

Analisis Sentimen *Omnibus Law* Pada Twitter Dengan Algoritma Klasifikasi Berbasis *Particle Swarm Optimization*

Annisa Elfina Augustia¹, Resi Taufan², Yuris Alkhalifi³, Windu Gata⁴

¹STMIK Nusa Mandiri
e-mail : 14002344@nusamandiri.ac.id

²STMIK Nusa Mandiri
e-mail : resitauf1508@nusamandiri.ac.id

³STMIK Nusa Mandiri
e-mail : 14002360@nusamandiri.ac.id

⁴STMIK Nusa Mandiri
e-mail : windu@nusamandiri.ac.id

Abstract - Twitter social media is a means to develop expression and appreciation of something online as a sign of technological progress and the ease of using the internet. The Omnibus Law, a work copyright law made by the Indonesian government, aims to increase investment so that it creates many new jobs. The Omnibus Law has become a hot topic on Twitter social media, responses from various circles of society both support and reject the existence of the law. The purpose of this study is to separate positive and negative sentiments and to measure opinions on the Omnibus Law with a classification algorithm. In this study, positive and negative sentiments were collected in the crawling process. The classification algorithm used is Support Vector Machine (SVM) and Naive Bayes (NB), optimization with Particle Swarm Optimization (PSO) is also used to improve accuracy. The results of the k-fold cross validation test with SVM and NB obtained an accuracy of 84.95% and 87.53% with an Area Under the Curve (AUC) value of 0.958 and 0.754, while the results of the test using an accuracy value of 86.53% and 90.12% with AUC values of 0.948 and 0.816.

Keywords: Omnibus Law, Sentiment Analysis, SVM, Naive Bayes, Particle Swarm Optimization.

Abstrak - Media sosial twitter menjadi sarana untuk mencurahkan ekspresi dan apresiasi terhadap sesuatu secara online sebagai tanda dari kemajuan teknologi dan kemudahan penggunaan internet. *Omnibus Law* undang-undang cipta kerja yang dibuat oleh pemerintah Indonesia bertujuan untuk meningkatkan investasi asing sehingga banyak menciptakan lapangan pekerjaan baru. *Omnibus Law* telah menjadi topik panas pada media sosial twitter, beragam tanggapan muncul dari berbagai kalangan masyarakat baik mendukung maupun menolak adanya undang-undang tersebut. Tujuan Penelitian ini agar sentimen positif dan sentimen negatif dapat dipisahkan dan untuk mengukur pendapat terhadap *Omnibus Law* dengan algoritma klasifikasi. Pada penelitian ini sentimen positif dan sentimen negatif dikumpulkan dalam proses *crawling*. Algoritma klasifikasi yang digunakan yaitu *Support Vector Machine* (SVM) dan *Naive Bayes* (NB), optimasi dengan *Particle Swarm Optimization* (PSO) juga digunakan untuk meningkatkan akurasi. Hasil pengujian *k-fold cross validation* dengan SVM dan NB mendapatkan akurasi 84,95% dan 87,53% dengan nilai *Area Under the Curve* (AUC) 0.958 dan 0.754, sedangkan hasil pengujian menggunakan masing-masing mendapatkan nilai akurasi 86,53% dan 90,12% dengan nilai AUC 0.948 dan 0.816.

Kata Kunci: *Omnibus Law*, Analisis Sentimen, SVM, *Naive Bayes*, *Particle Swarm Optimization*.

PENDAHULUAN

Omnibus Law adalah undang-undang yang menitikberatkan pada penyederhanaan jumlah regulasi karena sifatnya yang merevisi dan mencabut banyak undang-undang sekaligus. Penggunaan konsep *Omnibus Law* belum terakomodir di dalam Undang-Undang No. 12 Tahun 2011, namun penggunaan konsep ini bukanlah hal yang terlarang

(Putra, 2020). Rancangan undang-undang *Omnibus Law* menyebabkan konflik antara pemerintah dan kelompok gerakan masyarakat sipil karena kurangnya komunikasi politik, adanya politik *relation* dan minimnya partisipasi politik (Razy & Fedryansyah, 2020). Meski demikian menurut (Fitryantica, 2019) harmonisasi hukum perundang-undangan dari *Omnibus Law* sangat penting dalam perkembangan hukum di Indonesia. Undang-undang konsep



Omnibus Law bisa mengarah sebagai Undang-Undang payung karena mengatur secara menyeluruh dan kemudian mempunyai kekuatan terhadap aturan yang lain. Harmonisasi *Omnibus Law* menjadi pelindung bagi pejabat daerah yang ingin melakukan inovasi dan kreasi untuk kemajuan ekonomi dan investasi.

