

IMPLEMENTASI ALGORITMA MULTICLASS SVM PADA OPINI PUBLIK BERBAHASA INDONESIA DI TWITTER

Debby Alita*¹⁾, Yusra Fernando²⁾, Heni Sulistiani³⁾

^{1), 2)} Informatika, Universitas Teknokrat Indonesia

³⁾ Sistem Informasi Akuntansi, Universitas Teknokrat Indonesia

Jl. H.ZA Pagaralam, No 9-11, Labuhanratu, Bandar Lampung

Email : debbyalita@teknokrat.ac.id¹⁾, yusra.fernando@teknokrat.ac.id²⁾, henisulistiani@teknokrat.ac.id³⁾

Abstrak

Klasifikasi pada proses text mining dapat dikerjakan dengan menggunakan berbagai jenis metode klasifikasi yang salah satunya yaitu metode SVM. SVM merupakan singkatan dari Support Vector Machine, SVM bekerja dengan membagi dua kelompok kelas data menggunakan fungsi linear dalam sebuah ruang fitur berdimensi tinggi dengan proses menemukan garis pemisah (hyperplane) terbaik sehingga dapat menemukan ukuran margin yang maksimal antara ruang input dengan ruang ciri menggunakan kaidah kernel. SVM telah dikembangkan dengan menggabungkan semua data yang terdiri dari beberapa kelas kedalam sebuah bentuk optimasi untuk memecahkan permasalahan yang terdapat pada penelitian ini dengan jumlah kelas yang melebihi dari dua kelas dan akan diuji dengan berbagai jenis pendekatan multiclass yaitu SVM One Against One dan One Against Rest. Data merupakan opini publik berbahasa Indonesia yang didapatkan dari twitter berjumlah 2000 dataset mengenai jaringan telekomunikasi seluler dan layanan BPJS. Hasil penelitian ini didapatkan bahwa untuk penilaian kinerja metode multiclass SVM dengan tingkat akurasi yang lebih tinggi didapatkan dari kinerja metode SVM One Against Rest dengan nilai perbedaan sebesar 0,06 untuk proses klasifikasi tiga kelas yaitu positif, negatif dan netral. Dapat disimpulkan bahwa dalam proses klasifikasi yang memiliki lebih dari dua kelas dapat dilakukan dengan menggunakan metode klasifikasi SVM melalui pendekatan SVM One Against One dan One Against Rest dengan nilai akurasi yang lebih baik.

Kata kunci: Text Mining, Support Vector Machine, One Against One, One Against Rest, Klasifikasi

1. Pendahuluan

Support Vector Machine (SVM) merupakan salah satu metode yang sudah banyak diterapkan untuk berbagai jenis penelitian dibidang data dan text mining karena telah mampu menunjukkan performa yang lebih baik (Styawati and Mustofa, 2019). SVM bekerja dengan sistem pembelajaran yang menggunakan ruang hipotesis berupa fungsi-fungsi linier dalam sebuah ruang fitur berdimensi tinggi, hanya saja pada metode klasifikasi SVM hanya dapat mengklasifikasikan data kedalam dua kelas (Pramudita and Musdholifah, 2020). Optimasi

SVM terus dilakukan peningkatan sehingga SVM dapat mengklasifikasikan data menjadi lebih dari dua kelas yaitu dengan model pendekatan One Against One dan One Against Rest, kedua pendekatan tersebut dapat membantu proses klasifikasi multiclass pada penelitian ini yang memiliki lebih dari dua kelas klasifikasi. Data yang akan diklasifikasikan merupakan data yang berasal dari media sosial yang berisikan komentar-komentar publik.

