

Analisis Sentimen Penghapusan Ujian Nasional pada Twitter Menggunakan Support Vector Machine dan Naïve Bayes berbasis Particle Swarm Optimization

Yuris Alkhalifi¹, Windu Gata², Arfhan Prasetyo³, Imam Budiawan⁴

^{1,2} Ilmu Komputer, STMIK Nusa Mandiri Jakarta

³ Sistem Informasi Akuntansi, Universitas Bina Sarana Informatika Kampus Bogor

⁴ Sistem Informasi, STMIK Nusa Mandiri Jakarta

^{1,2,4} Jl. Kramat Raya No.18, RT.5/RW.7, Kwitang, Kec. Senen, Kota Jakarta Pusat, Daerah Khusus Ibukota Jakarta 10420

³ Jl. Raya Cilebut No.3a, RT.01/RW.04, RT01 RW04, Kec. Tanah Sereal, Kota Bogor, Jawa Barat 16165

14002360@nusamandiri.ac.id¹, windu@nusamandiri.ac.id², arfhan.afp.prasetyo@gmail.com³,

imam121281@gmail.com⁴

Abstrak – Twitter merupakan sebuah platform media sosial yang sering digunakan oleh pengguna di Indonesia. Dalam penggunaannya twitter digunakan sebagai platform yang membahas tentang opini publik, hiburan dan trending topik didunia salah satu perbincangan pada awal tahun 2020 yakni dihapusnya Ujian Nasional (UN) oleh Kementerian Pendidikan dan Kebudayaan Republik Indonesia (Mendikbud RI). Opini dan Sentimen pengguna di twitter pun sangat beragam, ada yang termasuk kedalam sentimen positif dan ada juga sentimen negatif. Untuk memilah mana yang termasuk kedalam sentimen positif dan negatif diperlukan sebuah rangkaian proses, salah satu proses yang dapat digunakan yakni *data mining*. Pada penelitian ini akan membahas tentang sentimen pengguna twitter dalam kasus penghapusan UN dengan menggunakan algoritma *Support Vector Machine* (SVM) dan *Naïve Bayes* (NB) berbasis *Particle Swarm Optimization* (PSO). Dalam 2 algoritma tersebut akan dibagi masing-masing kedalam 2 proses yakni menggunakan PSO dan tidak menggunakan PSO sehingga mempunyai 4 metode. Pengujian dilakukan menggunakan *k-Fold Cross Validation* untuk diperoleh nilai akurasi (*accuracy*), tabel *Confusion Matrix* dan *Area Under Curve*. Hasil pengujian diperoleh nilai akurasi 92,92% dan AUC sebesar 0,977 untuk SVM tanpa PSO. Lalu nilai akurasi 94,81% dan AUC sebesar 0,974 untuk SVM dengan PSO. Nilai akurasi 85,93% dan AUC sebesar 0,645 untuk NB tanpa PSO. Serta nilai akurasi 86,92% dan AUC sebesar 0,715 untuk NB dengan PSO. Dalam penelitian ini metode SVM dengan PSO yang paling baik untuk mengklasifikasikan sentimen positif dan negatif terkait penghapusan UN.

Kata Kunci – *Analisis Sentimen Twitter, Support Vector Machine, Naïve Bayes, Particle Swarm Optimization, Penghapusan UN*

PENDAHULUAN

Perkembangan teknologi saat ini sangat berkembang pesat, salah satunya perkembangan internet. Perkembangan Internet yang begitu cepat telah mengubah banyak aspek dalam proses komunikasi data komputer. Dengan hadirnya internet semakin banyaknya *platform-platform* yang bisa kita jumpai termasuk *platform* media sosial.

Media sosial berasal dari kata media dan sosial. Media merupakan alat atau sarana komunikasi seperti koran, radio, televisi, film, poster dan spanduk[1]. Sedangkan sosial merupakan berkenaan dengan masyarakat, memperhatikan kepentingan umum (suka menolong, menderma dan sebagainya)[2]. Menurut Mandibergh berpendapat bahwa media sosial adalah media yang mawadahi kerjasama diantara pengguna yang menghasilkan konten (*user generated content*)[3].

Salah satu platform media sosial yang populer khususnya di Indonesia yakni Twitter. Twitter merupakan salah satu media sosial yang didirikan oleh Jack Dorsey yang memungkinkan penggunaanya mengirim dan membaca pesan yang dikenal dengan sebutan kicauan (*tweet*)[4]. Dalam penggunaannya twitter rata-rata diakses oleh pengguna di Indonesia mencapai 59% dan menjadikannya menduduki peringkat ke-5 media sosial yang sering digunakan di Indonesia pada tahun 2020[5]. Hal tersebut menandakan bahwa twitter merupakan salah satu media sosial yang cukup berpengaruh untuk para pengguna di Indonesia.