Masyarakat Indonesia banyak yang menanggapi rancangan undang-undang tersebut melalui *twitter*. *Twitter* sebagai salah satu media sosial yang populer, digunakan masyarakat untuk mengungkapkan opini – opini (Zamachsari et al., 2020). Opini dan Sentimen pengguna di *twitter* dapat digolongkan menjadi sentiment positif dan negatif dengan suatu proses dan algoritma tertentu atau disebut data mining (Alkhalifi et al., 2020). Menurut (Cahyaningrum et al., 2020) *Twitter* merupakan media sosial yang menjadi wadah penggunaannya untuk saling berbagi respons atas suatu peristiwa melalui cuitan, pada penelitian nya cuitan tersebut diklasifikasikan menjadi sentimen positif, negatif dan netral. Penelitian tentang *text mining* yang bersumber dari komentar yang ditulis pada media social *twitter* dan *facebook* (Mihuandayani et al., 2018) pada penelitian tersebut performa diukur menggunakan parameter *precision*, *recall*, dan *F-measure*. Pada penelitian ini, *twitter* akan menjadi sumber data yang akan diproses untuk analisis sentimen. Cuitan dan komentar pada *twitter* diambil dengan proses *crawling* dengan menggunakan *software RapidMiner* kemudian dilakukan analisis sentimen dengan menggunakan algoritma klasifikasi.

Algoritma SVM pada analisis sentimen *twitter* untuk teks berbahasa Indonesia dengan fitur *unigram* dan fitur pembobotan *Term Frequency dan Invers Document Frequency* (TF-IDF) berhasil mendapatkan akurasi 86,81% (Putranti & Winarko, 2014). Sedangkan pada penelitian lain tentang analisis sentimen menggunakan algoritma SVM dan SVM+PSO pada *twitter* telah berhasil dilakukan dan mendapat akurasi sebesar 95,46% dan AUC 0,979 sedangkan pada SVM+PSO sebesar 96,04% dan AUC 0,993 (Que et al., 2020).

Algoritma klasifikasi lain yang banyak digunakan dalam analisis sentimen adalah NB, algoritma ini menghasilkan akurasi yang baik seperti pada penelitian (Gunawan et al., 2017) tentang ulasan aplikasi *mobile banking* user yang dilakukan dengan akurasi sebesar 94,4% dan penelitian tentang layanan *customer care* yang dilakukan (Saputra et al., 2020) dengan akurasi sebesar 78,88%. Optimasi PSO pada algoritma NB pada sentimen analisis terbukti dapat meningkatkan nilai akurasi sebesar 4,12% dari nilai akurasi tanpa PSO sebesar 86,62% menjadi 90,74% (Hayatin et al., 2020).

Tanggapan dari media sosial *twitter* sebagai bentuk reaksi berupa cuitan dan komentar tertulis baik bersifat penolakan (sentimen negatif) maupun dukungan (sentimen positif) pada rancangan undang-undang omnibus law tersebut menjadi permasalahan yang menarik untuk dilakukan penelitian tentang

sentimen analisis menggunakan algoritma klasifikasi, yaitu dengan algoritma SVM dan NB. Tujuan dari penelitian ini untuk melakukan analisis sentimen masyarakat Indonesia yang diwakili oleh opini pengguna *twitter* terhadap undang-undang *Omnibus Law*, selain itu penelitian ini juga untuk menguji kembali apakah penggunaan optimasi PSO pada algoritma SVM dan NB dapat meningkatkan nilai akurasi dibanding tanpa menggunakan PSO.

