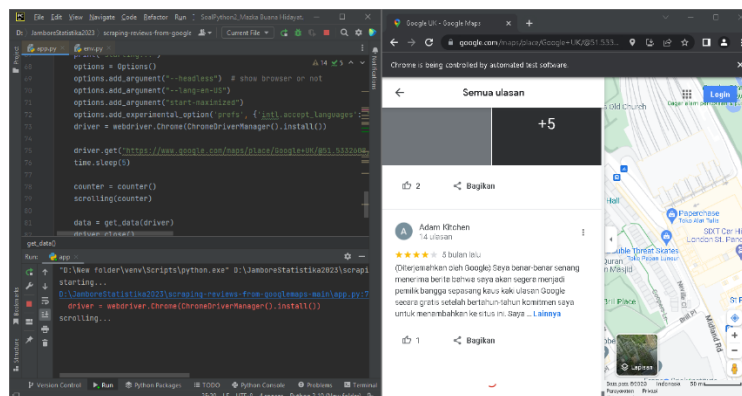
G. Hasil dan Visualisasi

Pada tahap ini dilakukan penarikan kesimpulan yang diperoleh berdasarkan hasil dari klasifikasi dengan menggunakan metode SVM dan *Random Forest*. Serta hasil dari proses visualisasi, metode Support Vector Machine dan *Random Forest* dalam penelitian untuk analisis sentimen review pada *Google Map* berdasarkan penelitian yang telah menghasilkan kesimpulan bahwa SVM *Optimized* menunjukkan kinerja terbaik klasifikasi analisis sentiment (Aruan, dkk., 2022).

Hasil Dan Pembahasan

A. Pengambilan Data

Pengambilan data ulasan pada Google Maps tentang review opini publik RSUD Moewardi menggunakan teknik *scrawling*. Bahasa pemrograman yang digunakan adalah Python menggunakan Library BeautifulSoup dan Selenium.



Gambar 2. Screenshot Crawling Google Maps

B. Preprocessing

- 1) Menghapus kolom yang tidak digunakan yaitu kolom yang tidak diperlukan, case folding(lowercase), dan remove emoji. Sehingga tersisa kolom yang akan digunakan untuk analisis selanjutnya.

```
[109] df.drop('author_id', inplace=True, axis=1)
```

	author_title	review_text	review_rating	review_datetime_utc	review_likes
0	Gylang Aryandaru	Rumah sakit umum daerah solo merupakan salah s...	4.0	02/25/2022 17:18:08	26.0
1	Abdurrahman Suparno	saya pasien rujukan bpjs dilayani cukup baik d...	4.0	08/19/2022 07:32:33	9.0
2	Ala Sandra	Pelayanan nya dibagian loket memang lama dan a...	4.0	07/13/2022 02:48:56	18.0
3	Anita Andriyana	Baikkkk bangeett puas bangeett disini, dar...	5.0	10/11/2022 01:09:55	1.0
4	Endraa	Administrasinya diperbaiki, minta resume medis...	1.0	07/21/2022 17:19:56	19.0
...
994	veronica wika	Rumah Sakit dengan pelayanan lengkap dan memua...	5.0	10/27/2020 02:04:56	0.0
995	bambang edy purwanto	maaf salah pencet...	1.0	05/25/2017 04:10:02	0.0
996	Tita Septiana	Bersih, pelayanan baik, terpadu	5.0	06/29/2021 05:57:30	0.0
997	Rina Puspita	Masjidnya dalam tahap renovasi.	5.0	09/05/2017 07:25:52	1.0
998	Tri Sunanto	Parkir gedung lantai 3	5.0	03/15/2018 12:29:46	0.0

993 rows x 5 columns

Gambar 3. Hapus Kolom Tidak Digunakan

- 2) Membersihkan data dengan menghapus tautan HTML dan emoji serta menghapus strip ruang putih seperti spasi, tab, dan baris baru yang berlebihan.

```
#hapus emoji
!pip install emot
import re
import pickle
from emoji import EMOJI_DATA
from emot.emo_unicode import UNICODE_EMOJI, EMOTICONS_EMO

import re
def remove_emoji(string):
    emoji_pattern = re.compile("["
        u"\U0001F600-\U0001F64F" # emoticons
        u"\U0001F300-\U0001F5FF" # symbols & pictographs
        u"\U0001F680-\U0001F6FF" # transport & map symbols
        u"\U0001F1E0-\U0001F1FF" # flags (iOS)
        u"\U00002702-\U000027B0"
        u"\U000024C2-\U0001F251"
        "]+", flags=re.UNICODE)
    return emoji_pattern.sub(r'', string)
df['review_text']=df['review_text'].apply(remove_emoji)
```