Media sosial merupakan aktivitas sosial yang menggunakan jaringan online berupa bahasa, gambar dan video (Purawinangun 2020) Aktivitas sosial tersebut dianggap sangat memudahkan seseorang dalam berdiskusi, berbisnis dan berkomentar secara bebas. (Isnain, Sihabuddin and Suyanto, 2020) menuliskan salah satu media sosial yang masih sering digunakan untuk mengekspresikan opini seseorang secara bebas adalah twitter. Twitter digunakan oleh semua kalangan untuk memberikan tanggapan netral, positif bahkan tanggapan negatif terhadap suatu trending topic (Alita, Priyanta and Rokhman, 2019). (Yunitasari, Musdholifah and Sari, 2019) mengungkapkan bahwa tanggapan/opini yang berasal dari twitter ini sangat mudah untuk di analisis secara realtime dengan dibantu oleh piranti lunak yang telah banyak digunakan sehingga banyak penelitian yang memperoleh data dari twitter. Data opini selanjutnya akan dianalisis untuk dijadikan sumber informasi.

Opinion Mining merupakan rangkaian kegiatan pengelolaan data tekstual yang dilakukan secara otomatis untuk memperoleh informasi yang akan digunakan sebagai bahan evaluasi perbaikan karena perkembangan yang sangat pesat khususnya dibidang e-commerce menuntut para stakeholder untuk meningkatkan mutu produk melalui pendapat publik tentang apa yang sedang masyarakat butuhkan saat ini (Pozzi et al., 2016).

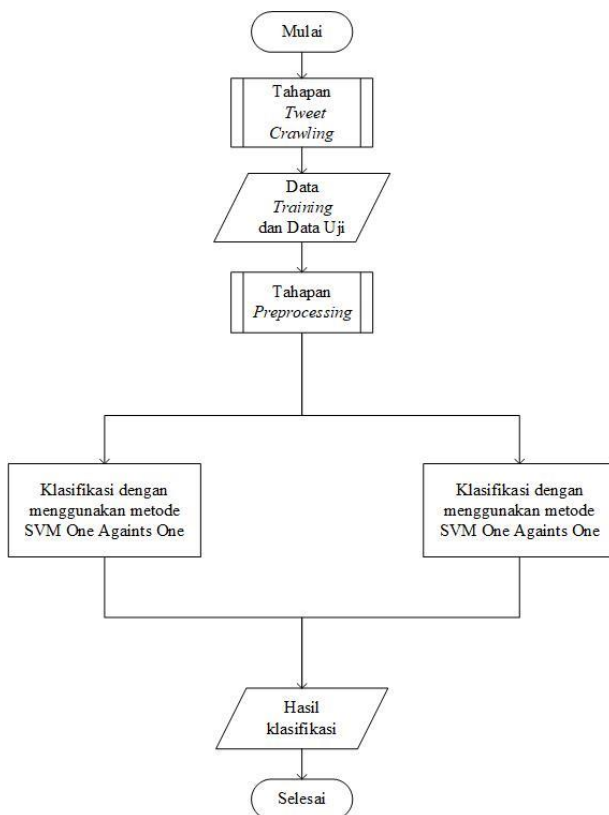
Tujuan utama penelitian ini adalah untuk melihat performa dari metode Support Vector Machine yang telah dioptimasi dan diterapkan pada proses analisis opini berbahasa Indonesia yang memiliki tiga kelas klasifikasi yaitu positif, negatif dan netral.

2. Pembahasan

Proses klasifikasi data opini dilakukan dengan menggunakan metode support vector machine yang telah dioptimasi untuk dapat mengklasifikasikan data multiclass dengan data label positif, negatif dan netral. Proses analisis opini terdiri dari 3 tahapan yaitu pengumpulan data, preprocessing dan klasifikasi.

2.1. Rancangan Proses

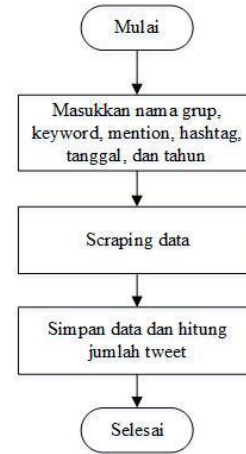
Rancangan proses merupakan penggambaran urutan alur kerja dari proses itu dimulai hingga proses tersebut berhenti. Alur kerja dari suatu proses dengan proses lainnya dapat digambarkan dengan suatu bagan yang disebut dengan flowchart. Pada penelitian ini, proses yang diusulkan pada sistem terdiri dari tahapan awal yaitu tweet crawling yang merupakan proses pengambilan data dari twitter kemudian setelah mendapatkan data dilanjutkan dengan proses labeling data yang dilakukan oleh mahasiswa Linguistik, selanjutnya dilakukan tahapan preprocessing yaitu pengolahan data text sesuai dengan kebutuhan dari proses analisis sentimen, selanjutnya dilakukan proses ekstraksi fitur yakni mempermudah dalam melakukan proses klasifikasi maka data text akan diubah kedalam bentuk numerik dan selanjutnya dilakukan tahapan klasifikasi menggunakan metode multiclass SVM dan terakhir dilakukan pengujian dan evaluasi. Secara garis besar langkah-langkah yang dilakukan dapat dilihat pada gambar 1