Twitter menjadi platform media sosial yang banyak membahas tentang opini publik, hiburan dan yang paling terkenal adalah tentang isu trending topik yang sedang terjadi didunia. Salah

satu trending topik yang diperbincangkan dalam Twitter pada awal tahun 2020 yakni dihapusnya Ujian Nasional (UN) oleh Kementerian Pendidikan dan Kebudayaan Republik Indonesia dalam masa darurat Virus Covid-19 (Coronavirus) di Indonesia[6]. Ujian Nasional merupakan sistem evaluasi standar yang dilakukan oleh Pusat Penilaian Pendidikan pada tingkat pendidikan dasar dan menengah secara nasional dan persamaan mutu tingkat pendidikan antar daerah[7].

Dalam penghapusan UN tahun 2020 ini pun menuai berbagai opini para pengguna Twitter, ada yang mendukung kebijakan tersebut dan ada yang tidak mendukung serta tidak sedikit pula ada yang menjadikannya sebagai bahan candaan berupa *tweet* dan *meme*. Penghapusan UN tersebut dinilai aneh khususnya bagi para calon lulusan karena UN seolah sudah menjadi tradisi turun temurun untuk menentukan lulus atau tidaknya di sekolah tersebut.

Melihat dari hal tersebut, maka dibutuhkan sebuah proses untuk menentukan mana *tweet* yang merupakan sentimen positif dan mana yang merupakan sentimen negatif. *Data mining* merupakan salah satu proses yang dapat digunakan untuk mengolah data *tweet* sentimen para pengguna twitter khususnya dalam kasus penghapusan UN untuk mengetahui opini publik yang ada di twitter guna sebagai bahan pertimbangan atau evaluasi untuk para pembuat kebijakan.

Penelitian sebelumnya untuk menentukan analisis sentimen twitter dalam operasi tangkap tangan KPK menurut masyarakat menggunakan Algoritma *Support Vector Machine* (SVM) dan Algoritma *Naïve Bayes* (NB) berbasis *Particle Swarm Optimization* (PSO). Hasil dari pembahasannya, di dapatkan akurasi sebesar 80,77% dan AUC sebesar 0,867 untuk Algoritma SVM dan akurasi sebesar 76,92% dan AUC sebesar 0,729 untuk Algoritma NB. Namun setelah menggunakan PSO hasil akurasi yang didapatkan sebesar 83,79% dan AUC sebesar 0,910 untuk Algoritma SVM dan akurasi sebesar 80,13% dan AUC sebesar 0,771 untuk Algoritma NB[8]. Adapun penelitian lainnya dalam analisis sentimen menggunakan perbandingan antara Algoritma NB tanpa *PSO* dan Algoritma NB berbasis *PSO*. Dalam penelitiannya membahas sentimen *e-wallet* pada Google Play yakni produk OVO, dari *review* pengguna pada Google Play akan dikategorikan apakah *review* tersebut termasuk *review* positif atau negatif. Setelah dilakukan pengujian dengan *10-Fold Cross Validation* diperoleh hasil dengan tingkat akurasi 82,30% dan AUC sebesar 0,780 untuk Algoritma *Naïve Bayes* tanpa *Particle Swarm Optimization* dan tingkat akurasi 83.60% dan AUC

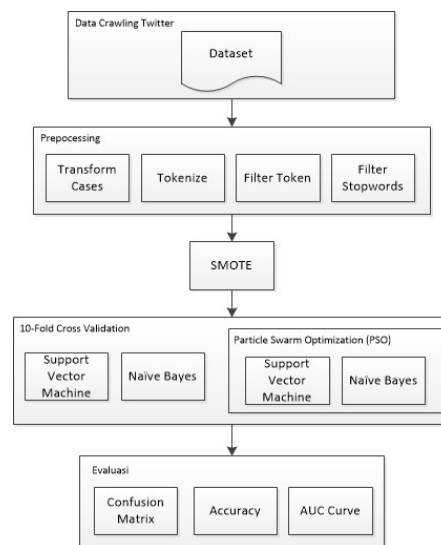
sebesar 0.801 Algoritma NB berbasis PSO[9]. Penelitian berikutnya komparasi metode klasifikasi pada analisis sentimen usaha waralaba berdasarkan data twitter. Algoritma yang digunakan yakni SVM, *Neural Network* (NN), *Decision Tree* dan *K-Nearest Neighbor* (k-NN). Dihilaskan nilai akurasi sebesar 83% untuk algoritma NN, 52% untuk algoritma k-NN, 83% untuk algoritma SVM dan nilai akurasi 81% untuk algoritma *Decision Tree*[10]. Penelitian lainnya tentang analisis sentimen registrasi ulang kartu sim pada twitter menggunakan algoritma SVM dan k-NN. Pada pembahasannya dihasilkan nilai akurasi sebesar 78,87% dan AUC 0,851 untuk algoritma SVM, lalu nilai akurasi 76,55% dan AUC 0,852 untuk algoritma k-NN[11].