Gambar 4. Remove Emoji

- 3) Melakukan stopwords removal yaitu menghapus kata-kata yang tidak penting. Lalu dilakukan stemming atau pengubahan kata dasar pada kata berdasarkan library sastrawi.

```
# Importing stop words from NLTK corpus and word tokenizer
import nltk
nltk.download('punkt')
from nltk.corpus import stopwords
from nltk.tokenize import word_tokenize

df['review_token'] = df['review_text'].apply(lambda x: word_tokenize(x))

[nltk_data] Downloading package punkt to /root/nltk_data...
[nltk_data] Package punkt is already up-to-date

[123] import nltk
nltk.download('stopwords')
stop_words = set(stopwords.words('Indonesian'))

# Created new columns of tokens - where stop words are being removed
df['review_token_filtered'] = df['review_token'].apply(lambda x: [word for word in x if not word in stop_words])

## Tokens columns with stop words and without stop words
df[['review_token', 'review_token_filtered']].head(10)

[nltk_data] Downloading package stopwords to /root/nltk_data...
```

Gambar 5. StopWords & Tokenizing

```
[124] !pip install sastrawi
Looking in indexes: https://pypi.org/simple, https://us-python.pkg.dev/colab-wheels/public/simple/
Requirement already satisfied: sastrawi in /usr/local/lib/python3.8/dist-packages (1.0.1)

from Sastrawi.Stemmer.StemmerFactory import StemmerFactory

factory = StemmerFactory()
stemmer = factory.create_stemmer()

df['review_stemmed'] = df['review_token_filtered'].apply(lambda x: ' '.join([stemmer.stem(i) for i in x]))
df['review_stemmed'].head(10)

0    rumah sakit daerah solo salah rumah sakit rujuk...
1    pasien rujuk bpjs layanan daftar ribet extra sab...
2    layanan bagi dokter andhonom ontri panggil d...
3    baikkkk bangeeetttt puas bangeeetttt 2018 rawat ...
4    administrasi baik resume medis aja lempar kesa...
5    mertua masuk rumah sakit rawat ruang x rujuk p...
6    baik karena bingung sih parkir memang tunjuk j...
7    fasilitas daftar online cepet pegawai priu wan...
8    tolong tukang fotokopi lantai 1 ajar attitude b...
9    layanan cepat profesional memediasi keluhan terimakasih...
Name: review_stemmed, dtype: object
```

Gambar 6. Stemming

- 4) Menghapus data yang terduplikat dan menghapus nilai null.
- 5) Melakukan tokenisasi adalah proses untuk membagi teks yang dapat berupa kalimat, paragraf atau dokumen, menjadi token-token/bagian-bagian tertentu.

C. Pembobotan

Pembobotan dilakukan menggunakan TF-IDF yaitu metode yang memberikan bobot hubungan suatu kata terhadap suatu dokumen berdasarkan konsep frekuensi kemunculan kata di dalam sebuah dokumen dan inverse frekuensi dokumen yang mengandung kata tersebut.

Term Frequency dihitung menggunakan Persamaan dengan term frequency ke-i adalah frekuensi kemunculan term ke-i dalam dokumen ke-j. Sedangkan Inverse Document Frequency (IDF) adalah logaritma dari rasio jumlah seluruh dokumen dalam korpus dengan jumlah dokumen yang memiliki term yang dimaksud seperti yang dituliskan secara matematis pada Persamaan:

$$(tf - idf)_{ij} = tf_i(d_j) * idf_i$$

Fungsi metode TFIDF adalah untuk mencari representasi nilai dari tiap-tiap dokumen dari suatu kumpulan data training (training set) dimana nantinya dibentuk suatu vektor Antara dokumen dengan kata (documents with terms) yang kemudian untuk kesamaan antar dokumen dengan cluster akan ditentukan oleh sebuah prototype vektor yang disebut juga dengan cluster centroid

D. Pelabelan

Setelah melakukan *Crawling* data dari ulasan Google Maps, didapatkan 893 data dari rentang waktu 2017-2022. Terdapat 846 data berupa ulasan sentimen positif yang diberi label positif dan 47 data berupa ulasan sentimen negatif yang diberi label 0. Untuk proses *labelling* menggunakan model VADER (Valence Aware Dictionary for Sentiment Reasoning).

