

Application Random Forest Method for Sentiment Analysis in Jamsostek Mobile Review

Penerapan Metode Random Forest untuk Analisis Sentimen Pada Review Aplikasi Jamsostek

Tasya Auliya Ulul Azmi¹, Lutfi Hakim², Dian Candra Rini Novitasari³, Wika Dianita Utami

^{1,2,3,4}Departement of Mathematics, UIN Sunan Ampel Surabaya, Indonesia

tasyaauliya16@gmail.com

*: *Penulis korespondensi (corresponding author)*

Informasi Artikel

Received: January 2023

Revised: January 2023

Accepted: January 2023

Published: February 2023

Abstract

Purpose: This study aims to monitor the service quality of JMO applications from time to time by classifying JMO user reviews into the class of positive, neutral, and negative sentiments.

Design/methodology/approach : The method used in this study is the random forest classification method. Data processing in this study uses feature extraction, TF-IDF and labeling with the lexicon-based method.

Findings/result: Based on the research results, it was found that the highest frequency of classification was the positive class with 17571 reviews compared to the neutral class with 8701 reviews and the negative class with 3876 reviews with an accuracy evaluation value of 93%, precision 88%, recall 93%, and f1-score 90%.

Originality/value/state of the art:

This study uses 150737 reviews that have been pre-processed using the random forest method and TF-IDF and lexicon-based feature extraction.

Abstrak

Tujuan: Penelitian ini bertujuan untuk memantau kualitas pelayanan aplikasi JMO dari waktu ke waktu dengan mengklasifikasikan review pengguna JMO dalam bentuk sentimen positif, netral, dan negatif.

Perancangan/metode/pendekatan: Metode yang digunakan pada penelitian ini adalah metode klasifikasi random forest. Pengolahan data pada penelitian ini menggunakan extraction feature, TF-IDF dan pelabelan dengan metode lexicon based.

Keywords: JMO; Sentimen Analysis; Machine Learning

Kata kunci: JMO; Analisis Sentimen; Machine Learning

Hasil: Berdasarkan hasil penelitian, diperoleh frekuensi terbesar klasifikasi adalah kelas positif sebanyak 17571 *review* dibandingkan dengan kelas netral sebanyak 8701 *review* dan kelas negatif sebanyak 3876 *review* dengan nilai evaluasi *accuracy* 93%, *precision* 88%, *recall* 93%, dan *f1-score* 90%.

Keaslian/ *state of the art*: Penelitian ini menggunakan 150737 *review* yang telah di *pre-processing* dengan menggunakan metode klasifikasi *random forest* dan *extraction feature*, TF-IDF dan metode *lexicon based*.

1. Pendahuluan

Pesatnya perkembangan teknologi saat ini sangatlah memberikan kemudahan bagi masyarakat, organisasi maupun perusahaan dalam menyelesaikan berbagai masalah. Teknologi yang terus menerus tumbuh dapat menghasilkan data dengan jumlah yang besar dimana data tersebut bisa memberikan informasi yang bermanfaat [1]. Adanya teknologi dan informasi muncul inovasi dalam mengumpulkan data dengan memanfaatkan teknologi, salah satunya berkembangnya metode pengumpulan data secara digital yang bisa dilakukan atau diakses melalui seluler, tablet maupun komputer sehingga memudahkan pengguna dalam mengumpulkan berkas yaitu Jamsostek Mobile (JMO) [2].

Aplikasi Jamsostek Mobile atau yang sering dikenal dengan JMO merupakan aplikasi layanan digital dari BPJS Ketenagakerjaan yang memudahkan pengguna untuk cek saldo jaminan hari tua (JHT), simulasi jumlah saldo JHT, cek rincian iuran JHT, pelacakan dan pengajuan klaim JHT, pengkinian data diri, pengaduan, dan masih banyak fitur-fitur yang bisa diakses di aplikasi JMO [3]. Seluruh fitur tersebut bertujuan agar mempermudah akses layanan masyarakat terhadap layanan yang ada di BPJS Ketenagakerjaan sehingga masyarakat tidak perlu repot untuk datang ke kantor BPJS Ketenagakerjaan agar pelayanan yang diberikan lebih efisien, efektif, menghemat biaya, dan waktu [4]. Aplikasi JMO bisa didownload di android maupun ios melalui *google play store*. *Google play store* adalah platform yang menyediakan aplikasi, musik, buku, dan film dengan berbagai kategori. *Google play store* juga memiliki banyak fitur salah satunya yaitu fitur rating dan *review* dimana pengguna dapat memberikan *review* berupa komentar dan rating mengenai kualitas pelayanan aplikasi tersebut yang bertujuan untuk mengetahui kekurangan dan kelebihan dari aplikasi [5]. Rating aplikasi JMO pada *review* telah menjadi 4,7 dengan ulasan sekitar 1 juta *review* yang rata-rata komentar bersifat positif dari pengguna. Meskipun aplikasi JMO ini memiliki penilaian yang baik tetapi adapula komentar yang bersifat negatif dalam *review* yang diberikan pengguna. Pada uraian tersebut diperlukan analisis sentimen.

