

Analysis of Public Sentiment of Covid-19 Dynamics on Social Media Using Support Vector Machine and Particle Swarm Optimization

¹Fakhri Muhammad, ²Chaerur Rozikin, ³Riza Ibnu Adam

^{1,2,3}Teknik Informatika, Fakultas Ilmu Komputer, Universitas Singaperbangsa Karawang
Email: fakhri.muhammad18054@student.unsika.ac.id

Abstract

Variants of covid in Indonesia continue to grow and make people required to stay at home and are not required to go out if they don't have important things, therefore many people who stay at home often play social media such as twitter, it is possible that many irresponsible people make opinions or hoaxes with a specific purpose to make tweets that are not in accordance with the facts, which are feared to make the public more panicked about the increase in this covid-19 variant. Therefore this study was conducted to classify tweets as positive, negative, and neutral. The methodology used is a text mining process with 4 modeling using Support Vector Machine and Particle Swarm Optimization. The results obtained from the 4th modeling produce an accuracy of 83% on the linear kernel. While the PSO modeling in scenario 4 with 90:10 data division resulted in the highest accuracy in linear and polynomial kernels of 86% and 87%, respectively. Other evaluation values also improved, such as precision to 90%, recall to 83% and f1-score to 86%.

Keyword: Text mining, Text Classification, Support Vector Machine, Particle Swarm Optimization

1. INTRODUCTION

Dinamika wabah covid-19 di indonesia terus meningkat 2 tahun kebelakang ini, menurut [1] kasus covid-19 sudah menginfeksi lebih dari 5.701.257 lebih jiwa dan membuat menurunnya ekonomi hampir diseluruh dunia [2]. Wabah ini sangat cepat menular umumnya bagi orang yang sedang tidak fit dan orang yang sudah tua “manula” rentang umur 50-60 tahun[3], Indonesia sempat mengalami 2 kali gelombang covid-19 dengan gelombang ke dua yaitu varian delta, bahkan ada media luar yang mengatakan bahwa indonesia sebagai epicentrum covid-19 di wilayah Asia pada gelombang ke dua ini[4]. Oleh sebab itu dengan masih meningkatnya kasus covid 19 membuat masyarakat dituntut untuk selalu menjaga prokes sesuai anjuran pemerintah dan melakukan *social distancing* dan pemberlakuan pembatasan kegiatan masyarakat di beberapa daerah dengan kasus covid-19 masih tinggi, dengan adanya program ini adalah suatu cara untuk mengurangi peningkatan kasus covid, dengan adanya ppkm ini membuat masyarakat mau tidak mau menggunakan media sosial yang sudah berkembang pesat untuk bersosialisasi berbasis virtual atau online. Peran dari media sosial sangat penting untuk mengetahui informasi terbaru tentang kasus covid-19 dan juga menjadi sarana pihak berwenang untuk mengedukasi masyarakat agar selalu menjaga imun dengan berolahraga, mematuhi prokes dengan tujuan yaitu agar wabah covid-19 cepat selesai dan keadaan bisa normal kembali[5].

Perkembangan teknologi setiap tahun terus berkembang[6] dengan memanfaatkan media sosial

twitter merupakan wadah yang banyak digunakan untuk mengekspresikan opini baik tentang wabah covid-19 dan lain hal, dengan adanya media sosial selain menyebarkan informasi positif juga bisa menjadi negatif jika tidak digunakan dengan bijak oleh penggunanya. Ada sebagian oknum yang memanfaatkan wabah covid 19 ini dengan tujuan tertentu kerap kali menyebarkan kepanikan, hoax, ujaran kebencian dan pengalihan opini publik[7]. Apabila kondisi seperti ini dibiarkan akan menjadi sangat fatal di era pandemic ini, jika suatu fakta yang telah di rilis oleh pemerintah dan dicoba untuk dibelokan kebenaran dalam media sosial twitter, hal ini akan menjadi bahaya karena dengan menyebar luasnya berita hoax oleh oknum tidak bertanggung jawab mereka bisa memanfaatkan keterbukaan akses yang ada di media sosial. Jika *sentiment* masyarakat terhadap berita yang berisi fakta menjadi berubah, nanti akan banyak orang-orang yang cenderung tidak mempercayai fakta tentang berita tersebut dan lebih memilih mengabaikan, yang mana hal ini sangat buruk dan bisa memperkeruh kondisi di era pandemi ini. Oleh sebab itu penting untuk mengetahui apakah benar perubahan *sentiment* masyarakat dalam menanggapi suatu berita di media sosial khususnya twitter dapat mempengaruhi persebaran wabah covid-19.