Gambar 1. Rancangan Proses Umum

2.2. Tweet Crawling

Tahapan tweet crawling ini adalah melakukan pengambilan data tweet dari twitter.com dengan menggunakan API yang tersedia pada twitter. Langkah-langkah dari tweet crawling dapat dilihat pada gambar 2.

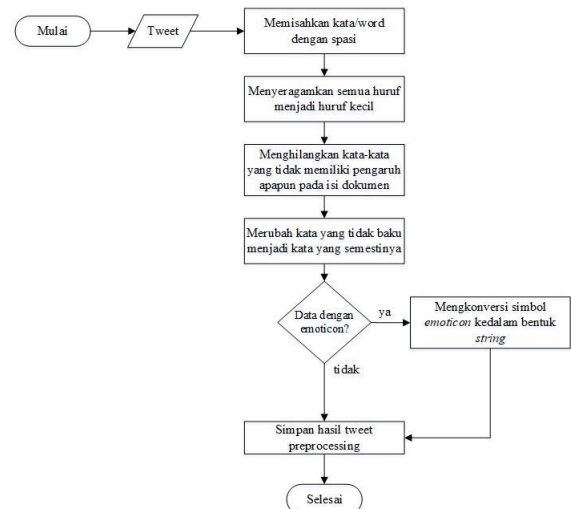


Gambar 2. Langkah-langkah tweet crawling

Pada tahap pengumpulan data dilakukan dengan cara scraping halaman pencarian twitter. Proses scraping dimulai dengan user memasukkan query pencarian twitter yang dapat berupa keyword, mention, hashtag, tanggal, dan tahun.

2.3 Tahapan Preprocessing

Tahapan preprocessing merupakan tahapan yang dilakukan sebelum proses klasifikasi yaitu untuk mengelola teks pada data tweet (Negara, Muhardi and Putri, 2020). Tahapan preprocessing yang digunakan pada penelitian ini dapat dilihat pada gambar 3



Gambar 3. Tahapan Preprocessing

1. Tokenisasi

Tokenisasi digunakan untuk pemberian jarak antar kata dan spasi.

2. *Case Folding*
Case folding digunakan untuk merubah semua huruf menjadi huruf kecil .
3. *Stopword*
Stopword digunakan untuk menghilangkan kata yang tidak memiliki arti dan tidak diperlukan dalam proses *preprocessing* kata seperti link url, kata-kata saat, akan, bagi, dengan, dan menghilangkan tweet mention (@debbyalita)
4. *Slangword*
5. *Slangword* digunakan untuk merubah kata yang tidak baku dan kata yang tidak ada di dalam kamus kedalam kata yang terdekat dengan kamus dengan melihat pola kemunculan kata-kata tidak baku.
6. *Convert Emoticon*
Convert emoticon digunakan untuk mengkonversikan *emoticon* kedalam bentuk *string* yang bersesuaian.

2.4 Ekstraksi Fitur

Proses ekstraksi fitur digunakan untuk menemukan fitur-fitur yang ada pada sebuah text , ekstraksi fitur yang digunakan pada penelitian ini adalah

1. Unigram

Pada proses unigram, data tweet akan dipotong-potong berdasarkan spasi dan dijadikan fitur.