Model yang digunakan untuk analisis sentimen teks yang peka terhadap polaritas (positif/negatif) dan intensitas (kekuatan) emosi. Ini tersedia dalam library NLTK dan dapat diterapkan langsung ke data teks yang tidak berlabel. Analisis sentimental VADER bergantung pada kamus yang memetakan fitur lexicon ke intensitas emosional yang dikenal sebagai skor sentimen. Skor sentimen suatu teks dapat diperoleh dengan menjumlahkan intensitas setiap kata dalam teks.

Misalnya, kata-kata seperti 'cinta', 'nikmati', 'bahagia', 'suka' semuanya menyampaikan sentimen positif. Juga VADER cukup canggih untuk memahami konteks dasar dari kata-kata ini, seperti "tidak suka" sebagai pernyataan negatif. Itu juga memahami kapitalisasi dan penekanan tanda baca, seperti "NIKMATI".

Tabel 1. Ulasan Setelah *Labelling*

No	Ulasan	Label
1.	<ul style="list-style-type: none">• pendaftaran antrian via online useless pasa akhirnya antrian online tidak dipakai• jadwal jam 830 datang jam 825 baru dilayani jam setengah 12• overall rencana awal bersalin di rs ini untuk bulan juli dibatalkan coba rs lain yang diharapkan pelayanan administrasi jauh lebih baik	Negatif
2.	kemarin siang setelah ke puskesmas dan beberapa klinik di kawasan kemuning akhirnya saya melesat ke igd rsdm untuk mendapatkan penanganan dagu si bungsu yang sobek dan harus dijahit masuk ke igd komunikasi perawat dokter dan petugas pendaftaran dengan kami boleh saya acungi jempol triase anamnese asesmen awal dan rekomendasi penanganan	Positif

	dikomunikasikan dengan baik ini penilaian obyektif mantan humas rumah sakit ya.	
--	---	--

E. Pengujian dan Analisis

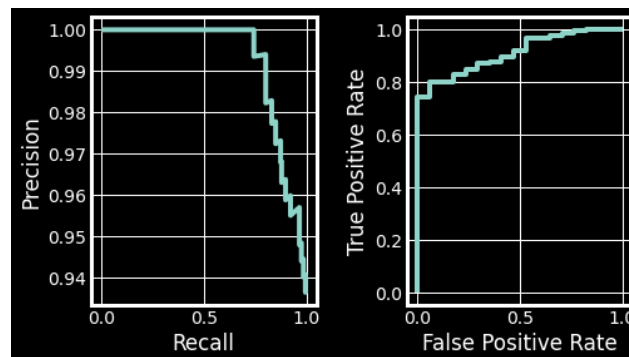
E.1 Klasifikasi dengan *Support Vector Machine* (SVM)

Bagian ini berisikan tentang hasil pengujian dari hasil klasifikasi menggunakan Support Vector Machine. Jenis pengujian yang dilakukan yaitu pengujian menggunakan data ketika sebelum dan sesudah di Model dengan TF-IDF. Berikut hasil pada pengujian pertama.

	precision	recall	f1-score	support
0	0.00	0.00	0.00	17
1	0.92	1.00	0.96	207
accuracy			0.92	224
macro avg	0.46	0.50	0.48	224
weighted avg	0.85	0.92	0.89	224
Train accuracy score: 0.9626307922272048				
Test accuracy score: 0.9241071428571429				
Train ROC-AUC score: 1.0				
Test ROC-AUC score: 0.9113384484228475				

Gambar 7. Hasil Evaluasi Pengujian Sebelum Menggunakan Model TF-IDF dan *Grid Search*

Dapat dilihat untuk akurasi nilai f-1 score yaitu 0.92, serta untuk nilai akurasi tes yaitu 0.92 dan nilai akurasi *train* yaitu 0.96. Dapat disimpulkan model yang dilatih dan dites memiliki hasil yang cukup baik dan data nya *fit*.