Sentiment analisis pada teks tweet dari media sosial twitter sudah banyak dilakukan di banyak bahasa dan metode yang digunakan baik berbasis *lexicon* dari kata positif, negatif sampai dengan menggunakan metode *machine learning*, *deep learning* baik secara *supervised* maupun *unsupervised* [8]. Analisis *sentiment* untuk bahasa indonesia data teks *tweet* banyak yang menggunakan pendekatan *supervised learning* yaitu klasifikasi teks jadi 3 kelompok yaitu positif, negatif, netral, pendekatan ini banyak digunakan karena kemudahan akses algoritma serta proses *fine tuning*[9]. Berdasarkan penelitian [10] menggunakan 3 algoritma *supervised learning* diantaranya SVM, Naïve Bayes, dan KNN untuk analisis tweet seputar covid-19 menghasilkan nilai performa algoritma SVM sebesar 90,1%, Naïve Bayes akurasi 79,2%, dan KNN sebesar 62,1% ini membuktikan algoritma *Support Vector Machine* paling baik diantara ke 2 algoritma lainnya. Lalu pada artikel [11] melakukan analisis *sentiment* terhadap pembelajaran tatap muka dilihat dari komentar media sosial *youtube* menggunakan *deep learning* menghasilkan berupa LSTM dan LDA untuk arsitektur NN menghasilkan akurasi nilai ROC AUC 84%. Selanjutnya pada [12] penelitian ini berusaha menganalisa *sentiment* topik viral dengan *rule-based* yaitu *lexicon positif, negatif* bahasa indonesia, dimana performa dari *rule based* tidak diukur dengan metode lain, namun metode ini cepat dan mudah dioperasikan karena tidak memerlukan data dengan label seperti pendekatan *supervised learning* dan tidak memerlukan proses pelatihan model terlebih dahulu. Menurut faried dkk tentang topik pemindahan ibu kota indonesia, Total tweet yang digunakan adalah 849 *tweets* dengan data yang diambil dengan metode *crawling* periode 23 april 2020 sampai 2 mei 2020. Dalam penelitiannya peneliti menambahkan fitur seleksi PSO untuk meningkatkan akurasi model yang dibangun sebelumnya, model sebelumnya hanya menggunakan *Support Vector Machine* dan Naïve Bayes, hasil uji coba yang dilakukan peneliti terbukti bahwa PSO dapat

meningkatkan hasil klasifikasi model sebelumnya [13].

Tujuan penelitian ini adalah menganalisis *sentiment* masyarakat tentang dinamika varian covid menggunakan SVM dan PSO. Tahapan penelitian ini menggunakan NLP untuk ekstraksi data *tweet* berisi judul berita untuk selanjutnya *sentiment* harian akan disandingkan dengan data kasus harian covid-19 di Indonesia. Analisis *sentiment* varian covid dalam data teks tweets twitter belum menunjukkan akurasi yang sangat baik [14], sehingga diperlukan model dalam proses klasifikasi. *Particle Swarm Optimization* merupakan salah satu metode yang dapat digunakan untuk meningkatkan akurasi seperti dalam penelitian [13] meningkat 3,99% dengan menggunakan *Support Vector Machine*. Berdasarkan latar belakang tersebut peneliti bertujuan untuk menerapkan algoritma *Support Vector Machine* berbasis *Particle Swarm Optimization* untuk menganalisis data teks *tweets* varian covid-19 dan meningkatkan model akurasi sebelum ditambah *Particle Swarm Optimization*. Pada penelitian ini menambahkan tahapan proses pada *pre-processing*, pengambilan data, *labelling*, penelitian ini diharapkan dapat membantu memberikan informasi pada masyarakat luas apakah polaritas opini di media sosial twitter berdampak secara nyata di era pandemi ini dan jika benar, dapat mengajak masyarakat untuk lebih berhati – hati dalam bersosial media agar tidak mudah percaya atau termakan upaya fabrikasi opini oknum-oknum yang tidak bertanggung jawab.