Maka dari penelitian sebelumnya Pada penelitian ini akan membahas tahapan yang dilalui untuk melakukan proses analisis sentimen pada media sosial twitter terhadap penghapusan Ujian Nasional. Mulai dari tahap *preprocessing data* sampai tahap analisis sentimen dengan melakukan perbandingan menggunakan Algoritma SVM dan Algoritma NB berbasis PSO serta melakukan pengukuran kualitas hasil analisis antara kedua algoritma tersebut.

METODE PENELITIAN

A. Model Penelitian

Dalam sebuah penelitian, terdapat model metode penelitian. Model Metode Penelitian pada penelitian ini dibuat sebagai alur perjalanan dari penelitian ini yang selanjutnya akan diolah untuk memudahkan dalam melakukan penelitian. Model metode penelitian dapat dilihat pada gambar 1 berikut ini :



Gambar 1. Model Penelitian Ujian Nasional

B. Teknik Pengumpulan Data

Dalam penelitian terdapat 2 teknik pengumpulan data yakni data Primer dan data Sekunder. Data primer merupakan Data yang diperoleh melalui pertanyaan lisan dengan menggunakan metode wawancara, observasi, kuesioner dll[12]. Sedangkan data sekunder adalah data dari sumber data lain atau data yang diperoleh dari obyek penelitian yaitu berupa dokumentasi[12]. Dalam penelitian ini, teknik pengumpulan data yang diambil adalah menggunakan data sekunder yakni data yang diperoleh dari dokumentasi twitter dengan cara pengambilan dari aplikasi RapidMiner 9.6 kedalam format Microsoft Excel yang berekstensi (.csv).

C. Pre-processing Data

Tahap selanjutnya adalah melakukan persiapan data yang sudah didapatkan agar dapat diolah pada saat melakukan pemodelan. Tahap *Preprocessing Data* mencakup kegiatan membangun data dan juga membersihkan data agar siap untuk dikelola ke tahap selanjutnya[9]. Berikut ini tahap *Pre-processing data* antara lain :

Transform Cases

Operator yang digunakan pada tahapan ini adalah untuk mengubah huruf kapital menjadi huruf kecil. Hal ini dilakukan agar ketikan dilakukan proses ke dalam model klasifikasi terdapat keseragaman huruf dan tidak terjadi kesalahan dalam proses tokenize.

Tokenize

Operator Tokenize pada teks adalah memecah sekumpulan kalimat menjadi potongan kata-kata yang nantinya akan dilakukan pembobotan.

Filter Token

Operator Filter Token yang dipakai pada tahapan ini adalah *Filter Token by Length* yakni memfilter dan membatasi karakter dari kata yang diambil dalam penelitian untuk meminimalisir kata yang tidak relevan.

Stopwords

Operator Stopwords yang digunakan dalam tahapan ini adalah *Indonesian Stopword Removal*, yakni menghapus kata-kata yang tidak relevan seperti kata penghubung “tetapi”, “untuk”, “dengan” dan kata lain yang tidak mengandung arti serta tidak ada kaitannya dengan sifat yang berhubungan dengan sentimen.

D. Model Pengujian Data

Setelah dilakukan *preprocessing data*, selanjutnya apabila kelas pada data berbeda jauh

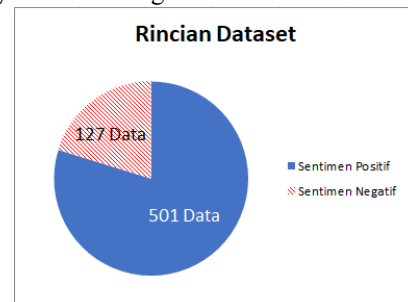
atau (*imbalancing*) data akan dilakukan pengimbangan (*balancing*) terlebih dahulu dengan menggunakan teknik SMOTE. Teknik SMOTE (*Synthetic Minority Over-sampling Technique*) menghasilkan hasil yang baik dan cara yang efektif untuk menangani ketidak seimbangan kelas yang mengalami overfitting pada teknik oversampling untuk memproses kelas minoritas (*positive*)[13].

Kemudian data akan dilakukan proses pemodelan dengan menggunakan Algoritma SVM. Algoritma *Support Vector Machine* (SVM) dan Algoritma *Naïve Bayes* (NB) berbasis *Particle Swarm Optimization* (PSO). Lalu data akan diuji dengan menggunakan *k-Fold Cross Validation* dan akan dievaluasi menggunakan kurva *Area Under Curve* (AUC) dan tabel *Confusion Matrix*.