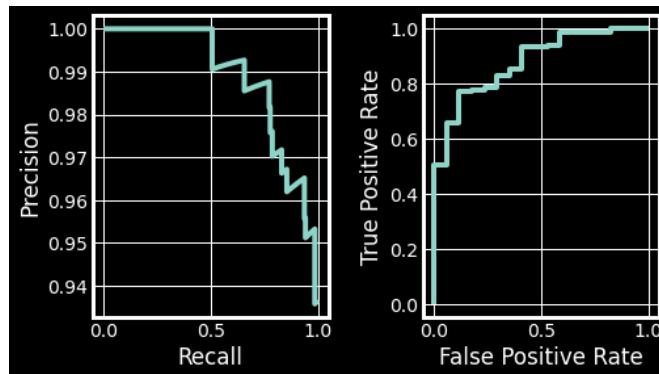
Gambar 8. Grafik *Recall* dan FPR Sebelum Menggunakan Model TF-IDF dan *Grid Search*

Selain itu untuk pengujian kedua menggunakan metode SVM dengan TF-IDF serta menggunakan *grid search* untuk mencari *best* parameter, berikut hasil evaluasinya.

	precision	recall	f1-score	support
0	0.00	0.00	0.00	17
1	0.92	1.00	0.96	207
accuracy			0.92	224
macro avg	0.46	0.50	0.48	224
weighted avg	0.85	0.92	0.89	224
Train accuracy score: 0.9551569506726457				
Test accuracy score: 0.9241071428571429				
Train ROC-AUC score: 1.0				
Test ROC-AUC score: 0.8792270531400966				
Are under Precision-Recall curve: 0.9605568445475638				
Area under ROC-AUC: 0.9886244232040304				

Gambar 9. Hasil Evaluasi Pengujian Sesudah Menggunakan Metode TF-IDF dan *Grid Search*

Dapat dilihat untuk akurasi nilai f-1 score yaitu 0.92, serta untuk nilai akurasi tes yaitu 0.92 dan nilai akurasi *train* yaitu 0.95. Dapat disimpulkan model yang dilatih dan dites memiliki hasil yang cukup baik dan data nya fit.



Gambar 10. Grafik Recall dan FPR Sesudah Menggunakan Model TF-IDF dan *Grid Search*

Sehingga perbedaan SVM dengan SVM Optimized tidak terlalu signifikan jika dilihat dari perbedaan angka f1-score, nilai akurasi tes, dan nilai akurasi *train*.

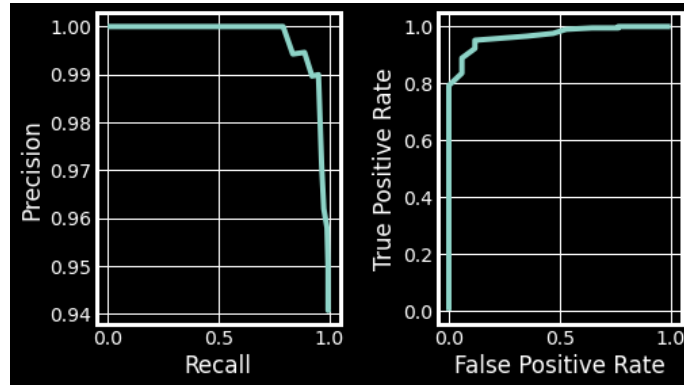
E.2 Klasifikasi dengan *Random Forest*

Perhitungan evaluasi metrics klasifikasi yang kita gunakan adalah *confusion metrix*, F1-score and akurasi. Berikut lampiran gambarnya.

	precision	recall	f1-score	support
0	0.00	0.00	0.00	17
1	0.92	1.00	0.96	207
accuracy			0.92	224
macro avg	0.46	0.50	0.48	224
weighted avg	0.85	0.92	0.89	224
Train accuracy score: 1.0				
Test accuracy score: 0.9241071428571429				
Train ROC-AUC score: 1.0				
Test ROC-AUC score: 0.9680306905370843				
Are under Precision-Recall curve: 0.9605568445475638				
Area under ROC-AUC: 0.9973167302785091				

Gambar 11. Hasil Evaluasi Pengujian Metode *Random Forest*

Dapat dilihat untuk akurasi nilai f-1 score yaitu 0.92, serta untuk nilai akurasi tes yaitu 0.92 dan nilai akurasi train yaitu 1. Dapat disimpulkan model yang dilatih dan dites memiliki hasil yang cukup baik dan data nya fit.