2. POS Tagging

Pada proses POS Tagging akan dilakukan pengkategorian kelas kata dan dalam proses penentuan fitur akan diambil kata yang dengan kategori tag kata sifat dan kata kerja

3. TF dan TF-IDF

TF dan TF-IDF merupakan proses pembobotan fitur kata pada tweet (Haryalesmana Wahid, 2016). TF IDF merupakan gabungan dari Term Frequency (TF) dan Invers Document Frequency (IDF) yang digunakan dalam menghitung bobot setiap kata (term) pada setiap dokumen.

2.5 Klasifikasi menggunakan Support Vector Machine

2.5.1 Representasi Data

Dalam penelitian ini untuk representasi data menggunakan format sparse data representation dengan vector sebagai inputnya. Format data input untuk klasifikasi SVM dalam penelitian ini adalah [+1 1:0,65 2:0.394]. Dengan masukan pertama +1 atau -1 menyatakan dua kelas. Angka kedua menyatakan dimensi (row_id) dan angka ketiga setelah tanda “:” menyatakan bobot dari term, setiap term didalam dokumen dipisahkan dengan spasi, [label row-id:bobot(row-id) row-idn:bobot(row-idn)].

Data training dan data uji akan diubah menjadi data vektor. Data training adalah data latih artikel yang sudah memiliki kelas. Sedangkan data uji adalah data uji artikel yang belum memiliki kelas. Data training digunakan untuk melakukan proses pembelajaran terhadap sistem.

Proses pembelajaran ini akan menghasilkan model baru yang akan digunakan pada klasifikasi.

2.5.2 Analisis Pross SVM Training

Pada proses pelatihan SVM bertujuan untuk menemukan vektor α dan konstanta b yang akan digunakan dalam menemukan fungsi pemisah (Hyperplane) dengan margin yang maksimal. Vektor $\alpha > 0$ dinamakan support vector dan menyatakan data training yang diperlukan untuk mewakili fungsi keputusan yang optimal. Konstanta b menentukan lokasi fungsi pemisah relative terhadap titik asal. Dalam proses pelatihan dibutuhkan satu set input-output data atau dalam kasus ini akan digunakan dokumen yang memiliki kelas positif, kelas negatif dan kelas netral (Klyueva, 2019). Pada penelitian ini menggunakan SVM non linear, maka perlu menggunakan fungsi kernel atau kernel trick Fungsi kernel yang digunakan dalam penelitian ini adalah fungsi kernel RBF.

2.5.3 Klasifikasi menggunakan SVM *One Against One*

Dengan metode ini perlu menemukan $k(k-1)/2$ fungsi pemisah dimana setiap fungsi dilatih dengan data dari dua kelas. Dalam metode ini diambil contoh kasus yang telah melalui tahap pembobotan TF-IDF. Nilai x pada tabel akan digunakan untuk perhitungan kernel. Tahap selanjutnya yaitu melakukan kernelisasi menggunakan fungsi RBF yang didefinisikan sebagai $K(x,y) = \exp(-\gamma \|x - y\|^2)$, $\gamma > 0$.

Selanjutnya dilakukan kernelisasi pada set data dari fitur dimensi lama sehingga mendapatkan set data dengan fitur baru dimensi tinggi. Dengan kernel $K(x,xi) = \exp(-\gamma \|x - y\|^2)$, $\gamma > 0$, dan set data berdimensi $N \times 1$ maka akan didapatkan dimensi baru $N \times N$, dimana N adalah banyaknya data.

Prinsip dari metode ini yaitu dibangun $(k-1)2$ buah model klasifikasi biner (k adalah jumlah kelas). Setiap model klasifikasi dilatih pada data dari dua kelas.

Terdapat beberapa metode untuk melakukan pengujian setelah keseluruhan $k(k-1)/2$ model klasifikasi selesai dibangun. Salah satunya adalah metode *voting* (Klyueva, 2019).