1. Sentimen Analisis

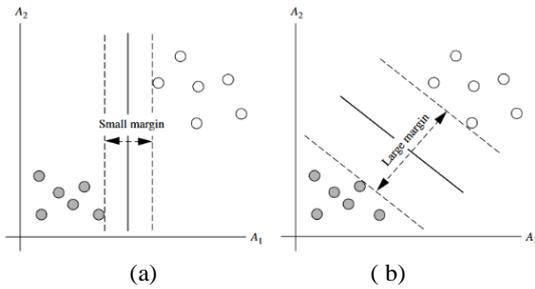
Menurut (Medhat et al., 2014) analisa sentimen atau *opinion mining* adalah studi komputasi mengenai pendapat, perilaku dan emosi seseorang terhadap entitas. Entitas tersebut dapat menggambarkan individu, kejadian atau topik. *Opinion mining* tidak memperhatikan topik dari teks tersebut tetapi lebih fokus kepada ekspresi yang digambarkan dari teks opini tersebut. Analisa sentimen juga dapat diartikan teknik komputasi pendapat, perasaan dan subjektivitas teks.

2. Pre-processing

Pre-processing adalah proses pembersihan dan penyiapan teks untuk klasifikasi karena pada teks online biasanya berisi banyak gangguan dan bagian yang tidak informatif seperti tag HTML, skrip, dan iklan. Selain itu, pada level kata, banyak kata-kata dalam teks tidak berdampak pada orientasi umum itu (Haddi et al., 2013). *Pre-processing* menjadi langkah penting pencarian informasi. Pencarian Informasi digunakan untuk memutuskan dokumen mana dalam suatu kumpulan data yang harus diambil untuk memenuhi pengguna kebutuhan informasi. *Pre-processing* dilakukan dalam 4 langkah-langkah dalam yaitu *case folding*, *tokenizing*, *Filtering*, *Stemming* (Hermawan & Bellanar Ismiati, 2020).

3. Support Vector Machine

SVM merupakan salah satu metode yang sudah banyak diterapkan untuk berbagai jenis penelitian di bidang data dan *text mining* karena telah mampu menunjukkan performa yang lebih baik (Styawati & Mustofa, 2019). Klasifikasi pada SVM dilakukan dengan garis pembatas (*hyperlane*) yang memisahkan antara kelas opini positif dan opini negatif. Secara intuitif, suatu garis pembatas yang baik adalah yang memiliki jarak terbesar ke titik data pelatihan terdekat dari setiap kelas, karena pada umumnya semakin besar margin, semakin rendah *error* generalisasi dari pemilah. Margin adalah jarak dari suatu titik vektor di suatu kelas terhadap *hyperplane* (Santoso et al., 2017) Ilustrasi *hyperplane* pada SVM dapat dilihat pada gambar 1 (Han et al., 2012)



Gambar 1. Ilustrasi *hyperplane* pada SVM

Pada gambar 1 bagian (a) memiliki fungsi pemisahan yang memisahkan Kelas 1 dan Kelas 2, namun tidak optimal. Sedangkan gambar 1 (b) memiliki fungsi pemisahan yang memisahkan Kelas 1 dan Kelas 2 secara efektif. Pada SVM, yang paling optimal fungsi pemisahan dicari seperti yang diilustrasikan pada gambar 1 (b).

4. Naive Bayes

Metode NB merupakan salah satu metode yang dapat mengklasifikasikan teks dengan algoritma sederhana tetapi memiliki akurasi yang tinggi (Ariadi & Fithriasari, 2015). NB merupakan salah satu implementasi sederhana dari teorema Bayes. Adapun rumus NB ditulis dalam persamaan (1) berikut.

$$P(B|A) = \frac{P(A|B)P(B)}{P(A)} \dots\dots\dots(1)$$

Keterangan :

- P(B|A) = Probabilitas kemunculan B ketika A diketahui
- P(A|B) = Probabilitas kemunculan A jika B diketahui.
- P(A) = Probabilitas kemunculan A.
- P(B) = probabilitas kemunculan B.