HASIL DAN PEMBAHASAN

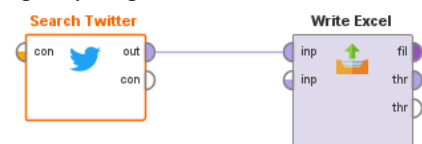
A. Pengambilan Dataset

Proses pengambilan data diambil dari proses *crawling* dari Aplikasi RapidMiner dengan menggunakan operator *search twitter* dengan query “Ujian Nasional”. Kemudian data akan disimpan kedalam format excel 628 data. Adapun rincian datanya adalah sebagai berikut:



Gambar 2. Rincian Dataset

Pengambilan *dataset* ini diambil dan dimasukkan kedalam format excel dengan menggunakan operator *write excel*. Adapun penerapannya seperti Gambar 3 dibawah ini.

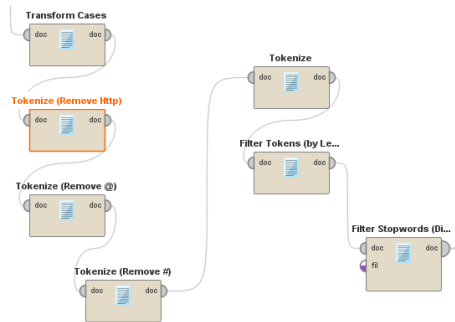


Gambar 3. Proses *Crawling Data* Twitter

B. Proses Pre-processing Data

Setelah data berhasil disimpan kedalam format excel, kemudian dilakukan pembersihan data (*cleansing data*) secara manual yakni menghilangkan kata “RT” (*Retweet*) pada bagian *Text*. Setelah itu dilakukan penghapusan duplikasi

data untuk menghindari data yang berulang. Sampai sini data tidak langsung bisa dimasukkan dalam pengolahan untuk sentimen analisis, maka dilanjutkan dengan tahapan *Data Preparation* seperti Gambar 4 dibawah ini.



Gambar 4. Operator Proses *Preprocessing Data*

Dalam proses *Preprocessing Data* diatas, dapat dilihat ada operator *transform cases*, *tokenize*, *filter token by Length* dan *Filter Stopwords*

Transform Cases

Operator yang digunakan dalam tahapan ini adalah mengubah huruf kapital yang ada pada *text* diubah menjadi huruf kecil semua. Hal ini dilakukan untuk keseragaman huruf dan menghindari eror dalam proses *tokenize*. Penerapannya dapat dilihat pada Tabel 1.

Tabel 1. Penerapan *Transform Cases*

Sebelum	Sesudah
@Reyhan_Ihza46: Ujian Nasional aja ditiadakan, Skripsi ga bisa ditiadakan juga?	@Reyhan_Ihza46: ujian nasional aja ditiadakan, skripsi ga bisa ditiadakan juga?

Tokenize

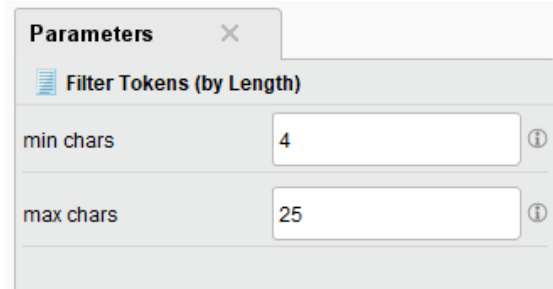
Pada operator ini, terdapat 2 parameter yang akan dipakai yakni mode *regular expression* dan mode *non letters*. Pada mode *regular expression* akan menghilangkan tanda baca, angka, simbol dan karakter khusus atau apapun yang bukan huruf yakni akan menghilangkan karakter “@”, “#” dan tautan yang berupa kata “http/https”. Dan untuk mode *non letters* setiap kalimat akan dipecah menjadi kata perkata untuk diketahui bobotnya. Penerapannya dapat dilihat pada Tabel 2.

Tabel 2. Penerapan *Tokenize*

Sebelum	Sesudah
@Reyhan_Ihza46: ujian nasional aja ditiadakan, skripsi ga bisa ditiadakan juga?	reyhan ihza ujian nasional aja ditiadakan, skripsi ga bisa ditiadakan juga

Filter Token by Length

Operator ini berfungsi untuk menghilangkan sejumlah kata (setelah proses *tokenize*) dengan panjang karakter tertentu. Pada penelitian ini Panjang minimum dan maksimum karakter yang digunakan adalah 4 karakter dan panjang maksimum 25 karakter. Artinya kata yang karakternya kurang dari 4 karakter dan lebih dari 25 karakter maka kata tersebut akan dihilangkan. Penerapannya dapat dilihat pada Gambar 5.