Analisis sentimen merupakan *text mining* berupa opini, pendapat, atau emosi pengguna dalam bentuk komentar, *review*, umpan balik, dan memberikan informasi yang diperlukan untuk berbagai tujuan [6]. Analisis sentimen adalah proses memahami, mengekstraksi, dan mengolah data tekstual secara otomatis untuk mendapatkan informasi [7]. Analisis sentimen berguna untuk memantau kualitas aplikasi JMO dari waktu ke waktu melalui *review* sehingga dapat meningkatkan kualitas pelayanan BPJS Ketenagakerjaan agar dapat memudahkan pekerjaan pegawai dan juga pengguna [8]. Terdapat dua pendekatan yang digunakan untuk menganalisis

yakni *machine learning* dan *knowledge-based*. *Machine learning* adalah sentimen yang membutuhkan dataset untuk mengumpulkan, mengklasifikasi sampel dalam sampel dataset, dan proses train yang membutuhkan waktu. Sedangkan, *knowledge-based* adalah pendekatan analisis sentimen pada level kata [9]. Analisis sentimen juga merupakan klasifikasi teks untuk menentukan suatu teks tersebut memiliki makna positif, netral atau negatif [10].

Beberapa kajian penelitian sebelumnya yang dianggap relevan untuk penelitian ini yaitu penelitian dari Siti dan Amsal yang berjudul “*Analysis Of User Satisfaction Level Of Jamsostek Mobile Application In Aceh Using End User Computing Satisfaction*”. Pada penelitian ini terdapat 107 responden dengan hasil yaitu bahwa nilai konten dan format berpengaruh signifikan terhadap kepuasan pelanggan. Nilai signifikan konten $0.005 < 0.05$ dan format $0.035 < 0.05$. Sedangkan, variabel akurasi, kemudahan pengguna, dan ketepatan waktu tidak berpengaruh signifikan terhadap kepuasan pengguna sebab angka signifikan lebih besar dari 0.05 dan $t_{hitung} < t_{tabel}$ atau nilai t uji hipotesis dari perhitungan statistik lebih kecil dari nilai yang terdapat pada Tabel T menggukan uji T [11].

Penelitian oleh Boma, Irwan dan Siti yang berjudul “*Analisis Sentimen Pelanggan Hotel di Purwokerto Menggunakan Metode Random Forest dan TF-IDF*”. Pada penelitian ini dataset yang digunakan untuk analisis sentimen ini adalah data komentar pelanggan hotel di Purwokerto yang diambil dari lama [tripadvisor.co.id](https://www.tripadvisor.co.id) dengan mengklasifikasikan *review* pelanggan dalam dua kelas yakni *review* positif dan negatif. Hasil dari penelitian tersebut menunjukkan akurasi model dari *random forest* mencapai angka 87,23% [12].

Berdasarkan penelitian yang telah dijelaskan, maka penelitian ini bertujuan untuk memantau kualitas aplikasi JMO dari waktu ke waktu sehingga dapat meningkatkan kualitas pelayanan BPJS Ketenagakerjaan maka dibutuhkan analisis sentimen aplikasi Jamsostek Mobile (JMO) dengan menggunakan metode *random forest*. Data yang digunakan berasal dari *review* aplikasi JMO pada *google play store*. Ditransformasikan data *review* pengguna dari bentuk tekstual ke dalam bentuk kuantitas menggunakan *feature extraction* yaitu *lexicon based* dan *Term Frequency-Invers Document Frequency* (TF-IDF).

2. Metode/Perancangan

Data pada penelitian ini diambil dari *review* aplikasi JMO dari *google play store* dengan menggunakan *software python*. Metode yang digunakan adalah *random forest* untuk mengetahui hasil akurasi dalam analisis sentimen pada *review* pengguna aplikasi JMO. Berikut merupakan *flowchart* dari penelitian ini.