2. METHODS

Metodologi yang digunakan dalam penelitian ini adalah *proses text mining*. *Sub-task* dari proses *text mining* yaitu *categorization*, *information retrieval*, *POS tagging*, *Clustering*, dan lainnya. Proses ini bisa masuk kedalam kategori “*Knowledge Discovery Database*” (KDD), yaitu proses untuk menemukan informasi yang benar, baru dan juga menemukan pola dari data yang besar, kompleks yang berguna dan dimengerti. Dalam melakukan penelitian ini, peneliti menggunakan beberapa metode yang bisa menunjang penelitian ini[15].

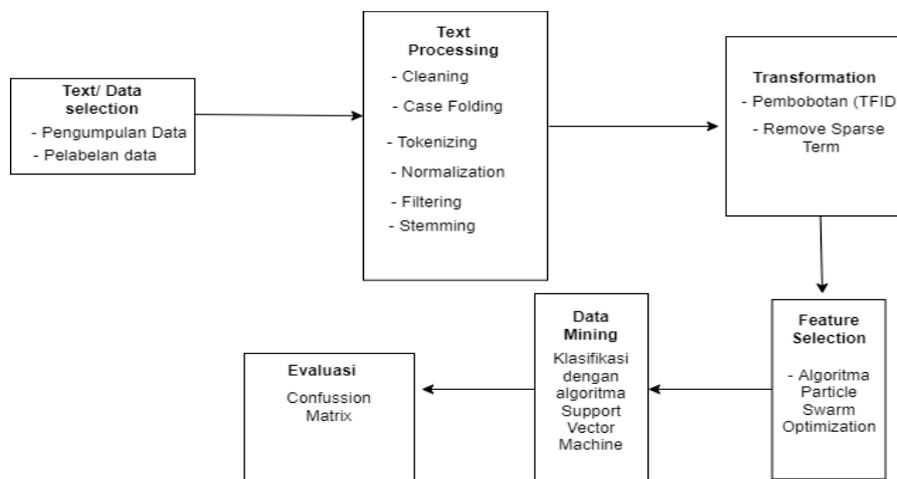


Figure 1. Alur Penelitian Text Mining

3. RESULTS AND DISCUSSION

3.1 Hasil Penelitian

Pada penelitian ini menggunakan alur proses *text mining*, penelitian yang dilakukan akan menghasilkan hasil analisis *sentiment* masyarakat tentang opini mereka di sosial media twitter tentang varian covid baik itu bernilai positif, negatif, dan netral. Dan juga akan dilakukan proses komparasi yaitu penerapan algoritma *support vector machine* dan *support vector machine* berbasis *particle swarm optimization* untuk melihat mana hasil yang terbaik dari skenario yang telah dibuat. Data yang digunakan diambil dari bulan Januari sampai oktober 2021.

3.2 Data Selection

Pada penelitian ini data tweets yang diambil berjumlah 8993 tweets yang diambil dari tanggal 1 Januari 2021 sampai 31 Oktober 2021 tentang varian covid, data tweets ini sudah dilakukan *labelling* dengan menggunakan *library* yang di dapat dari penelitian [16] juga dibantu dengan *ruled based* rumus skor polaritas yang juga didapatkan dari penelitian [17].

| ID | Username | Text | Keyword | Date | Cleaned Text | Sentiment |
|---------------------|---------------|------------------------|--------------------------|------------------------|--------------------------|-----------|
| 1344932201998082048 | anak_ancol | @cnbcindonesia Si c... | Beijing Lockdown Viru... | 2021-01-01 09:03:19... | si covid pulkam mau i... | Neutral |
| 1344931082311847937 | iman_manul | @cnbcindonesia Alah... | Beijing Lockdown Viru... | 2021-01-01 08:58:52... | alah asalnya dr wuha... | Neutral |
| 1344947417200803840 | gorarhiezv | @cnbcindonesia cou... | Beijing Lockdown Viru... | 2021-01-01 10:03:46... | counter attack | Neutral |
| 1344997445059575812 | shirohige2025 | @cnbcindonesia uda... | Beijing Lockdown Viru... | 2021-01-01 13:22:34... | udah bener ini dikem... | Neutral |
| 1344953012591104000 | fahry54 | @cnbcindonesia akhi... | Beijing Lockdown Viru... | 2021-01-01 10:26:00... | akhir nya kita swase... | Positive |
| 1344948509053743104 | ManofBravo | @cnbcindonesia 🤔... | Beijing Lockdown Viru... | 2021-01-01 10:08:07... | | Neutral |
| 1344981800695648256 | setyotomo16 | @cnbcindonesia Kok... | Beijing Lockdown Viru... | 2021-01-01 12:20:24... | kok impor lah wong c... | Neutral |
| 1344950347693846528 | Andre_wow26 | @cnbcindonesia gim... | Beijing Lockdown Viru... | 2021-01-01 10:15:25... | gimana cara nya | Neutral |
| 1345000331537653762 | iammsofyan | @cnbcindonesia Buk... | Beijing Lockdown Viru... | 2021-01-01 13:34:02... | bukan impor tp ciretur | Neutral |
| 1345201617918857216 | adhi_kusumo | @cnbcindonesia Pen... | Maia Estianty Hari S... | 2021-01-02 02:53:53... | penguasa | Positive |