Gambar 5. Penerapan *Filter Token By Length*

Filter Stopwords

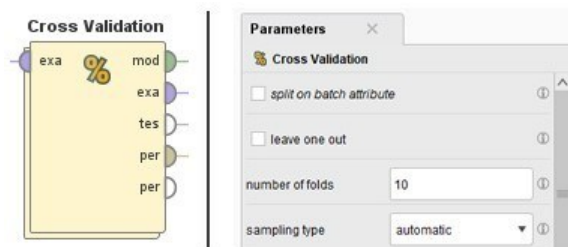
Terakhir adalah operator *Filter Stopwords*. Operator ini berfungsi untuk menghilangkan kata-kata dalam Bahasa Indonesia yang tidak relevan, tidak memiliki makna dan tidak ada hubungan dengan isi *text* seperti kata penghubung “dan”, “tetapi”, “akan” dan lain-lain dengan file *Stopword* Bahasa Indonesia yang sudah disiapkan. Penerapannya dapat dilihat pada Tabel 3.

Tabel 3. Penerapan *Filter Stopwords*

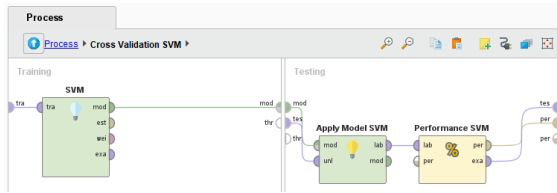
Sebelum	Sesudah
reyhan ihza ujian nasional aja ditiadakan, skripsi ga bisa ditiadakan juga	reyhan ihza ujian nasional ditiadakan, skripsi ditiadakan

C. Proses Pengujian Algoritma

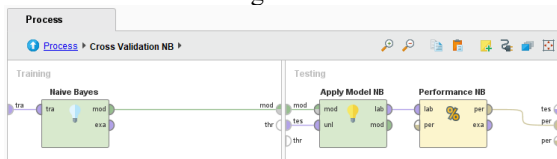
Pada proses ini, 4 metode pada algoritma yang digunakan yakni SVM, SVM + PSO, NB dan NB + PSO akan dilakukan pengujian dengan menggunakan *k-Fold Cross Validation* dengan nilai *k=10*. Operator yang digunakan dapat dilihat pada Gambar 6.



Gambar 6. Operator *K-Fold Cross Validation*
 Adapun didalam operator *Cross Validation* tersebut terdapat beberapa operator dan proses diantaranya dapat dilihat pada Gambar 7 dan Gambar 8.



Gambar 7. Proses *K-Fold Cross Validation* Algoritma SVM



Gambar 8. Proses *K-Fold Cross Validation* Algoritma NB

D. Hasil Pengujian dan Evaluasi

Setelah data di proses, Langkah selanjutnya adalah melakukan evaluasi yakni menggunakan Tabel *Confusion Matrix*, *Accuracy* dan Kurva AUC (*Area Under Curve*).

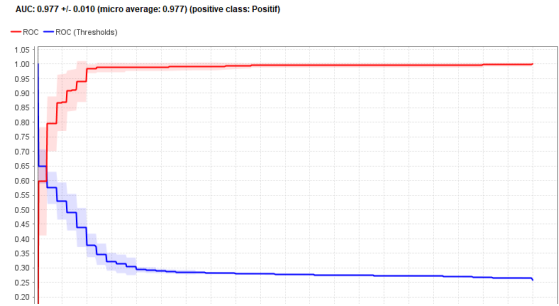
Algoritma SVM Tanpa PSO

Dari tahapan-tahapan yang dilakukan Tabel *Confusion Matrix* pada Algoritma SVM tanpa PSO dapat dilihat pada Tabel 4.

Tabel 4. Hasil Evaluasi Algoritma SVM tanpa PSO

<i>Accuracy</i> : : 92,91% +/- 1,71%			
	<i>True Positif</i>	<i>True Negatif</i>	<i>Class Precision</i>
<i>Pred. Positif</i>	458	28	94,24%
<i>Pred. Negatif</i>	43	473	91,67%
<i>Class Recall</i>	91,42%	94,41%	

Diketahui bahwa hasil evaluasi dari Algoritma SVM tanpa PSO diatas, jumlah *True Positive* (TP) sebesar 458 *record* dan *False Negative* (FN) sebanyak 43 *record* diklasifikasikan sebagai SENTIMEN POSITIF namun seharusnya SENTIMEN NEGATIF. Berikutnya jumlah *True Negative* (TN) sebesar 473 *record* diklasifikasikan sebagai SENTIMEN NEGATIF, dan 28 *record False Positive* (FP) diklasifikasikan sebagai SENTIMEN NEGATIF namun seharusnya SENTIMEN POSITIF. Dan mendapatkan tingkat akurasi sebesar 92,91% dengan toleransi kesalahan sebesar 1,71%. Lalu untuk nilai AUC nya sebesar 0,980 yang dapat dilihat pada Gambar 9.