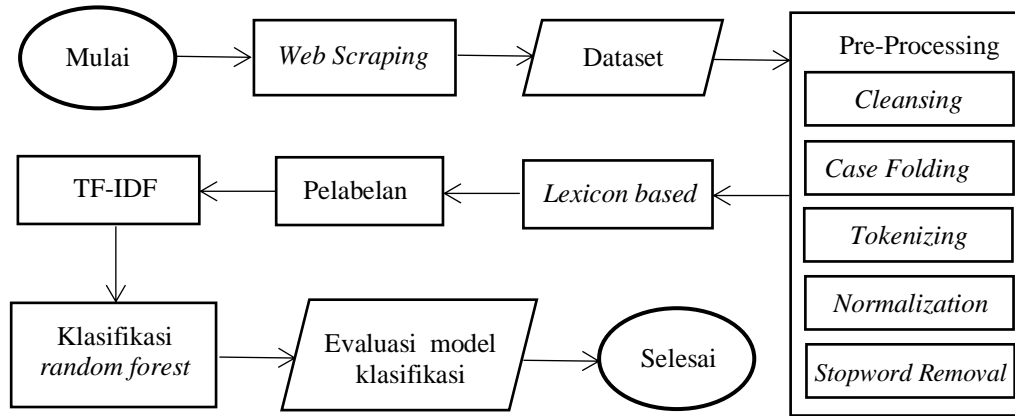
2.1. Web Scraping

Pengumpulan data *review* dari aplikasi JMO menggunakan proses *web scraping* dari *website google play store* yang terdiri dari *username*, *rating*, *time* dan *review* menggunakan *library google-play-scraper* yang kemudian disimpan dalam bentuk *csv* [13].

2.2. Pre-Processing

Text pre-processing termasuk kedalam *text mining*. *Text mining* merupakan analisis data bahasa alami yang menggunakan teknik dan alat untuk menemukan, merancang, dan mengekstrak informasi yang tidak terstruktur. *Text mining* berfungsi untuk mengubah kata dalam data yang tidak berstruktur menjadi data yang informasi yang lebih terstruktur dan mudah dianalisa [14].

Perbedaan antara *text mining* dan *data mining* adalah terletak pada *preprocessing*, *preprocessing*.



Gambar 1. Flowchart

Pada *data mining* memiliki fokus pada penomoran dan normalisasi data, sedangkan *preprocessing* pada *text mining* memiliki fokus pada ekstraksi fitur dan identifikasi [15]. *Text pre-processing* merupakan proses pembersihan dan tranformasi data sebelum data tersebut diolah. Terdapat 5 tahapan dalam *text pre-processing* yaitu *cleansing*, *case folding*, *tokenizing*, *normalization* dan *stopword removal* [16]. *Cleansing* merupakan proses penghapusan simbol, link, tab, hastag dan nomor. *Case folding* merupakan tahap mengubah teks menjadi huruf kecil [17]. *Tokenizing* merupakan langkah pemecahan kalimat menjadi token atau potongan kata. *Normalization* merupakan proses normalisasi kata-kata singkatan dan kata-kata gaul. *Stopword removal* merupakan tahap menghilangkan kata yang tidak penting [18].

2.3. Lexicon

Lexicon adalah pemilihan kata pada dokumen dalam analisis sentimen. *Review* yang mengandung kata yang terdapat pada kamus akan dilabeli sebagai *review* opini [19]. Kamus kata terdiri dari dua kamus yakni kamus kumpulan kata sentimen positif dan sentimen negatif. *Vader sentiment* adalah metode analisis yang digunakan oleh *lexicon based*. Hasil dari *vader* berupa 3 kelas yaitu positif, netral, dan negatif [20].

2.4. TF-IDF

TF-IDF merupakan metode pembobotan kata melalui ekstraksi ciri dari teks. TF atau *term frequency* merupakan frekuensi dari *term* yang muncul dalam sebuah dokumen, sedangkan IDF atau *inversed document frequency* merupakan ukuran besarnya kepentingan suatu term yang muncul dalam dataset [16]. Rumus untuk TF-IDF dapat dilihat pada Persamaan 3, 4 dan 5.

$$tf = tf_{ij} \quad (1)$$

$$idf = \log \frac{N}{df_i} \quad (2)$$

$$tf\ idf_{ij} = tf \times idf \quad (3)$$

Keterangan :

tf_{ij} = jumlah munculnya *term*.

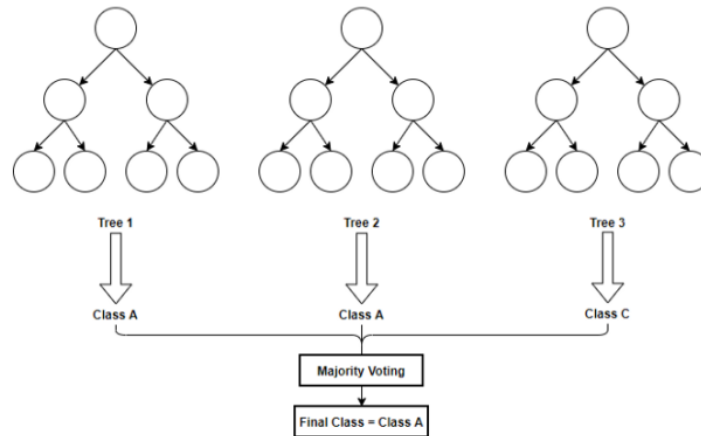
N = jumlah dokumen.

df_i = jumlah frekuensi kata yang terpilih.