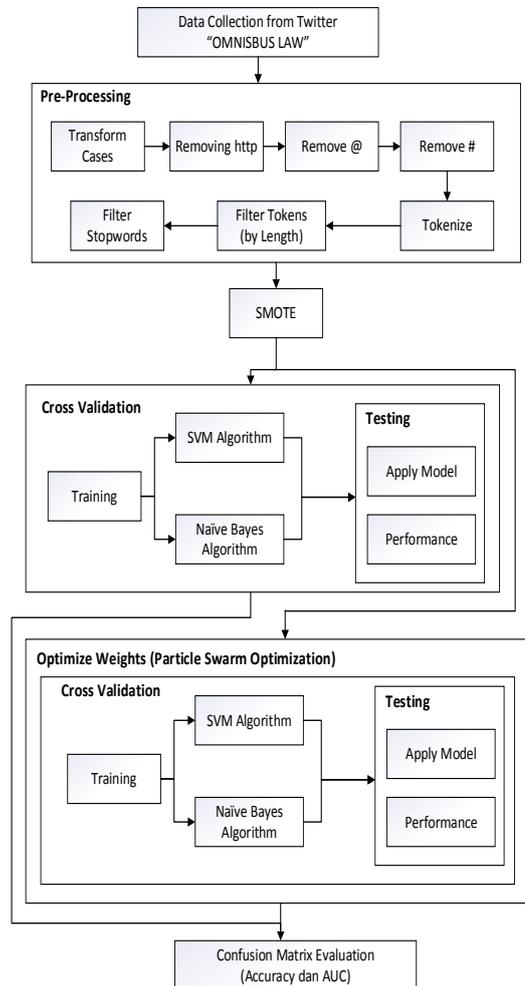
4. Particle Swarm Optimization

PSO merupakan algoritma berbasis populasi yang mengeksplorasi individu dalam pencarian, pada PSO populasi disebut *swarm* dan individu disebut *particle* (Kusmarna et al., 2015). PSO yaitu pencarian solusi optimal secara global dalam ruang pencarian melalui interaksi individu dalam segerombolan partikel dengan cara melakukan seleksi terhadap atribut yang ada (Achyani, 2018). Setiap partikel bergerak dalam ruang tertentu kemudian mencari posisi terbaik yang dilaluinya dengan memperbaharui posisinya, setiap partikel menyampaikan informasi posisi terbaik ke partikel lain dan menyesuaikan posisi dan kecepatan berdasarkan informasi posisi terbaik yang diterima (Istighfarin et al., 2020).

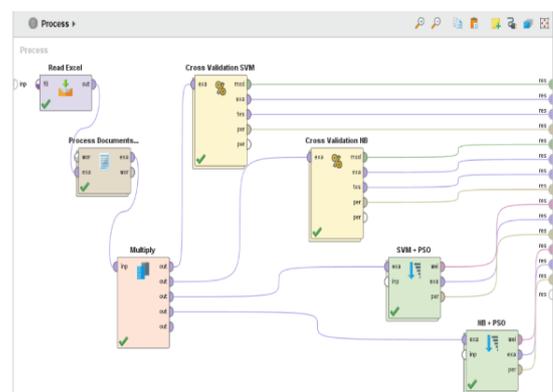
METODOLOGI PENELITIAN

Pada penelitian ini langkah-langkah direpresentasikan dalam sebuah model kerangka kerja penelitian pada gambar 2. Berdasarkan kerangka kerja penelitian tersebut, dilakukan

pembuatan model analisis data. Tool yang digunakan untuk pemodelan ini adalah *RapidMiner Studio* versi 9.2. model analisis data yang dibuat sebagaimana disajikan pada gambar 3.



Gambar 2. Kerangka Kerja Analisis Sentimen



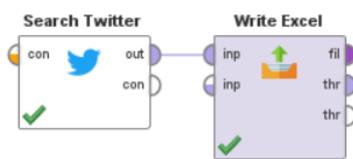
Gambar 3. Desain Model Komparasi SVM dan NB berbasis PSO

1. Teknik Pengumpulan Data

Pada penelitian ini data yang didapatkan dari sumber yang sudah ada (data sekunder). Pengumpulan data

twitter dilakukan dengan *crawling tool RapidMiner* menggunakan operator *search twitter* dengan memasukkan parameter *query* “*Omnibus Law*”, dengan jumlah *record* sebanyak 2.000 dan dengan tipe terkini atau terpopuler.

Proses *crawling* data dari *twitter* disajikan pada gambar 3. Data yang didapatkan pada proses *crawling* sebanyak 2.000 baris dan disimpan ke dalam bentuk *file excell*. Setelah itu data pada *file excell* tersebut dihilangkan data duplikat nya dengan dilakukan penghapusan “RT” (*Retweet*) pada kolom text secara manual sebelum dilakukan *pre-processing*. Penghapusan data duplikat sebanyak 668 baris sehingga tersisa 1.332 baris yang akan dilakukan *pre-processing*.