Gambar 9. Grafik AUC Algoritma SVM tanpa PSO

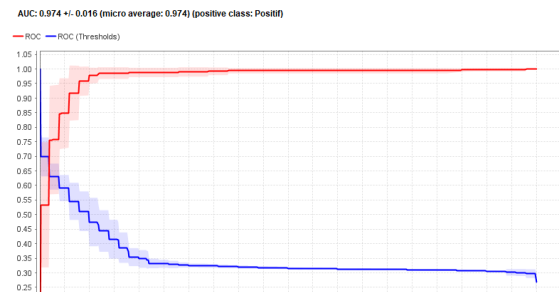
Algoritma SVM Dengan PSO

Dari tahapan-tahapan yang dilakukan Tabel *Confusion Matrix* pada Algoritma SVM dengan PSO dapat dilihat pada Tabel 5.

Tabel 5. Hasil Evaluasi Algoritma SVM dengan PSO

<i>Accuracy</i> : : 94,81% +/- 1,82%			
	<i>True Positif</i>	<i>True Negatif</i>	<i>Class Precision</i>
<i>Pred. Positif</i>	488	9	97,26%
<i>Pred. Negatif</i>	3	462	92,60%
<i>Class Recall</i>	92,22%	97,41%	

Diketahui bahwa hasil evaluasi dari Algoritma SVM dengan PSO diatas, jumlah *True Positive* (TP) sebesar 488 *record* dan *False Negative* (FN) sebanyak 13 *record* diklasifikasikan sebagai SENTIMEN POSITIF namun seharusnya SENTIMEN NEGATIF. Berikutnya jumlah *True Negative* (TN) sebesar 462 *record* diklasifikasikan sebagai SENTIMEN NEGATIF, dan 39 *record False Positive* (FP) diklasifikasikan sebagai SENTIMEN NEGATIF namun seharusnya SENTIMEN POSITIF. Dan mendapatkan tingkat akurasi sebesar 94,81% dengan toleransi kesalahan sebesar 1,82%. Lalu untuk nilai AUC nya sebesar 0,974 yang dapat dilihat pada Gambar 10.



Gambar 10. Grafik AUC Algoritma SVM dengan PSO

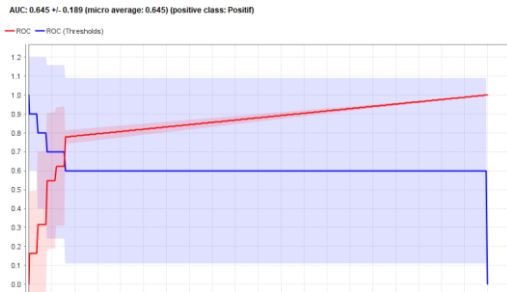
Algoritma NB Tanpa PSO

Dari tahapan-tahapan yang dilakukan Tabel *Confusion Matrix* pada Algoritma NB tanpa PSO dapat dilihat pada Tabel 6.

Tabel 6. Hasil Evaluasi Algoritma NB tanpa PSO

Accuracy: : 85,93% +/- 3,10%			
	True Positif	True Negatif	Class Precision
Pred. Positif	379	19	79,80%
Pred. Negatif	122	482	95,23%
Class Recall	96,21%	95,65%	

Diketahui bahwa hasil evaluasi dari Algoritma NB tanpa PSO diatas, jumlah True Positive (TP) sebesar 379 record dan False Negative (FN) sebanyak 19 record diklasifikasikan sebagai SENTIMEN POSITIF namun seharusnya SENTIMEN NEGATIF. Berikutnya jumlah True Negative (TN) sebesar 482 record diklasifikasikan sebagai SENTIMEN NEGATIF, dan 122 record False Positive (FP) diklasifikasikan sebagai SENTIMEN NEGATIF namun seharusnya SENTIMEN POSITIF. Dan mendapatkan tingkat akurasi sebesar 85,93% dengan toleransi kesalahan sebesar 3,10%. Lalu untuk nilai AUC nya sebesar 0,645 yang dapat dilihat pada Gambar 11.