Figure 2. Data Tweets Varian Covid

Setelah dilakukan pelabelan data yang sudah terkumpul tersebut dalam format csv akan dibagi menjadi 2 yaitu data *training* dan data *testing* dengan skenario pengujian diantaranya (60:40), (70:30), (80:30), (90:10).

3.3 Pre-processing

Pada tahap *pre-processing* ada beberapa tahap yang wajib dilakukan yaitu *cleaning*, *case folding*, *tokenizing*, *normalization*, *filtering*, *stemming*. Berikut penjabarannya sebagai berikut.

1. Cleaning

Pada tahap ini tweets yang diperoleh dari hasil *scrapping* yang sudah ada dalam format csv akan membersihkan emoji, *username*, *url*, hastag, spasi yang berlebih, angka, tanda baca. Berikut adalah hasil dari *cleaning*.

| Text | setelah cleaning |
|--|--|
| @cnbcindonesia Si covid pulkam mau imlekan 😊 | Si covid pulkam mau imlekan |
| @cnbcindonesia Alah asalnya dr wuhan koq, ngemeng... | Alah asalnya dr wuhan koq ngemeng buat nyalah... |
| @cnbcindonesia counter attack 😊 | counter atack |
| @cnbcindonesia udah bener ini dikembalikan lg ... | udah bener ini dikembalikan lg ke yg punya |
| @cnbcindonesia akhir nya kita swasembada 🙏 Sw... | akhir nya kita swasembada Swasembada Covid |

Figure 3. Sebelum dan Sesudah Cleaning

2. Case folding

Pada tahap ini tweets yang diperoleh diubah menjadi huruf kecil, jadi semua huruf kapital akan diubah menjadi huruf kecil. Berikut hasil dari *case folding*.

| sebelum case folding | setelah case folding |
|---|--|
| 0 Si covid pulkam mau imlekan | si covid pulkam mau imlekan |
| 1 Alah asalnya dr wuhan koq ngemeng buat nyalah... | alah asalnya dr wuhan koq ngemeng buat nyalah... |
| 2 counter atack | counter atack |
| 3 udah bener ini dikembalikan lg ke yg punya | udah bener ini dikembalikan lg ke yg punya |
| 4 akhir nya kita swasembada Swasembada Covid | akhir nya kita swasembada swasembada covid |
| ... | ... |
| 8988 Beruntung kita gak sampai lockdown | beruntung kita gak sampai lockdown |
| 8989 Zaman terawan san sehari berarti ini zama ape... | zaman terawan san sehari berarti ini zama ape... |
| 8990 cluster pilkoplo | cluster pilkoplo |
| 8991 Vaksin polio | vaksin polio |
| 8992 wow | wow |

Figure 4. Sebelum dan Sesudah Case folding

3. Tokenizing

Pada tahap ini dilakukan pemecahan suatu teks menjadi kata-kata, jadi pada tahap ini kalimat tweets menjadi terpisah atau menjadi potongan-potongan kata berdasarkan tiap kata penyusunnya. Berikut hasil dari proses *tokenizing*.

| sebelum tokenizing | setelah tokenizing |
|--|---|
| si covid pulkam mau imlekan | [si, covid, pulkam, mau, imlekan] |
| alah asalnya dr wuhan koq ngemeng buat nyalah... | [alah, asalnya, dr, wuhan, koq, ngemeng, buat, ...] |
| counter atack | [counter, atack] |
| udah bener ini dikembalikan lg ke yg punya | [udah, bener, ini, dikembalikan, lg, ke, yg, p...] |
| akhir nya kita swasembada swasembada covid | [akhir, nya, kita, swasembada, swasembada, covid] |
| ... | ... |
| beruntung kita gak sampai lockdown | [beruntung, kita, gak, sampai, lockdown] |
| zaman terawan san sehari berarti ini zama ape... | [zaman, terawan, san, sehari, berarti, ini, za...] |
| cluster pilkoplo | [cluster, pilkoplo] |
| vaksin polio | [vaksin, polio] |
| wow | [wow] |