Gambar 4. Proses Crawling data dari *Twitter*

2. Pre-Processing

Data sekunder yang didapat pada proses *crawling* biasanya tidak langsung siap digunakan untuk proses pengujian, maka dari itu perlu dilakukan *pre-processing*. Tahap ini dilakukan untuk mengubah data yang diperoleh dari sumber data menjadi dataset agar siap diproses lebih lanjut ke tahap pengujian data. *Pre-processing* bertujuan untuk menghilangkan noise serta melakukan beberapa tahapan untuk mengubah data yang belum terstruktur menjadi terstruktur atau mengubah teks menjadi term indeks yang mewakili sebuah dokumen sehingga siap diproses lebih lanjut. Tahapan - tahapan *pre-processing* disajikan pada gambar 5.



Gambar 5. Tahapan-tahapan dari *Pre-Processing*

Beberapa tahapan yang dilakukan dalam *pre-processing* sesuai dengan gambar tersebut yaitu:

Transform cases dilakukan untuk menyeragamkan semua teks menjadi huruf kecil. Hal ini bertujuan agar tidak terjadi masalah ketika dilakukan proses *tokenize*. *Remove http* yaitu proses menghilangkan URL yang biasanya terdapat pada *text tweet*. Pada tabel 1 menyajikan *text* sebelum dan sesudah diproses *Remove http*.

Remove Annotation bertujuan untuk menghapus tanda *annotation* yang terdapat pada *text tweet*. *Annotation* merupakan salah satu noise yang tidak memiliki arti. Contoh teks yang telah dihapus tanda *annotasinya*

pada proses *Remove Annotation* disajikan pada Tabel 1. *Remove Hashtag* dilakukan untuk menghapus kata kunci *hashtag*. Contoh teks sebelum dan sesudah dilakukan proses penghapusan *hashtag* disajikan pada tabel 1.

Proses *tokenize* merupakan proses pemotongan *string input* berdasarkan tiap kata penyusunnya. Tahapan ini dilakukan untuk memisahkan kata per kata dari suatu teks kalimat. Tahap ini bertujuan agar mendapat kata-kata yang panjangnya antara 4 sampai dengan 25 karakter. Pada tahapan ini dilakukan pemilihan kata-kata penting dari hasil token, yaitu kata-kata apa saja yang akan digunakan untuk mewakili dokumen, sehingga pada proses ini akan menghapus kata-kata yang tidak memiliki makna sesuai dengan pendekatan *bag-of-words*.

Tabel 1. Contoh *Tweet* Sebelum dan Sesudah *Pre-processing*

No	Sebelum	Sesudah
1	@Marulitua_: Omnibus Law Dijadikan Solusi Perbaiki Ekonomi Pasca Pandemi	Omnibus Law dijadikan solusi perbaiki ekonomi pasca pandemi
2	"Pada tingkat implementasinya nanti, Omnibus Law Cipta Kerja membawa kesejahteraan pada buruh & karyawan sesuai dengan kinerja serta sinergi yg baik antara karyawan & perusahaan.	pada tingkat implementasinya nanti Omnibus Law cipta kerja membawa kesejahteraan pada buruh karyawan sesuai dengan kinerja serta sinergi yg baik antara karyawan perusahaan
3	"Gerindra Dukung Omnibus Law, Prabowo Yakin Jokowi Pikirkan Rakyat Kalahkan Pandemi #RakyatBersamaOmnibusLaw https://t.co/FSWpb1bUOY https://t.co/HytA7PpJr7"	gerindra dukung Omnibus Law prabowo yakin jokowi pikirkan rakyat kalahkan pandemi

Sumber: *Twitter.com* (2020)

2. Model Pengujian Data

Setelah dilakukan *pre-processing* tahap berikutnya dilakukan *Synthetic Minority Oversampling Technique* (SMOTE) agar tidak terjadi *Imbalance* data atau suatu kondisi dimana antar kelas memiliki selisih yang signifikan, SMOTE perlu dilakukan untuk menghindari *Imbalance* karena pemodelan dengan algoritma pemodelan dengan algoritma yang tidak memperhatikan ketidakseimbangan data akan didominasi oleh kelas mayor dan tidak memperhatikan kelas minornya. Hal ini akan menyebabkan hasil penelitian pada batas-batas keputusan yang diragukan dalam klasifikasi (Heranova, 2019). Tahap berikutnya dilakukan pemodelan dengan algoritma SVM, NB, SVM+PSO

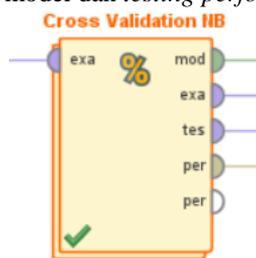
dan NB+PSO. Pada penelitian ini digunakan kernel *linear splines in one-dimension* (dot). Selanjutnya dilakukan pengujian dengan *k-fold cross validation* untuk mengevaluasi model atau algoritma yang bertujuan memisahkan data menjadi data *training* dan data validasi. Model akan dilatih oleh data training dan divalidasi oleh data validasi.