$tf\ idf_{ij}$ = bobot dokumen.

2.5. Random forest

Random forest merupakan metode yang termasuk dalam klasifikasi dan juga regresi yang termasuk dalam kelompok *ensemble learning* [21]. Metode *random forest* merupakan pengembangan dari metode *decision tree* yang mana proses *train* data yang dihasilkan memiliki banyak pohon dan setiap pohon memiliki cara yang sama dengan pohon berkembang seiring bertambahnya dataset [22]. Membuat pohon keputusan digunakan metode *Classification and Regression Tree* (CART) agar didapatkan keputusan dengan ukuran yang maksimum dan tidak dipangkas sehingga dapat menghasilkan kumpulan pohon (*tree*) yang disebut *forest* [9].



Gambar 2. Alur *Random Forest Classifier*

Random forest terdiri dari 3 *node* yaitu *root*, *internal*, dan *leaf node*. *Root node* merupakan akar dari pembentukan pohon, *internal node* merupakan percabangan pohon. Sedangkan, *leaf node* merupakan *node* terakhir yang dibentuk. *Entropy* atau *gini impurity* merupakan perhitungan untuk menghitung ketidakmurnian atribut [23]. Berikut formulasi dari metode *random forest*.

$$Gini = 1 - \sum_{i=1}^C (p_i)^2 \quad (4)$$

$$gain(A) = Gini\ impurity - Gini\ Average \quad (5)$$

Keterangan :

C = jumlah kelas klasifikasi (jumlah atribut).

p_i = jumlah sampel kelas i.

2.6. Evaluasi

Confusion matrix merupakan cara untuk menyajikan hasil akurasi dari model dan sumber informasi apakah model yang digunakan bekerja dengan baik atau sebaliknya [15]. *Confusion matrix* digunakan untuk mengukur performa atau kinerja model dengan menghitung nilai akurasi terhadap *accuracy*, *precision*, *recall*, dan *F1-score* [6]. *F1-score* merupakan perbandingan dari nilai *recall* dan juga nilai *precision* dalam penelitian. *Accuracy* merupakan tingkat keakuratan model dalam klasifikasi atau tingkat kedekatan prediksi dengan nilai sebenarnya. Berikut formulasi *accuracy*.

$$accuracy = \sum_{i=1}^l \frac{TP_i + TN_i}{TP_i + TN_i + FP_i + FN_i} \times 100\% \quad (6)$$

Precision merupakan tingkat ketepatan dalam prediksi *true positive* terhadap keseluruhan data hasil prediksi positif. Rumus perhitungan *precision*.

$$presisi = \frac{\sum_{i=1}^l TP_i}{\sum_{i=1}^l (FP_i + TP_i)} \times 100\% \quad (7)$$

Recall merupakan perbandingan ketepatan ratio prediksi *true positive* dengan keseluruhan data *true positive*. Berikut formulasi *recall*.

$$recall = \frac{\sum_{i=1}^l TP_i}{\sum_{i=1}^l (TP_i + FN_i)} \times 100\% \quad (8)$$

Keterangan :

i = jumlah kelas

TP_i = true positive (jumlah data positif yang terklasifikasi benar untuk kelas ke-i).

TN_i = true negative (jumlah data negatif yang terklasifikasi benar untuk kelas ke-i).

FP_i = false positive (jumlah data positif yang terklasifikasi salah untuk kelas ke-i).

FN_i = false positive (jumlah data negatif yang terklasifikasi salah untuk kelas ke-i).

3. Hasil dan Pembahasan

Data yang diambil untuk penelitian analisis sentimen aplikasi JMO ini menggunakan teknik *web scraping* dari *website goole play store* didapatkan sebanyak 279.877 *review* yang paling relevan mulai dari 20 Februari 2018 sampai 17 November 2022. Proses *web scraping* dilakukan dengan menginstal *library google-play-scraper*. Contoh hasil *web scraping* sebagai berikut.