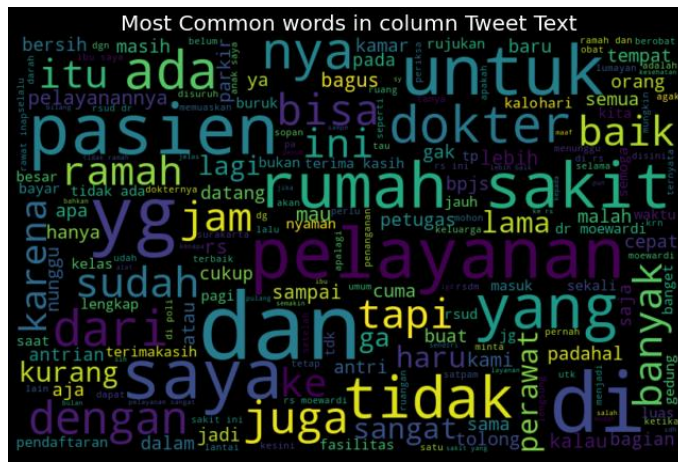


Gambar 12. Grafik Recall dan FPR Dengan Metode *Random Forest*

F. Visualisasi

Visualisasi data bertujuan untuk menggambarkan sifat dari hasil analisis sehingga mendapatkan kesimpulan yang berguna bagi masyarakat atau pekerja di RSUD Moewardi untuk dapat meningkatkan pelayanannya.

a. Visualisasi data dengan Word Cloud



Gambar 13. Word Cloud Sentiment Keseluruhan Ulasan Google Maps

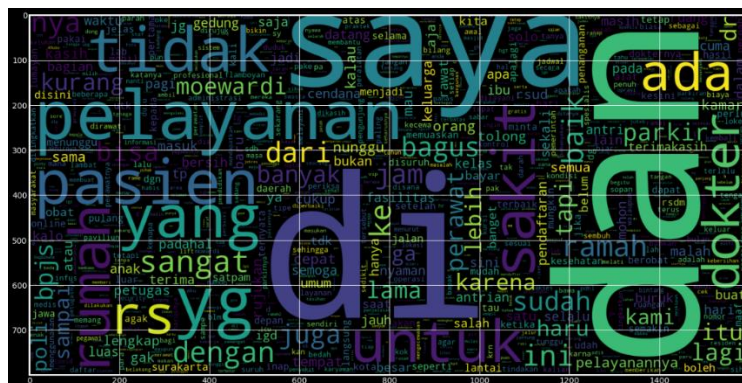
b. Visualisasi Data Sentimen Positif Dengan Word Cloud



Gambar 14. Word Cloud Sentiment Positif Ulasan Google Maps

Dalam visualisasi word cloud sentimen positif didapatkan dua kata yang frekuensinya muncul paling banyak selain keyword yaitu “pelayanan” Kata ini merupakan salah satu kata kunci pemicu sentimen positif. Hal ini bisa dikatakan bahwa ulasan dari *Google Maps* memudahkan masyarakat dalam mendapatkan informasi terkait ulasan tempat tersebut tentang pelayanan dan sarana prasarana RSUD Moewardi.

c. Visualisasi Data Sentimen Negatif Dengan Word Cloud



Gambar 15. Word Cloud Sentiment Negatif Ulasan *Google Maps*

Dalam visualisasi word cloud sentimen negatif didapatkan dua kata yang frekuensinya muncul paling banyak selain keyword yaitu, “pelayanan” dan “pasien”. Hal ini bisa dikatakan bahwa terdapat masalah pada pelayanan yang kurang maksimal menyebabkan pasien merasakan ketidaknyamanan .