Tabel 1. Contoh 3 SVM Biner dengan Metode *One Against One*

| $Y_i = 1$ | $Y_i = -1$ | Hipotesis |
|-----------|------------|-------------------------------|
| Kelas 1 | Kelas 2 | $f_{12}(x)=(w_{12})x+ b_{12}$ |
| Kelas 1 | Kelas 3 | $f_{13}(x)=(w_{13})x+ b_{13}$ |
| Kelas 2 | Kelas 3 | $f_{23}(x)=(w_{23})x+ b_{23}$ |

2.5.4 Klasifikasi menggunakan SVM *One Against Rest*

Pada proses klasifikasi k -kelas, ditemukan k fungsi pemisah dimana k adalah banyaknya kelas. Misalkan ada fungsi pemisah dengan nama ρ . Dalam metode ini, ρ_i dilatih dengan semua data dari kelas ke- i dengan label +1 dan semua data dari kelas lain dengan label -1. Dalam

tahap ini diambil contoh kasus yang telah melalui tahap *preprocessing* dan pembobotan TF-IDF.

Tahap selanjutnya adalah menghitung nilai x . Nilai x pada tabel akan digunakan untuk perhitungan kernel atau perhitungan *dot product*. Nilai x didapat dari nilai term pada dokumen kemudian melakukan kernelisasi menggunakan fungsi kernel RBF yang didefinisikan sebagai $K(x,y) = \exp(-\gamma \|x - y\|^2)$, $\gamma > 0$

Formulasi yang digunakan adalah dualitas Lagrange multiplier yang sudah dimodifikasi untuk x dengan fungsi kernelnya, menjadi $w = \sum_{i=1}^n \alpha_i y_i \phi(x_i)$ dengan keterangan $W = \text{vector}$ bobot, $\alpha_i = 0,5$ (konstanta), $y_i = \text{kelas data}$, $x_i = \text{data latihan}$.

Selanjutnya harus dilakukan kernelisasi pada set data dari fitur dimensi lama sehingga mendapatkan set data dengan fitur baru dimensi tinggi. Dengan kernel $K(x,y) = \exp(-\gamma \|x - y\|^2)$, $\gamma > 0$, dan set data berdimensi $N \times 1$ maka akan didapatkan dimensi baru $N \times N$, dimana N adalah banyaknya data.

Pada dasarnya prinsip dari metode ini adalah dibangun k buah model SVM biner (k adalah jumlah kelas). Setiap model klasifikasi ke- i dilatih dengan menggunakan keseluruhan data, untuk mencari solusi permasalahan. Contohnya, terdapat permasalahan klasifikasi dengan 3 buah kelas. Untuk pelatihan digunakan 3 buah SVM biner seperti pada tabel 2

Tabel 2. Contoh 3 SVM Biner dengan metode One Agiants Rest

| $Y_i = 1$ | $Y_i = -1$ | Hipotesis |
|-----------|---------------|----------------------|
| Kelas 1 | Bukan Kelas 1 | $F1(x) = (w1)x + b1$ |
| Kelas 2 | Bukan Kelas 2 | $F2(x) = (w2)x + b2$ |
| Kelas 3 | Bukan Kelas 3 | $F3(x) = (w3)x + b3$ |

2.6 Hasil dan Pembahasan

Total data twitter yang digunakan sebanyak 2.281 data tweet sentimen berbahasa Indonesia yang diperoleh dari hasil *scraping* data. Pada tahapan *preprocessing* terjadi penurunan data dikarenakan terdapat data yang tidak sesuai dengan aturan tahapan dari proses *preprocessing*. Ukuran kualitas yang akan diukur untuk menentukan kinerja dari sistem adalah akurasi, presisi, *recall*, dan *F1score* kemudian akan dihitung jumlah rata-rata ukuran kualitas.

Pada pengujian analisis sentimen menggunakan metode *Support Vector Machine* yang telah dioptimasi agar bias diklasifikasi dengan menggunakan tiga kelas klasifikasi yaitu SVM *One Against One* dan SVM *One Against Rest* dengan menggunakan *10 Fold Cross Validation* yang datanya dilakukan secara acak dan tidak acak, untuk hasil akurasi dari kedua metode SVM dapat dilihat pada tabel berikut:

Tabel 3. Nilai Akurasi Pengujian Analisis Sentimen

| Metode | Akurasi | |
|-----------------------------|---------|------------|
| | Acak | Tidak Acak |
| SVM <i>One Against One</i> | 91.94% | 92.00% |
| SVM <i>One Against Rest</i> | 92.00% | 92.00% |