Tabel 1. Hasil *web scraping*

	<i>Uername</i>	<i>Rating</i>	<i>Time</i>	<i>Review</i>	
1.	Rina Irfanudin	5	11/17/2022 1:43	cukup bagus apk nya bpjs	
2.	akmal khan 1992	5	11/17/2022 1:43	kasih bintang 5 dulu kalau jelek di kurangi ya	
3.	Ade Munawar	5	11/17/2022 1:42	Sangat cepat dan mudah	
4.	<u>Dhimas</u> Ciptadi	1	11/17/2022 1:42	Bintang setengah ada ga?	
5.	Dhiemas Prasetyo	Aji	1	11/17/2022 1:4	Komplain lewat Email gak dibales. Call center 175 sok ngartis.
⋮	⋮	⋮	⋮	⋮	
279875.	Pengguna Google	3	2/20/2018 13:32	Coba dlu nnti bintangny d tmbh klo bgs	
279876.	Pengguna Google	5	2/20/2018 11:59	Aplikasinya bagus, bpjs ketenagakerjaan memang terbaik	

	<i>Uername</i>	<i>Rating</i>	<i>Time</i>	<i>Review</i>
279877.	Pengguna Google	5	2/20/2018 9:56	Cucokk mantep

Data hasil web scraping pada Tabel 1, kemudian disimpan dalam format CSV. Dataset tersebut berisi *username*, *rating*, *time*, dan *review*. Pada tahap selanjutnya dilakukan pembersihan data pada dataset atau disebut dengan *pre-processing*. Pada tahap *pre-processing* ini terdapat 5 tahap. Pada tahap *pre-processing* yang pertama yakni *cleansing* atau proses penghapusan *noise* seperti simbol, link, tab, hastag, dan nomor. Pada Tabel 2 merupakan contoh hasil dari *cleansing*.

Tabel 2. Hasil *cleansing*

	Sebelum	Sesudah
1.	cukup bagus apk nya bpjs	cukup bagus apk nya bpjs
2.	kasih bintang 5 dulu kalau jelek di kurangi ya	kasih bintang dulu kalau jelek di kurangi ya
3.	Sangat cepat dan mudah	Sangat cepat dan mudah
4.	Bintang setengah ada ga?	Bintang setengah ada ga?
5.	Komplain lewat Email gak dibales. Call center 175 sok ngartis.	Komplain lewat Email gak dibales. Call center sok ngartis.

Tahap kedua *pre-processing* yaitu *case folding* atau mengubah data teks menjadi huruf kecil (*lower*). Berikut contoh dari hasil *case folding*.

Tabel 3. Hasil *case folding*

	Sebelum	Sesudah
1.	cukup bagus apk nya bpjs	cukup bagus apk nya bpjs
2.	kasih bintang dulu kalau jelek di kurangi ya	kasih bintang dulu kalau jelek di kurangi ya
3.	Sangat cepat dan mudah	sangat cepat dan mudah
4.	Bintang setengah ada ga?	bintang setengah ada ga?
5.	Komplain lewat Email gak dibales. Call center sok ngartis.	komplain lewat email gak dibales. call center sok ngartis

Tahap ketiga *pre-processing* yaitu *Tokenizing* atau pemecahan kalimat menjadi potongan kata (*token*) dan menghilangkan tanda baca. Contoh hasil *tokenizing* sebagai berikut.

Tabel 4. Hasil *tokenizing*

	Sebelum	Sesudah
1.	cukup bagus apk nya bpjs	[cukup, bagus, apk, nya, bpjs]
2.	kasih bintang dulu kalau jelek di kurangi ya	[kasih, bintang, dulu, kalau, jelek, di, kurangi, ya]
3.	sangat cepat dan mudah	[sangat, cepat, dan, mudah]
4.	bintang setengah ada ga?	[bintang, setengah, ada, ga]
5.	komplain lewat email gak dibales. call center sok ngartis	[komplain, lewat, email, gak, dibales, call, center, sok, ngartis]

Tahap keempat *pre-processing* yaitu *normalization* atau normalisasi kata-kata singkatan dan kata-kata gaul. Pada Tabel 5 merupakan contoh hasil dari *normalization*.

Tabel 5. Hasil *normalization*

	Sebelum	Sesudah
1.	[cukup, bagus, apk, nya, bpjs]	[cukup, bagus, aplikasi, nya, bpjs]
2.	[kasih, bintang, dulu, kalau, jelek, di, kurangi, ya]	[kasih, bintang, dulu, kalau, jelek, di, kurangi, ya]
3.	[sangat, cepat, dan, mudah]	[sangat, cepat, dan, mudah]
4.	[bintang, setengah, ada, ga]	[bintang, setengah, ada, ga]
5.	[komplain, lewat, email, gak, dibales, call, center, sok, ngartis]	[komplain, lewat, email, gak, dibales, call, center, sok, ngartis]

Tahap terakhir *pre-processing* yaitu *stopword removal* merupakan tahap menghilangkan kata yang tidak penting. Berikut contoh dari hasil *stopword removal*.