Kesimpulan

Berdasarkan hasil analisis dan pembahasan mengenai analisis sentimen ulasan *Google Maps*, penulis dapat menarik kesimpulan bahwa klasifikasi menggunakan metode *Support Vector Machine* (SVM) memiliki hasil *accuracy* yang tidak jauh berbeda dengan hasil *accuracy* dari metode *Random Forest*. Sehingga model klasifikasi sentimen analisis ulasan *Google Maps* menggunakan metode *Support Vector Machine* (SVM) Optimized jika dipertimbangkan dari semua score dan grafik. Metode ini dapat digunakan pihak masyarakat dan pekerja di RSUD Moewardi untuk mengetahui pendapat orang-orang mengenai RSUD Moewardi di *Google Maps*. Selain itu, melalui visualisasi dan eksplorasi data yang telah dilakukan, didapatkan bahwa kata yang sering muncul dalam sentimen positif yaitu “pelayanan”. Sedangkan untuk sentimen negatif yaitu “pelayanan” dan “pasien”. Kata-kata ini juga bisa digunakan untuk menjadi peningkatan kualitas atau evaluasi bagi pekerja RSUD Moewardi. Dari hasil analisis dan kesimpulan, peneliti memberikan beberapa saran. Bagi pihak pemerintah dan pekerja RSUD Moewardi, diharapkan mampu memperbaiki dan mengembangkan pelayanan kepada masyarakat dan terus memberikan hasil yang maksimal untuk merawat pasien.

Bagi masyarakat diharapkan untuk meningkatkan kesadaran diri untuk selalu menggunakan Sarana dan prasarana RSUD Moewardi dengan bijaksana. Selain itu, masyarakat harus selalu menjaga kebersihan dan ketentraman rumah sakit.. Serta untuk penelitian selanjutnya perlu dilakukan pengembangan penelitian menggunakan metode lain. Tidak hanya itu, data yang digunakan untuk penelitian perlu ditambah agar mendapatkan hasil yang lebih akurat serta tidak melakukan labeling ulasan secara manual.

Daftar Pustaka

- Anjasmoros, M. T., Istiadi, & Marisa, F. (2020). Analisis Sentimen Aplikasi Go-Jek Menggunakan Metode Svm Dan Nbc (Studi Kasus: Komentar Pada Play Store). Conference on Innovation and Application of Science and Technology (CIASTECH) 2020.
- Aruan, J. D. C., Rahayudi, B., & Ridok, A. (2022). Analisis Sentimen Opini Masyarakat terhadap Pelayanan Rumah Sakit Umum Daerah menggunakan Metode Support Vector Machine dan Term Frequency -Inverse Document Frequency. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, 6(5), 2072–2078.
- Fauzi, M. A. (2018). Random forest approach for sentiment analysis in Indonesian language. *Indonesian Journal of Electrical Engineering and Computer Science*, 12(1), 46–50. doi: 10.11591/ijeecs.v12.i1.
- Fitri, V. A., Andreswari, R. dan Hasibuan, M. A. (2019). Sentiment analysis of social media Twitter with case of Anti-LGBT campaign in Indonesia using Naïve Bayes, decision tree, and random forest algorithm. *Procedia Computer Science*, 161, 765–772. doi: 10.1016/j.procs.2019.11.181.
- Johnson, F., & Gupta, S. (2012). Web Content Mining Techniques: A Survey. *International Journal of Computer Applications*, 47(11), 44–50. <https://doi.org/10.5120/7236-0266>
- Maulana, A. dan Pratiwi, H. (2019). Sentiment analysis of public towards infrastructure development in Indonesia on Twitter media using boosting support vector machine method. *AIP Conference Proceedings*, 2202(2019). doi: 10.1063/1.5141695.
- Mining, A. O. (2020). Ringkasan Jumlah Aspek Ulasan Hotel untuk Pembentukan Dataset Sentimen Analisis Berbasis Aspek., 3(2), pp. 62–66.
- Karmayasa O., Mahendra I. B. 2010. Implementasi Vector Space Model dan Beberapa Notasi Metode Term Frequency Inverse Document Frequency (TF-IDF) pada Sistem Temu Kembali Informasi.

Nayak, A. (2016). Comparative study of Naïve Bayes , Support Vector Machine and Random Forest Classifiers in Sentiment Analysis of Twitter feeds. International Journal of Advanced Studies in Computer Science and Engineering, 5(1), pp. 14–17.