Gambar 11. Grafik AUC Algoritma NB tanpa PSO

Algoritma NB dengan PSO

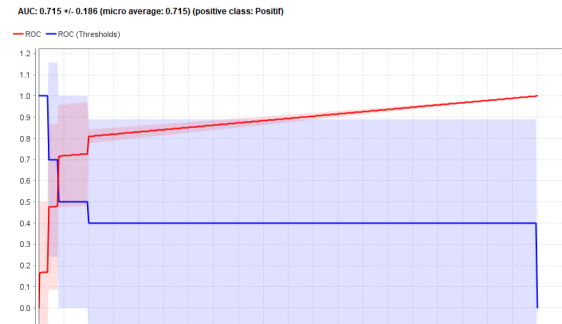
Dari tahapan-tahapan yang dilakukan Tabel *Confusion Matrix* pada Algoritma NB dengan PSO dapat dilihat pada Tabel 7.

Tabel 7. Hasil Evaluasi Algoritma NB dengan PSO

Accuracy: : 86,82% +/- 3,33%			
	True Positif	True Negatif	Class Precision
Pred. Positif	384	15	80,60%
Pred. Negatif	117	486	96,24%
Class Recall	97,01%	76,65%	

Diketahui bahwa hasil evaluasi dari Algoritma NB dengan PSO diatas, jumlah True Positive (TP) sebesar 384 record dan False

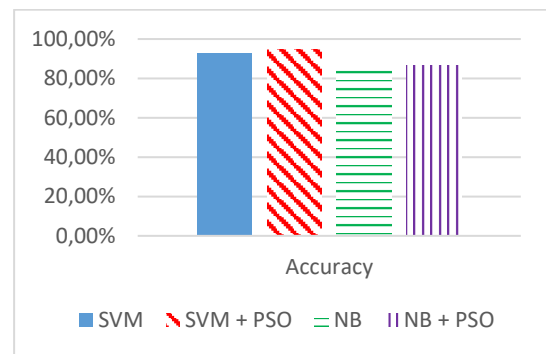
Negative (FN) sebanyak 15 record diklasifikasikan sebagai SENTIMEN POSITIF namun seharusnya SENTIMEN NEGATIF. Berikutnya jumlah True Negative (TN) sebesar 486 record diklasifikasikan sebagai SENTIMEN NEGATIF, dan 117 record False Positive (FP) diklasifikasikan sebagai SENTIMEN NEGATIF namun seharusnya SENTIMEN POSITIF. Dan mendapatkan tingkat akurasi sebesar 85,93% dengan toleransi kesalahan sebesar 3,10%. Lalu untuk nilai AUC nya sebesar 0,715 yang dapat dilihat pada Gambar 12.



Gambar 12. Grafik AUC Algoritma NB dengan PSO

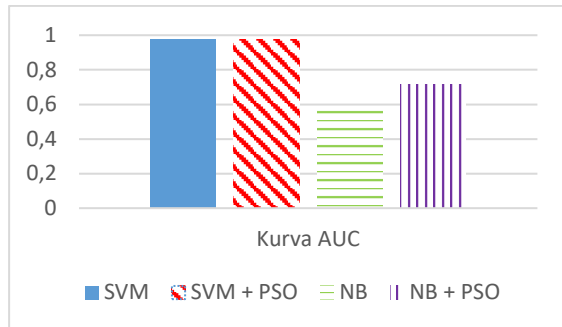
E. Perbandingan Hasil Pengujian

Dari hasil evaluasi 4 pengujian diatas didapatkan selisih Accuracy dan kurva AUC. Adapun perbedaan hasil evaluasi tingkat akurasi (Accuracy) dapat dilihat pada grafik di Gambar 13 berikut ini:



Gambar 13. Grafik Perbandingan Accuracy

Dapat dilihat pada Gambar 13 bahwa metode SVM tanpa PSO mendapatkan accuracy sebesar 92,92%, SVM dengan PSO sebesar 94,81%, NB tanpa PSO sebesar 85,9%, NB dengan PSO sebesar 86,92%. Adapun grafik selisih Kurva AUC dapat dilihat pada Gambar 14.



Gambar 14. Grafik Perbandingan Kurva AUC

Dapat dilihat pada Gambar 14 bahwa metode SVM tanpa PSO mendapatkan kurva AUC sebesar 0,977, SVM dengan PSO sebesar 0,974, NB tanpa PSO sebesar 0,645 dan NB dengan PSO sebesar 0,715.

KESIMPULAN DAN SARAN

Kesimpulan yang didapatkan setelah melakukan penelitian dengan 4 metode dengan menggunakan model Algoritma *Support Vector Machine* (SVM) dan *Naive Bayes* (NB) berbasis *Particle Swarm Optimization* (PSO) yakni SVM tanpa PSO, SVM dengan PSO, NB tanpa PSO dan NB dengan PSO. Maka pada penelitian ini dapat disimpulkan bahwa Algoritma SVM dengan PSO mendapatkan *accuracy* terbesar dibanding 3 metode lain. Namun untuk kurva AUC, metode yang paling besar yakni metode SVM tanpa PSO.