Tabel 6. Hasil *stopword removal*

	Sebelum	Sesudah
1.	[cukup, bagus, apk, nya, bpjs]	bagus aplikasi bpjs
2.	[kasih, bintang, dulu, kalau, jelek, di, kurangi, ya]	kasih bintang jelek kurangi
3.	[sangat, cepat, dan, mudah]	cepat mudah
4.	[bintang, setengah, ada, ga]	bintang ga
5.	[komplain, lewat, email, gak, dibales, call, center, sok, ngartis]	komplain email gak dibales call center sok ngartis

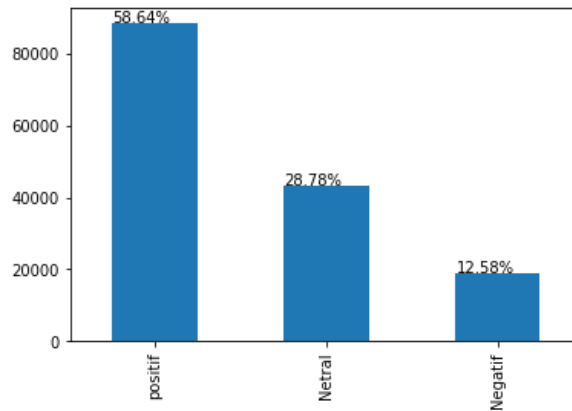
Hasil dari proses *pre-processing* kemudian dilakukan pelabelan dengan metode *lexicon based* menggunakan *library vader sentiment*. Contoh hasil pelabelan ditunjukkan pada Tabel 7.

Tabel 7. Hasil pelabelan

	Review	Score	Sentiment
1.	bagus aplikasi bpjs	0.4215	Positif
2.	kasih bintang jelek kurangi	0.1179	Positif
3.	cepat mudah	0.4404	Positif
4.	bintang ga	-0.2960	Negatif
5.	komplain email gak dibales call center sok ngartis	-0.3182	Negatif

Perhitungan skor pada metode *lexicon based* dengan cara menjumlahkan nilai bobot per kata yang mengandung sentiment dimana dibagi menjadi 3 kelas yakni positif, netral, dan negatif [18].

- 1) Jika $score > 0$ maka *sentiment* bernilai positif.
- 2) Jika $score = 0$ maka *sentiment* bernilai netral.
- 3) Jika $score < 0$ maka *sentiment* bernilai negatif.



Gambar 3. Diagram jumlah pelabelan

Jumlah pelabelan sesuai kelas pada Gambar 3 menunjukkan bahwa jumlah data *review* yang sudah melewati proses *pre-processing* berjumlah 150737 *review* dengan frekuensi sentimen yang paling tinggi adalah positif dengan jumlah 88395 *review*. Setelah itu, TF-IDF atau pembobotan kata dari frekuensi *term* yang sering muncul yang dapat ditunjukkan dengan hasil *wordcloud* ketiga kelas pada Gambar 4, 5, dan 6.



Gambar 4. Wordcloud positif



Gambar 5. Wordcloud netral



Gambar 6. Wordcloud negative

Pada tahap pemodelan klasifikasi menggunakan 80% data *training* dan 20% data *testing* dengan metode *random forest*. Hasil klasifikasi metode *random forest* yang akan digunakan untuk mengukur performa atau kinerja model dari model yang diuji dengan menggunakan hasil *confusion matrix* yang merupakan hasil data *review* yang dirubah menjadi data numerik.

Tabel 8. Confusion matrix random forest

	Negatif	Netral	Positif
Negatif	3527	213	136
Netral	273	8369	59
Positif	1179	245	16147

Hasil *confusion matrix* pada Tabel 8 diperoleh nilai *true positive*, *true negative*, *false positive*, dan *false negative* secara berurutan sebesar 16147, 12382, 195, dan 1424. Hasil evaluasi dari model klasifikasi *random forest* menghasilkan *confusion report* yang berasal dari perhitungan semua matriks yang diperlukan untuk evaluasi.

Tabel 9. Confusion report random forest

Precision	Recall	F1-score	Support

Negatif	0.71	0.91	0.80	3876
Netral	0.95	0.96	0.95	8701
Positif	0.99	0.92	0.95	17571
<i>Accuracy</i>	0.93	0.93	0.93	30148
<i>Macro avg</i>	0.88	0.93	0.90	30148
<i>Wighted avg</i>	0.94	0.93	0.93	30148

Hasil dari evaluasi metode *random forest* pada Tabel 9 diperoleh nilai *accuracy* sebesar 93% yang menggambarkan tingkat akurat model dalam klasifikasi. Nilai *precision* mencapai 88% yang artinya akurasi antara data aktual dan prediksi yang diberikan oleh model sebesar 88%. Keberhasilan model pada penelitian ini atau disebut sebagai *recall* bernilai 93%. Sedangkan *f1-score* atau rata-rata *precision* dan *recall* yang dibobotkan bernilai 90%.