Pada penelitian ini, hasil dari tahap pengujian akan dievaluasi menggunakan table *Confusion Matrix* yaitu *accuracy* dan *Area Under the Curve* (AUC). *Confusion matrix* merupakan matriks yang menampilkan prediksi klasifikasi dan klasifikasi *actual* (Mahardhika et al., 2015). *Accuracy* digunakan untuk evaluasi karena untuk mengetahui rasio prediksi benar (*true positif* dan *true negatif*) dengan data keseluruhan (Faisal et al., 2020). Nilai AUC menggambarkan hasil keseluruhan pengukuran atas kesesuaian dari model yang digunakan. Nilai AUC yang semakin besar memiliki arti bahwa variabel yang diteliti semakin baik dalam memprediksi kejadian (Maskoen & Purnama, 2018).

3. Pengujian Data

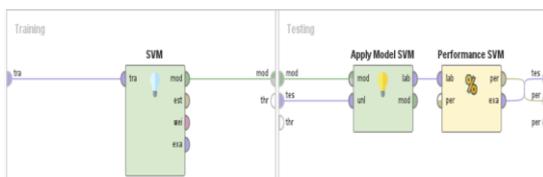
Tahapan pengujian data menggunakan pengujian data menggunakan algoritma SVM, NB, SVM+PSO dan NB+PSO. Pada penelitian ini nilai *k-fold cross validation* disajikan pada gambar 6, diatur pada angka 10 yang artinya dataset dibagi menjadi 10 area, dengan masing-masing aspek memberikan informasi yang sama persentase setiap jenis data. 9 area data digunakan dalam metode pelatihan sehingga membentuk model dan 1 area digunakan dalam prosedur pengujian.

Pada model *cross-validation* ini ada beberapa operator yaitu algoritma untuk training sehingga menghasilkan model dan *testing performance* model

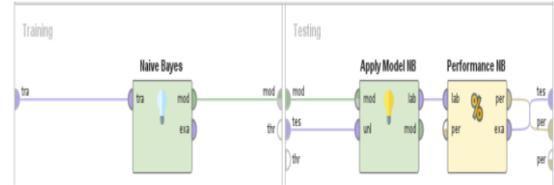


Gambar 6. K-Folds Cross Validation

Operator training tersebut menghasilkan model dan *testing performance* sebagaimana ditampilkan pada gambar 7 dan gambar 8.



Gambar 7. Training dan Testing Support Vector Machine



Gambar 8. Training dan Testing Naive Bayes

HASIL DAN PEMBAHASAN

Tahap berikutnya melakukan evaluasi atas hasil pengujian data. Evaluasi dilakukan menggunakan nilai akurasi pada tabel *confusion matrix* dan nilai AUC pada kurva ROC.

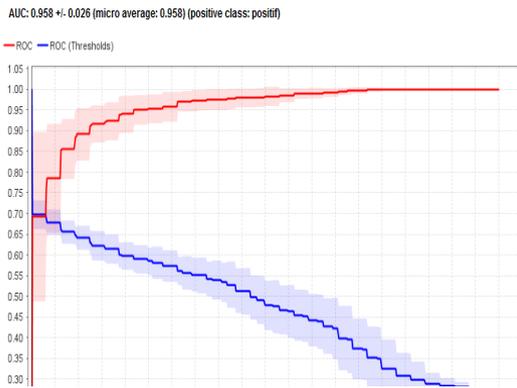
1. Algoritma Support Vector Machine

Confusion matrix yang dihasilkan dari pengujian algoritma SVM sebagaimana disajikan pada tabel 2.