Adapun saran untuk perbaikan dan peningkatan penelitian selanjutnya adalah mencoba dengan algoritma lain seperti *k-NN*, *Decision Tree* dan lain-lain, serta menggunakan *dataset* yang lebih banyak guna mengoptimalkan hasil penelitian.

UCAPAN TERIMA KASIH

Terimakasih banyak diucapkan kepada Dosen Pengampu Mata Kuliah *Advanced Information Retrieval* yakni Bapak Dr. Windu Gata, M.Kom atas bimbingannya dalam mata kuliah tersebut, juga kepada Bapak Arfhan Prasetyo, M.Kom & Bapak Imam Budiawan, M.Kom serta rekan-rekan kelas 14.2A.01 STMIK Nusa Mandiri Tahun 2020 sehingga jurnal ini dapat disusun dan dipublikasikan.

REFERENSI

- [1] Departemen Pendidikan Nasional, "Kamus Besar Bahasa Indonesia," Jakarta, 2005.
- [2] E. Setiawan, "Kamus Besar Bahasa Indonesia Daring (Online)." <https://kbbi.web.id/sosial> (accessed Apr. 17, 2020).

- [3] I. A. Ratnamulyani and B. I. Maksudi, "Peran Media Sosial Dalam Peningkatan Partisipasi Pemilih Pemula Dikalangan Pelajar Di Kabupaten Bogor," *Sosiohumaniora*, vol. 20, no. 2, pp. 154–161, 2018, doi: 10.24198/sosiohumaniora.v20i2.13965.
- [4] Wikipedia, "Twitter." <https://id.wikipedia.org/wiki/Twitter> (accessed Apr. 17, 2020).
- [5] Katadata Indonesia, "10 Media Sosial yang Paling Sering Digunakan di Indonesia," 2020. <https://databoks.katadata.co.id/datapublish/2020/02/26/10-media-sosial-yang-paling-sering-digunakan-di-indonesia> (accessed Apr. 17, 2020).
- [6] Kementerian Pendidikan dan Kebudayaan Republik Indonesia, *Surat Edaran Nomor 4 Tahun 2020 Tentang Pelaksanaan Kebijakan Pendidikan Dalam Masa Darurat Penyebaran Coronavirus Disease (Covid-19)*. Indonesia, 2020.
- [7] D. A. Purnamasari, "Strategi Meningkatkan Hasil Ujian Nasional Mata Pelajaran Ekonomi," *J. Ilm. Pendidik. Ekon. IKIP Veteran Semarang*, vol. 1, no. 2, pp. 38–48, 2013, [Online]. Available: <https://www.neliti.com/publications/37069/strategi-meningkatkan-hasil-ujian-nasional-mata-pelajaran-ekonomi>.
- [8] Hernawati and W. Gata, "Sentimen Analisis Operasi Tangkap Tangan KPK Menurut Masyarakat Menggunakan Algoritma Support Vector Machine, Naive Bayes Berbasis Particle Swarm Optimization," *Fakt. Exacta*, vol. 12, no. 3, pp. 230–243, 2019, doi: 10.30998/faktorexacta.v12i3.4992.
- [9] S. A. Saputra, D. Rosiyadi, W. Gata, and S. M. Husain, "Sentiment Analysis Analisis Sentimen E-Wallet Pada Google Play Menggunakan Algoritma Naive Bayes Berbasis Particle Swarm Optimization," *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 3, no. 3, pp. 377–382, 2019, doi: 10.29207/resti.v3i3.1118.
- [10] T. Mardiana, H. Syahreva, and T. Tuslaela, "Komparasi Metode Klasifikasi Pada Analisis Sentimen Usaha Waralaba Berdasarkan Data Twitter," *J. Pilar Nusa Mandiri*, vol. 15, no. 2, pp. 267–274, 2019, doi: 10.33480/pilar.v15i2.752.
- [11] M. Kusmira, "Analisis Sentimen Registrasi

- Ulang Kartu SIM pada Twitter Menggunakan Algoritma SVM dan K-NN,” *Inti Nusa Mandiri*, vol. 14, no. 1, pp. 105–110, 2019, [Online]. Available: <http://ejournal.nusamandiri.ac.id/index.php/inti/article/view/541/551>.
- [12] E. Ermawati, “Algoritma Klasifikasi C4.5 Berbasis Particle Swarm Optimization Untuk Prediksi Penerima Bantuan Pangan Non Tunai,” *J. Sist. Inf.*, vol. 8, no. September, pp. 513–528, 2019.
- [13] S. A. Putri, “Prediksi Cacat Software Dengan Teknik Sampel Dan Seleksi Fitur Pada Bayesian Network,” *J. Kaji. Ilm.*, vol. 19, no. 1, p. 17, 2019, doi: 10.31599/jki.v19i1.314.