Figure 5. Sebelum dan Sesudah Tokenizing

4. Normalization

Pada tahap ini dilakukan proses mengubah suatu kata dalam kalimat menjadi lebih tepat. Berikut hasil proses *normalization*.

| sebelum normalisasi | setelah normalisasi |
|--|---|
| [si, covid, pulkam, mau, imlekan] | [si, corona, pulang, kampung, mau, imlekan] |
| [counter, atack] | [serangan, balik] |
| [udah, bener, ini, dikembalikan, lg, ke, yg, p...] | [sudah, benar, ini, dikembalikan, lagi, ke, ya...] |
| [akhir, nya, kita, swasembada, swasembada, covid] | [akhir, ya, kita, swasembada, swasembada, corona] |
| ... | ... |
| [beruntung, kita, gak, sampai, lockdown] | [beruntung, kita, tidak, sampai, lockdown] |
| [zaman, terawan, san, sehari, berarti, ini, za...] | [zaman, dokter, terawan, san, sehari, berarti, ...] |

Figure 6. Sebelum dan Sesudah Normalization

5. Filtering

Pada tahap ini dilakukan proses penghilangan kata yang tidak penting atau kata sambung dengan menggunakan *teknik removed stopword* yang ada pada *library* sastrawi. Berikut hasil dari proses *filtering*.

| sebelum filtering | setelah filtering |
|--|------------------------------------|
| [si, corona, pulang, kampung, mau, imlekan] | [corona, pulang, kampung, imlekan] |
| [serangan, balik] | [serangan, balik] |
| [sudah, benar, ini, dikembalikan, lagi, ke, ya...] | [benar, dikembalikan, punya] |
| [akhir, ya, kita, swasembada, swasembada, corona] | [akhir, corona] |

Figure 7. Sebelum dan Sesudah Filtering

6. Stemming

Pada tahap ini melakukan proses mengubah kata berimbuhan kembali menjadi kata dasarnya, proses ini juga menggunakan *library* sastrawi. Hasil dari proses *stemming* sebagai berikut.

| sebelum stemming | setelah stemming |
|------------------------------------|----------------------------------|
| [corona, pulang, kampung, imlekan] | [corona, pulang, kampung, imlek] |
| [serangan, balik] | [serang, balik] |
| [benar, dikembalikan, punya] | [benar, kembali, punya] |
| [akhir, corona] | [akhir, corona] |
| ... | ... |
| [beruntung] | [untung] |

Figure 8.. Sebelum dan Sesudah Stemming

3.4 Transformation

Pada tahapan proses ini dilakukan proses pembobotan kata dengan mencari nilai *Term Frekuensi* dengan TF-IDF. Proses pembobotan nilai pada setiap kata akan dihitung probabilitasnya dengan mencari nilai TF dan *Inverse Document Frequency* (IDF) dengan rumus yang telah ditentukan, setelah didapatkan kedua nilai tersebut akan dikalikan, selanjutnya data tersebut diubah menjadi data numerik dalam proses *vectorize*. Berikut hasil proses TF -IDF pada **Gambar 9**.