Akurasi digunakan untuk mengukur performa klasifikasi berdasarkan keakuratan dari metode klasifikasi yang digunakan. Berdasarkan Tabel 3 dapat dilihat bahwa himpunan dataset yang dilkaukan secara acak dan tidak acak sama-sama memiliki nilai akurasi yang lebih baik dan untuk metodenya yang memiliki nilai akurasi lebih baik adalah SVM *One Against Rest* dengan nilai sebesar 92.00%. Kemudian untuk nilai presisi dapat dilihat pada Tabel 4 sebagai berikut:

Tabel 4. Nilai Presisi Pengujian Analisis Sentimen

| Metode | Presisi | |
|-----------------------------|---------|------------|
| | Acak | Tidak Acak |
| SVM <i>One Against One</i> | 86.93% | 86.78% |
| SVM <i>One Against Rest</i> | 86.92% | 86.78% |

Presisi merupakan tingkat ketepatan antara informasi yang diminta oleh pengguna dengan jawaban yang diberikan oleh system. Berdasarkan hasil dari Tabel 4 himpunan data set yang dilakukan secara acak memiliki performa nilai yang lebih baik dari himpunan dataset yang dilakukan secara tidak acak. Kemudian untuk penilaian kinerja metode terbaik diperoleh oleh metode SVM *One Against One* dengan perbedaan nilai yang hanya sedikit sebesar 0,1 dengan nilai 86,93%. Nilai *recall* pada pengujian analisis sentiment dapat dilihat pada Tabel 5 sebagai berikut:

Tabel 5. Nilai Recall Pengujian Analisis Sentimen

| Metode | Recall | |
|-----------------------------|--------|------------|
| | Acak | Tidak Acak |
| SVM <i>One Against One</i> | 92.00% | 92.00% |
| SVM <i>One Against Rest</i> | 91.94% | 92.00% |

Recall merupakan tingkat keberhasilan system dalam menemukan kembali sebuah informasi. Berdasarkan hasil perhitungan nilai *recall* pada Tabel 5 dapat dilihat bahwa metode SVM *One Against One* memiliki nilai lebih baik dengan himpunan data set yang dilakukan secara acak dengan nilai sebesar 92.00%. Jika dilakukan dengan teknik dataset secara tidak acak kedua metode memiliki nilai yang sama besarnya dengan kata lain

memiliki kualitas kinerja yang sama. Kemudian untuk nilai F1Score dapat dilihat pada Tabel 6 berikut:

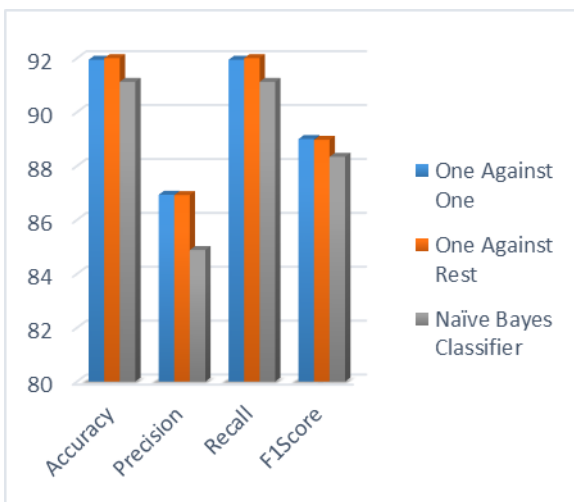
Tabel 6. Nilai F1Score Pengujian Analisis Sentimen

| Metode | F1Score | |
|-----------------------------|---------|------------|
| | Acak | Tidak Acak |
| SVM <i>One Against One</i> | 89.00% | 88.88% |
| SVM <i>One Against Rest</i> | 88.97% | 88.88% |

F1Score merupakan harmonic mean dari nilai recall dan precision ini perhitungan ini berguna untuk mengetahui seberapa presisi dan handalnya performa system dalam mengklasifikasikan kelas. Berdasarkan Tabel diatas didapatkan hasil bahwa metode SVM *One Against One* lebih unggul dengan nilai 89.00% untuk himpunan dataset yang dilakukan secara acak. Pada prose pengujian analisis sentimen menggunakan metode *Naïve Bayes Classifier* dengan teknik pengujian yang sama yaitu *10Fold cross validation* yang dilakukan secara acak dan tidak acak menghasilkan ukuran kualitas sebagai berikut:

Tabel 7. Pengujian analisis sentimen menggunakan metode *Naïve Bayes Classifier*

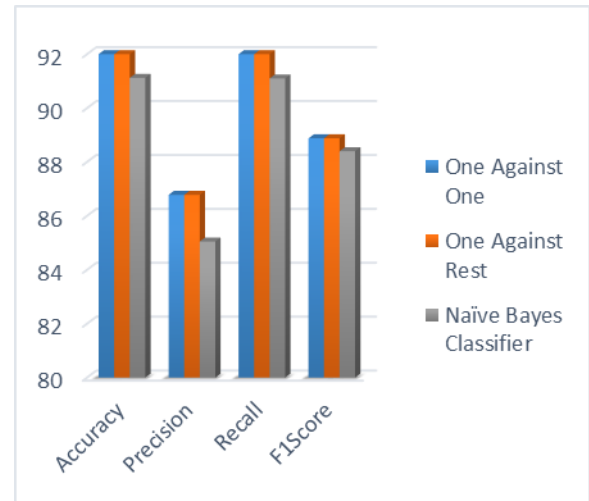
| Ukuran Kualitas | Acak | Tidak Acak |
|-----------------|--------|------------|
| Akurasi | 91.12% | 91.12% |
| Presisi | 84.88% | 85.05% |
| Recall | 91.10% | 91.10% |
| F1Score | 88.34% | 88.40% |



Gambar 4 Grafik nilai ukuran kualitas pada analisis sentiment yang dilakukan secara acak

Pada gambar 4 menampilkan grafik perbandingan pengukuran kualitas dari 3 metode klasifikasi yaitu *One Against One*, *One Against Rest* dan *Naïve Bayes Classifier* yang dilakukan secara acak bahwa metode

SVM multiclass menunjukkan hasil yang lebih baik dibandingkan dengan metode lain dan untuk kedua metode *multiclass* SVM menunjukkan bahwa metode SVM *One Against Rest* memiliki nilai akurasi yang lebih baik dibandingkan dengan metode lainnya.



Gambar 5 Grafik nilai ukuran kualitas pada analisis sentiment yang dilakukan secara tidak acak

Pada gambar 5 menampilkan grafik perbandingan pengukuran kualitas dari 3 metode klasifikasi yaitu *One Against One*, *One Against Rest* dan *Naïve Bayes Classifier* yang dilakukan secara tidak acak bahwa metode SVM multiclass menunjukkan hasil yang lebih baik dibandingkan dengan metode lain dan untuk kedua metode multiclass SVM menunjukkan bahwa kedua metode sama-sama memiliki nilai ukuran kualitas yang lebih baik dibandingkan dengan metode lainnya. Teknik dataset yang dilakukan secara acak dan tidak acak memiliki pengaruh terhadap hasil pengukuran nilai kualitas metode, dari kedua grafik diatas menunjukkan bahwa teknik dataset yang dilakukan secara acak pada saat proses klasifikasi menghasilkan nilai ukuran kualitas yang lebih baik dibandingkan dengan teknik yang dilakukan secara tidak acak.

3. Kesimpulan

Penelitian tiga kelas klasifikasi dapat dilakukan dengan menggunakan metode SVM yang telah dioptimasi yaitu dengan SVM *One Against One* dan SVM *One Against Rest* dengan hasil SVM *One Against One* lebih unggul untuk nilai presisi, recall dan F1Score, sedangkan untuk akurasi SVM *One Against Rest* lebih unggul dengan nilai perbedaan 0,06. Jika dibandingkan dengan metode lainnya yaitu metode klasifikasi *Naïve Bayes Classifier* pada proses analisis sentimen metode multiclass SVM tetap memiliki nilai lebih unggul dan untuk teknik dataset yang dilakukan secara acak memiliki nilai yang lebih baik. Penelitian selanjutnya disarankan untuk mencoba metode DAGSVM (*Directed Acyclic Graph*

Support Vector Machine) yang merupakan metode SVM Multiclass lainnya yang bisa dibandingkan dengan kedua metode pada penelitian ini dan juga disarankan untuk membuat penterjemah bahasa agar penelitian bisa bersifat continue seperti pada penelitian (Abidin Z, 2017).