4. Kesimpulan dan Saran

Berdasarkan hasil penelitian ini, terdapat 3 kelas yaitu positif, netral, dan negatif yang memiliki akurasi yang sangat tinggi sebesar 93% yang menunjukkan bahwa nilai akurat model klasifikasi analisis sentimen terhadap *review* pengguna aplikasi JMO menggunakan metode *random forest* sangat baik. Diketahui pula frekuensi dari *review* positif lebih tinggi dibandingkan dengan kelas lainnya. Hal ini menunjukkan bahwa aplikasi JMO memiliki pelayanan yang baik sehingga dapat menjadi sumber informasi bagi calon pengguna lainnya.

Daftar Pustaka

- [1] N. C. Agustina, D. Herlina Citra, W. Purnama, C. Nisa, and A. Rozi Kurnia, "Implementation of Naïve Bayes Algorithm for Sentiment Analysis of Shopee Reviews on Google Play Store," vol. 2, no. April, pp. 47–54, 2022.
- [2] B. S. Arisoemaryo *et al.*, "Analisis Tingkat Kepuasan Pengguna Aplikasi Jamsostek Mobile Menggunakan Metode End User Computing Satisfaction," vol. 4, no. 1, pp. 110–117, 2022, [Online]. Available: <https://ejurnal.ars.ac.id/index.php/jti/article/view/724>
- [3] I. Fahlevi, "Efektivitas Aplikasi Jamsostek Mobile (JMO) Dalam Proses Pencairan Jaminan Hari Tua (JHT) Kepada Peserta BPJS Ketenagakerjaan Cabang Meulaboh," vol. 6, 2022.
- [4] N. F. Fano, "Pembuatan Model Pemeringkatan Ulasan Menggunakan Metode Random Forest Regression," *JATISI (Jurnal Tek. Inform. dan Sist. Informasi)*, vol. 8, no. 2, pp. 664–675, 2021, doi: 10.35957/jatisi.v8i2.777.
- [5] M. Farkhan, D. Samudera, and V. I. Pertiwi, "Inovasi Pelayanan Publik Melalui Jamsostek Mobile (Jmo) (Studi Kasus Di Bpjs Ketenagakerjaan Cabang Rungkut Kota Surabaya)," *JPAP (Jurnal Penelit. Adm. Publik)*, vol. 8, no. Vol 8 No 1 (2022): Jurnal Penelitian Administrasi Publik (JPAP), pp. 152–172, 2022.
- [6] Y. Al Amrani, M. Lazaar, and K. E. El Kadirp, "Random forest and support vector machine based hybrid approach to sentiment analysis," *Procedia Comput. Sci.*, vol. 127, pp. 511–520, 2018, doi: 10.1016/j.procs.2018.01.150.