| | |
|--------------|---------------------|
| (0, 2527) | 0.7031609740567474 |
| (0, 2990) | 0.47359025852526215 |
| (0, 5495) | 0.46044322720559977 |
| (0, 1221) | 0.2631899430336084 |
| (1, 4881) | 0.2256596035250557 |
| (1, 5854) | 0.31646586645995023 |
| (1, 935) | 0.25528669801645193 |
| (1, 4645) | 0.3729548430491121 |
| (1, 3357) | 0.41217002904797523 |
| (1, 7301) | 0.44619953383279776 |
| (1, 365) | 0.3888411198284667 |
| (1, 179) | 0.35282419454632935 |
| (2, 517) | 0.6458259219359783 |
| (2, 6124) | 0.7634846943819789 |
| (3, 5509) | 0.4959237910684101 |
| (3, 3170) | 0.7605276610462014 |
| (3, 681) | 0.41911486520514896 |
| (4, 113) | 0.8322285187989065 |
| (4, 1221) | 0.5544327664357764 |
| (5, 4339) | 0.42329382799291526 |
| (5, 1223) | 0.578034360998279 |
| (5, 5572) | 0.31561535746854147 |
| (5, 2530) | 0.5275659879689019 |
| (5, 3362) | 0.3297873362590153 |
| (6, 1071) | 1.0 |
| : | : |
| (8243, 4363) | 0.39412043328771834 |
| (8243, 3748) | 0.2782813295087745 |
| (8243, 1155) | 0.23716969691840473 |
| (8243, 4339) | 0.4914163361669266 |
| (8243, 2990) | 0.2654464692685795 |
| (8243, 5495) | 0.25807758238303635 |
| (8243, 1221) | 0.14751747922944938 |
| (8244, 5663) | 1.0 |
| (8245, 268) | 0.4693031722200944 |
| (8245, 2485) | 0.6493532423595052 |
| (8245, 2031) | 0.5984103100560302 |
| (8246, 6528) | 1.0 |
| (8247, 7063) | 1.0 |

Figure 9. Hasil frequency TF-IDF

3.5 Data mining

Pada tahap ini dilakukan proses pemodelan dengan algoritma *support vector machine* (SVM) dan algoritma *support vector machine* yang dioptimasi dengan *particle swarm optimization*(PSO). Pada pemodelan ini menggunakan 4 skenario dengan pembagian data *training* dan *testing* masing-masing yaitu, 60:40, 70:30, 80:20, 90:10. Dari skenario yang telah dibuat bertujuan untuk mendapatkan nilai terbaik jika jumlah data *training* lebih besar dari data *testing*. Hasil proses klasifikasi akan menghasilkan nilai *precision*, *akurasi*, *recall*. Berikut gambaran hasil pembagian data *training* dan *testing*.

Tabel 1. Pembagian Data Training dan Testing

| Skenario | Data Training(%) | Data Testing(%) |
|------------|--------------------|-------------------|
| Skenario 1 | Data Training(60%) | Data Testing(40%) |
| | 600 | 400 |
| Skenario 2 | Data Training(70%) | Data Testing(30%) |
| | 700 | 300 |
| Skenario 3 | Data Training(80%) | Data Testing(20%) |
| | 800 | 200 |

| | | |
|------------|----------------------------|---------------------------|
| Skenario 4 | Data <i>Training</i> (90%) | Data <i>Testing</i> (10%) |
| | 900 | 100 |

Tabel 2 Hasil Perbandingan Evaluasi Support Vector Machine

| Skenario | Kernel | Akurasi(%) |
|------------|-------------------|------------|
| Skenario 1 | <i>Linear</i> | 81 |
| | <i>Polynomial</i> | 72 |
| Skenario 2 | <i>Linear</i> | 82 |
| | <i>Polynomial</i> | 73 |
| Skenario 3 | <i>Linear</i> | 82 |
| | <i>Polynomial</i> | 74 |
| Skenario 4 | <i>Linear</i> | 83 |
| | <i>Polynomial</i> | 74 |

Tabel 3. Hasil Perbandingan Presisi, Recall. F1-score Support Vector Machine

| Skenario | Presisi | | | Recall | | | F1-Score | | |
|------------|---------|------|------|--------|------|------|----------|------|------|
| | -1 | 0 | 1 | -1 | 0 | 1 | -1 | 0 | 1 |
| Skenario 1 | 0.84 | 0.77 | 0.88 | 0.58 | 0.93 | 0.73 | 0.68 | 0.84 | 0.80 |
| | 1.00000 | 0.65 | 0.95 | 0.29 | 0.99 | 0.49 | 0.45 | 0.79 | 0.64 |
| Skenario 2 | 0.87 | 0.79 | 0.89 | 0.60 | 0.94 | 0.74 | 0.71 | 0.86 | 0.81 |
| | 0.99 | 0.67 | 0.95 | 0.32 | 0.99 | 0.51 | 0.48 | 0.79 | 0.67 |
| Skenario 3 | 0.83 | 0.80 | 0.89 | 0.61 | 0.93 | 0.76 | 0.70 | 0.86 | 0.82 |
| | 0.99 | 0.67 | 0.96 | 0.33 | 0.99 | 0.54 | 0.50 | 0.80 | 0.69 |
| Skenario 4 | 0.88 | 0.79 | 0.88 | 0.62 | 0.93 | 0.75 | 0.73 | 0.86 | 0.81 |
| | 0.98 | 0.68 | 0.94 | 0.35 | 0.98 | 0.55 | 0.52 | 0.81 | 0.70 |