Tabel 2. *Confusion Matrix* Algoritma SVM

Accuracy: 84,95% +/- 3.93% (micro average: 84,95%)			
	True Negatif	True positif	Class Precision
Pred. negatif	166	11	93,79%
Pred. positif	152	754	83,22%
Class recall	52,20%	98,56%	

Pada tabel 2 menunjukkan bahwa nilai akurasi sebesar 84,95% dengan toleransi kesalahan sebesar 3,93%, dengan nilai *true negatif* 166 records dan *true positif* 754 records. Hal tersebut menunjukkan bahwa data yang diklasifikasikan sesuai dengan sentimen negatif sebanyak 166 records dan data yang diklasifikasikan sesuai dengan sentimen positif sebanyak 754 records. Sedangkan sebanyak 152 records seharusnya sentimen negatif tetapi diklasifikasikan positif (*false positif*) dan sebanyak 11 records yang seharusnya diklasifikasikan sentimen positif namun diklasifikasikan negatif (*false positif*). Berdasarkan hasil pengujian *performance* menghasilkan Kurva Receiver Operating Characteristic (ROC) seperti pada gambar 9 dan nilai *Area Under Curve* (AUC) yang didapatkan sebesar 0,958 (*good classification*).



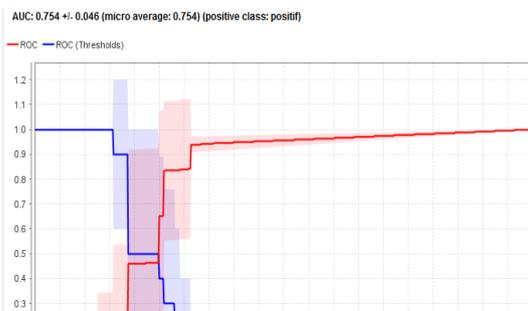
Gambar 9. Training dan Testing Naive Bayes

2. Algoritma Naive Bayes

Confusion matrix yang dihasilkan dari pengujian algoritma NB sebagaimana disajikan pada tabel 3. Pada tabel 3 menunjukkan bahwa nilai akurasi sebesar 87,54% dengan toleransi kesalahan sebesar 1,88%, dengan nilai *true negatif* 249 records dan *true positif* 699 records. Hal tersebut menunjukkan bahwa data yang diklasifikasikan sesuai dengan sentimen negatif sebanyak 249 records dan data yang diklasifikasikan sesuai dengan sentimen positif sebanyak 699 records. Sedangkan sebanyak 69 records seharusnya sentimen negatif tetapi diklasifikasikan positif (*false positif*) dan sebanyak 66 records yang seharusnya diklasifikasikan sentimen positif namun diklasifikasikan negatif (*false positif*). Berdasarkan hasil pengujian *performance* menghasilkan *Kurva ROC* seperti pada gambar 10 dan nilai AUC yang didapatkan sebesar 0,817 (*good classification*).

Tabel 3 *Confusion Matrix* Algoritma NB

Accuracy: 87,54% +/-1.88% (micro average:87,53%)			
	True Negatif	True Positif	Class Precision
Pred. negatif	249	66	79,05%
Pred. positif	69	699	91,02%
Class recall	78,30%	91,37%	



Gambar 10. Training dan Testing Naive Bayes

3. SVM+PSO

Confusion matrix yang dihasilkan dari pengujian algoritma SVM+PSO sebagaimana disajikan pada tabel 4.

Tabel 4 *Confusion Matrix* Algoritma SVM+PSO

Accuracy: 86,52% +/-3.13% (micro average:86,52%)			
	True Negatif	True Positif	Class Precision
Pred. negatif	181	9	95,26%
Pred. positif	137	756	84,66%
Class recall	56,92%	98,82%	

Pada Tabel 4 menunjukkan bahwa nilai akurasi sebesar 86,52% dengan toleransi kesalahan sebesar 3,13%, dengan nilai *true negatif* 181 records dan *true positif* 756 records. Hal tersebut menunjukkan bahwa data yang diklasifikasikan sesuai dengan sentimen negatif sebanyak 181 records dan data yang diklasifikasikan sesuai dengan sentimen positif sebanyak 754 records. Sedangkan sebanyak 137 records seharusnya sentimen negatif tetapi diklasifikasikan positif (*false positif*) dan sebanyak 9 records yang seharusnya diklasifikasikan sentimen positif namun diklasifikasikan negatif (*false positif*). Berdasarkan hasil pengujian *performance* menghasilkan *Kurva ROC* seperti pada gambar 11 dan nilai AUC yang didapatkan sebesar 0,948 (*good classification*).