- [7] A. Ahmed Shokeralla, F. Al Rahman Idress Al Sameeh, A. Gibreel Mohammed Musa, F. ELrhman Elsmih, and A. A. Shokeralla, "Prediction the daily number of confirmed cases of covid-19 in Sudan with arima and holt winter exponential smoothing," *Artic. Int. J. Dev. Res.*, no. August, 2020, doi: 10.37118/ijdr.19811.08.2020.
- [8] J. Y. B. Yin, N. H. M. Saad, and Z. Yaacob, "Exploring Sentiment Analysis on E-Commerce Business: Lazada and Shopee," *TEM J.*, vol. 11, no. 4, pp. 1508–1519, 2022, doi: 10.18421/TEM114-11.
- [9] M. Y. Aldean, P. Paradise, and N. A. Setya Nugraha, "Analisis Sentimen Masyarakat Terhadap Vaksinasi Covid-19 di Twitter Menggunakan Metode Random Forest Classifier (Studi Kasus: Vaksin Sinovac)," *J. Informatics, Inf. Syst. Softw. Eng. Appl.*, vol. 4, no. 2, pp. 64–72, 2022, doi: 10.20895/inista.v4i2.575.
- [10] I. N. Kabiru, P. K. Sari, S. Prodi, and M. Bisnis, "Analisa Konten Media Sosial E-Commerce Pada Instagram Menggunakan Metode Sentimen Analysis Dan Lda-Based Topic Modeling (Studi Kasus : Shopee Indonesia) Analysis of Content Social Media E-Commerce in Instagram Using Sentiment Analysis and Lda Based Topi," *e-Proceeding Manag.*, vol. 6, no. 1, pp. 12–19, 2019.
- [11] Y. Kelpin and K. Ardila, "Analysis of User Satisfaction Level Of E-Pkh Application Using End User Computing Satisfaction (EUCS) Method," *Adpebi Sci. Ser.*, pp. 869–876, 2022, [Online]. Available: <http://series.adpebi.com/index.php/AICMEST/article/view/63>
- [12] T. S. Kasus, P. Pada, and S. Tripadvisor, "Analisis Sentimen Pelanggan Hotel di Purwokerto Menggunakan Metode Random Forest dan TF-IDF," vol. 8106, pp. 21–29, 2021, doi: 10.20895/INISTA.V3.
- [13] E. Fitri, "Analisis Sentimen Terhadap Aplikasi Ruangguru Menggunakan Algoritma Naive Bayes, Random Forest Dan Support Vector Machine," *J. Transform.*, vol. 18, no. 1, p. 71, 2020, doi: 10.26623/transformatika.v18i1.2317.
- [14] O. Somantri and D. Dairoh, "Analisis Sentimen Penilaian Tempat Tujuan Wisata Kota Tegal Berbasis Text Mining," *J. Edukasi dan Penelit. Inform.*, vol. 5, no. 2, p. 191, 2019, doi: 10.26418/jp.v5i2.32661.
- [15] M. Dimas Lutfiyanto, E. B. Setiawan, and S. Si, "Expansion Feature dengan Word2Vec untuk Analisis Sentimen pada Opini Politik di Twitter dengan Klasifikasi Support Vector Machine, Naive Bayes, dan Random Forest," vol. 8, no. 5, pp. 10399–10410, 2016.
- [16] R. D. Himawan and E. Eliyani, "Perbandingan Akurasi Analisis Sentimen Tweet terhadap Pemerintah Provinsi DKI Jakarta di Masa Pandemi," *J. Edukasi dan Penelit. Inform.*, vol. 7, no. 1, p. 58, 2021, doi: 10.26418/jp.v7i1.41728.
- [17] R. Batra and S. M. Daudpota, "Integrating StockTwits with sentiment analysis for better prediction of stock price movement," *2018 Int. Conf. Comput. Math. Eng. Technol. Innov. Integr. Socioecon. Dev. iCoMET 2018 - Proc.*, vol. 2018-Janua, pp. 1–5, 2018, doi: 10.1109/ICOMET.2018.8346382.
- [18] T. Rahman, F. E. M. Agustin, and N. F. Rozy, "Normalization of Unstructured Indonesian Tweet Text for Presidential Candidates Sentiment Analysis," *2019 7th Int.*

- Conf. Cyber IT Serv. Manag. CITSM 2019*, 2019, doi: 10.1109/CITSM47753.2019.8965324.
- [19] Y. Azhar, “Metode Lexicon-Learning Based Untuk Identifikasi Tweet Opini Berbahasa Indonesia,” *J. Nas. Pendidik. Tek. Inform.*, vol. 6, no. 3, p. 237, 2018, doi: 10.23887/janapati.v6i3.11739.
- [20] F. Amaliah and I. K. Dwi Nuryana, “Perbandingan Akurasi Metode Lexicon Based Dan Naive Bayes Classifier Pada Analisis Sentimen Pendapat Masyarakat Terhadap Aplikasi Investasi Pada Media Twitter,” *J. Informatics Comput. Sci.*, vol. 3, no. 03, pp. 384–393, 2022, doi: 10.26740/jinacs.v3n03.p384-393.
- [21] A. S. Neogi, K. A. Garg, R. K. Mishra, and Y. K. Dwivedi, “Sentiment analysis and classification of Indian farmers’ protest using twitter data,” *Int. J. Inf. Manag. Data Insights*, vol. 1, no. 2, p. 100019, 2021, doi: 10.1016/j.jjime.2021.100019.
- [22] D. Irawan, E. B. Perkasa, Y. Yurindra, D. Wahyuningsih, and E. Helmud, “Perbandingan Klasifikasi SMS Berbasis Support Vector Machine, Naive Bayes Classifier, Random Forest dan Bagging Classifier,” *J. Sisfokom (Sistem Inf. dan Komputer)*, vol. 10, no. 3, pp. 432–437, 2021, doi: 10.32736/sisfokom.v10i3.1302.
- [23] A. N. Izza, D. E. Ratnawati, W. Hayuhardhika, N. Putra, U. Brawijaya, and P. Korespondensi, “Analisis Sentimen Objek Wisata Di Provinsi Sulawesi Sentiment Analysis of Tourist Attractions in the Province of South Sulawesi Based on Visitor Reviews,” vol. 3, no. 2, pp. 97–105, 2022.