Tabel 4. Hasil Perbandingan Evaluasi Support Vector Machine Dengan Pso

| Skenario | Kernel | Akurasi (%) |
|------------|-------------------|-------------|
| Skenario 1 | <i>Linear</i> | 84 |
| | <i>Polynomial</i> | 84 |
| Skenario 2 | <i>Linear</i> | 85 |
| | <i>Polynomial</i> | 85 |
| Skenario 3 | <i>Linear</i> | 85 |
| | <i>Polynomial</i> | 86 |

| Skenario | <i>Linear</i> | | 86 | | | | | |
|----------|-------------------|------|--------|------|------|----------|------|------|
| | <i>Polynomial</i> | | 87 | | | | | |
| Presisi | | | Recall | | | F1-Score | | |
| -1 | 0 | 1 | -1 | 0 | 1 | -1 | 0 | 1 |
| 0.81 | 0.83 | 0.88 | 0.72 | 0.90 | 0.80 | 0.76 | 0.86 | 0.84 |
| 0.82 | 0.82 | 0.89 | 0.67 | 0.92 | 0.80 | 0.74 | 0.87 | 0.84 |
| 0.82 | 0.84 | 0.88 | 0.71 | 0.91 | 0.82 | 0.76 | 0.87 | 0.85 |
| 0.80 | 0.85 | 0.88 | 0.71 | 0.90 | 0.83 | 0.75 | 0.87 | 0.86 |
| 0.81 | 0.84 | 0.90 | 0.70 | 0.91 | 0.84 | 0.75 | 0.88 | 0.87 |
| 0.82 | 0.85 | 0.91 | 0.71 | 0.92 | 0.85 | 0.76 | 0.88 | 0.88 |
| 0.84 | 0.84 | 0.90 | 0.72 | 0.92 | 0.83 | 0.78 | 0.88 | 0.86 |
| 0.85 | 0.85 | 0.93 | 0.74 | 0.94 | 0.83 | 0.79 | 0.89 | 0.88 |

Berdasarkan tabel diatas dapat dilihat bahwa hasil akurasi tertinggi dari svm yang telah dilakukan proses optimasi dengan *pso* yaitu kernel *linear* dan *polynomial* pada skenario 4 dengan pembagian data 90:10 senilai 86% dan 87%. Nilai evaluasi yang lain juga mengalami peningkatan seperti presisi menjadi 90%, *recall* menjadi 83% dan *f1-score* menjadi 86%. Hasil akurasi terkecil dari optimasi ini ada pada skenario 1 dengan pembagian data 60:40 pada kernel *linear* dan *polynomial* dengan akurasi 84%.

4. CONCLUSION

Dari hasil penelitian yang telah dilakukan dapat disimpulkan beberapa hal, diantaranya yaitu Klasifikasi hasil analisis *sentiment* varian covid dengan svm lalu dilakukan proses *preprocessing* seperti *cleaning*, *case folding*, *tokenizing*, *normalization*, *stemming*, pembobotan kata dengan (TF-IDF), selanjutnya dilakukan pemodelan dengan algoritma *support vector machine* dioptimasi dengan *particle swarm optimization*. Kemudian dengan membandingkan hasil model svm tanpa *pso* dan svm dengan *pso*, rasio perbandingan data training dan testing pada skenario 4 yaitu 90:10 dengan optimasi PSO memiliki performa yang paling tinggi, dapat disimpulkan bahwa penggunaan optimasi PSO berpengaruh dalam hasil performa uji model yang telah dilakukan.