4. Ucapan Terimakasih

Tim Peneliti mengucapkan terimakasih yang sebesar-besarnya kepada Kementerian Riset dan Teknologi / Badan Riset dan Inovasi Nasional (KEMENRISTEK/BRIN) yang telah memberikan dukungan dan pembiayaan pada penelitian kami melalui skema Penelitian Dosen Pemula (PDP) tahun 2020 dengan nomor kontrak 052/UTI/LPPM/E.1.3/VII/2020

Daftar Pustaka

- ALITA, D., PRIYANTA, S. AND ROKHMAN, N. (2019) 'Analysis of Emoticon and Sarcasm Effect on Sentiment Analysis of Indonesian Language on Twitter', *Journal of Information Systems Engineering and Business Intelligence*, 5(2), p. 100. doi: 10.20473/jisebi.5.2.100-109.
- HARYALESMANA WAHID, D. (2016) 'Peringkasan Sentimen Esktraktif di Twitter Menggunakan Hybrid TF-IDF dan Cosine Similarity', *Ijccs*, 10(2), pp. 1978–1520.
- ISNAIN, A. R., SIHABUDDIN, A. AND SUYANTO, Y. (2020) 'Bidirectional Long Short Term Memory Method and Word2vec Extraction Approach for Hate Speech Detection', *IJCCS (Indonesian Journal of Computing and Cybernetics Systems)*, 14(2), p. 169. doi: 10.22146/ijccs.51743.
- KLYUEVA, I. (2019) 'Improving Quality of the Multiclass SVM Classification Based on the Feature Engineering', *Proceedings - 2019 1st International Conference on Control Systems, Mathematical Modelling, Automation and Energy Efficiency, SUMMA 2019*. IEEE, pp. 491–494. doi: 10.1109/SUMMA48161.2019.8947599.
- NEGARA, A. B. P., MUHARDI, H. AND PUTRI, I. M. (2020) 'Analisis Sentimen Maskapai Penerbangan Menggunakan Metode Naive Bayes dan Seleksi Fitur Information Gain', *Jurnal Teknologi Informasi dan Ilmu Komputer*, 7(3), p. 599. doi: 10.25126/jtiik.2020711947.
- POZZI, F. A. ET AL. (2016) *Sentiment Analysis in Social Networks, Sentiment Analysis in Social Networks*. Available at: <https://www.scopus.com/inward/record.uri?eid=2-s2.0-85017156686&partnerID=40&md5=db0d99c4de36ad3b0cb6c7c02a3792e1>.
- PRAMUDITA, D. A. AND MUSDHOLIFAH, A. (2020) 'GSA to Obtain SVM Kernel Parameter for Thyroid Nodule Classification', *IJCCS (Indonesian Journal of Computing and Cybernetics Systems)*, 14(1), p. 11. doi: 10.22146/ijccs.41215.
- STYAWATI, S. AND MUSTOFA, K. (2019) 'A Support Vector Machine-Firefly Algorithm for Movie Opinion Data Classification', *IJCCS (Indonesian Journal of Computing and Cybernetics Systems)*, 13(3), p. 219. doi: 10.22146/ijccs.41302.
- YUNITASARI, Y., MUSDHOLIFAH, A. AND SARI, A. K. (2019) 'Sarcasm Detection For Sentiment Analysis in Indonesian Tweets', *IJCCS (Indonesian Journal of Computing and Cybernetics Systems)*, 13(1), p. 53. doi: 10.22146/ijccs.41136.
- ABIDIN, Z. (2017). Penerapan Neural Machine Translation untuk Eksperimen Penerjemahan secara Otomatis pada Bahasa Lampung–Indonesia. In *Prosiding Seminar Nasional Metode Kuantitatif* (No. 1).