REFERENCES

- [1] F. Muhammad, N. M. Maghfur, and A. Voutama, "Sentiment Analysis Dataset on COVID-19 Variant News," *Systematics*, vol. 4, no. 1, pp. 382–391, 2022.
- [2] N. Akbulaev, I. Mammadov, and V. Aliyev, *Economic Impact of COVID-19*. 2020.
- [3] G. Xu *et al.*, "Since January 2020 Elsevier has created a COVID-19 resource centre with free

- information in English and Mandarin on the novel coronavirus COVID- 19 . The COVID-19 resource centre is hosted on Elsevier Connect , the company ' s public news and information ,” no. January, 2020.
- [4] O. Dyer, “Covid-19: Indonesia becomes Asia’s new pandemic epicentre as delta variant spreads,” *BMJ*, vol. 374, no. July, p. n1815, 2021, doi: 10.1136/bmj.n1815.
- [5] A. Wong, S. Ho, O. Olusanya, M. V. Antonini, and D. Lyness, “The use of social media and online communications in times of pandemic COVID-19,” *J. Intensive Care Soc.*, vol. 22, no. 3, pp. 255–260, 2021, doi: 10.1177/1751143720966280.
- [6] E. Novalia, J. Na’am, G. W. Nurcahyo, and A. Voutama, “Website Implementation with the Monte Carlo Method as a Media for Predicting Sales of Cashier Applications,” *Systematics*, vol. 2, no. 3, pp. 118–131, 2020.
- [7] H. Sahni and H. Sharma, “Role of social media during the COVID-19 pandemic: Beneficial, destructive, or reconstructive?,” *Int. J. Acad. Med.*, vol. 6, no. 2, pp. 70–75, 2020, doi: 10.4103/IJAM.IJAM_50_20.
- [8] R. B. Koyel Chakraborty, Siddhartha Bhattacharyya, “A Survey of Sentiment Analysis from Social Media Data,” vol. 7, pp. 450–464, 2020, [Online]. Available: <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8951256&isnumber=9055454>.
- [9] E. Miranda, M. Aryuni, R. Hariyanto, and E. S. Surya, “Sentiment Analysis using Sentiwordnet and Machine Learning Approach (Indonesia general election opinion from the twitter content),” *Proc. 2019 Int. Conf. Inf. Manag. Technol. ICIMTech 2019*, vol. 1, no. June, pp. 62–67, 2019, doi: 10.1109/ICIMTech.2019.8843734.
- [10] F. Sodik and I. Kharisudin, “Analisis Sentimen dengan SVM , NAIVE BAYES dan KNN untuk Studi Tanggapan Masyarakat Indonesia Terhadap Pandemi Covid-19 pada Media Sosial Twitter,” *Prisma*, vol. 4, pp. 628–634, 2021.
- [11] F. Z. Ahmad, M. Fauzi, S. Arifandy, M. R. Caesarardhi, and N. Aini, “Bagaimana Masyarakat Menyikapi Pembelajaran Tatap Muka : Analisis Komentar Masyarakat pada Media Sosial Youtube Menggunakan Algoritma Deep Learning Sekuensial dan LDA,” vol. 4, no. 2, pp. 40–46, 2021.
- [12] R. Arief and K. Imanuel, “Analisis Sentimen Topik Viral Desa Penari Pada Media Sosial Twitter Dengan Metode Lexicon Based,” *J. Ilm. Matrik*, vol. 21, no. 3, pp. 242–250, 2019, doi: 10.33557/jurnalatrik.v21i3.727.
- [13] P. Arsi, R. Wahyudi, and R. Waluyo, “Optimasi SVM Berbasis PSO pada Analisis Sentimen Wacana Pindah Ibu Kota Indonesia,” *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 5, no. 2, pp. 231–237, 2021, doi: 10.29207/resti.v5i2.2698.
- [14] N. M. Maghfur, F. Muhammad, and A. Voutama, “Analysis of the Relationship between Public

- Sentiment on Social Media and Indonesian Covid-19 Dynamics,” vol. 3, no. 3, pp. 336–345, 2021.
- [15] A. Voutama and D. Wahyono, “Perancangan Sistem Informasi Transaksi Penjualan pada Toko Bata Kota Solok,” *Systematics*, vol. 2, no. 1, p. 39, 2020, doi: 10.35706/sys.v2i1.3637.
- [16] D. H. Wahid and A. SN, “Peringkasan Sentimen Esktraktif di Twitter Menggunakan Hybrid TF-IDF dan Cosine Similarity,” *IJCCS (Indonesian J. Comput. Cybern. Syst.*, vol. 10, no. 2, p. 207, 2016, doi: 10.22146/ijccs.16625.
- [17] Z. Wu and D. C. Ong, “Context-Guided BERT for Targeted Aspect-Based Sentiment Analysis,” 2020, [Online]. Available: <http://arxiv.org/abs/2010.07523>.