Gambar 11. Grafik ROC Algoritma SVM+PSO

4. NB+PSO

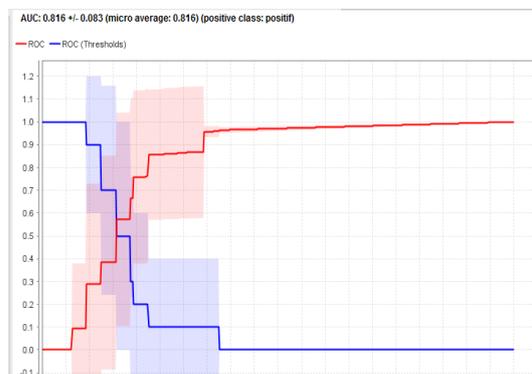
Confusion matrix yang dihasilkan dari pengujian algoritma NB+PSO sebagaimana disajikan pada Tabel 5.

Tabel 5. *Confusion Matrix* Algoritma NB+PSO

Accuracy: 90,12% +/-3.96% (micro average:90,12%)			
	True Negatif	True Positif	Class Precision
Pred. negatif	265	54	83,07%
Pred. positif	53	711	93,06%
Class recall	83,33%	92,94%	

Pada tabel 5 menunjukkan bahwa nilai akurasi sebesar 90,12% dengan toleransi kesalahan sebesar 3,96%, dengan nilai *true negatif* 265 records dan *true positif* 711 records. Hal tersebut menunjukkan bahwa data yang diklasifikasikan sesuai dengan sentimen negatif sebanyak 265 records dan data yang diklasifikasikan sesuai dengan sentimen positif sebanyak 711 records. Sedangkan sebanyak 53 records seharusnya sentimen negatif tetapi diklasifikasikan positif (*false positif*) dan sebanyak 54 records yang seharusnya diklasifikasikan sentimen positif namun diklasifikasikan negatif (*false positif*). Berdasarkan hasil pengujian *performance* menghasilkan *Kurva ROC* seperti pada gambar 12 dan nilai *Area Under AUC* yang didapatkan sebesar 0,816 (*good classification*).

Pada tabel 6 disajikan perbandingan *accuracy* dan AUC keempat algoritma. Hasil pengujian dengan Algoritma SVM mendapatkan nilai *accuracy* 84,95% dengan nilai AUC sebesar 0,958, berikutnya pengujian dengan algoritma *Naïve Bayes* menghasilkan *accuracy* 87,54% dengan nilai AUC sebesar 0,817, dengan hasil ini *accuracy* kedua algoritma tersebut terpaut selisih sebesar 2,59%. Sedangkan hasil pengujian pada algoritma SVM yang telah dioptimasi dengan PSO mendapatkan hasil 86,52% dengan nilai AUC 0,948%, berikutnya hasil pengujian pada algoritma *Naïve Bayes* yang telah dioptimasi mendapatkan nilai *accuracy* sebesar 90,12% dengan nilai AUC 0,816, dengan hasil ini *accuracy* kedua algoritma yang telah dioptimasi dengan PSO tersebut terpaut selisih sebesar 3,60%.

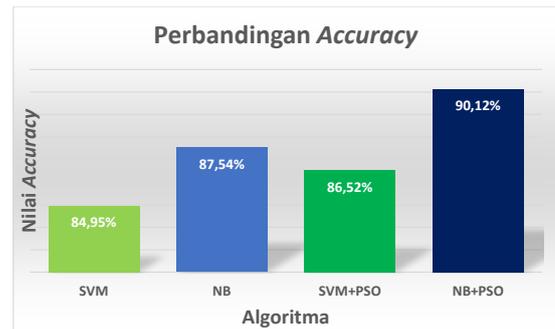


Gambar 12. Grafik ROC Algoritma NB+PSO

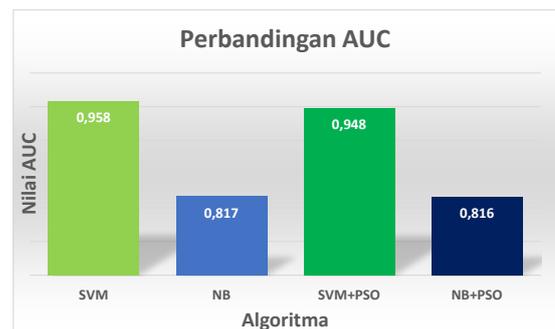
Tabel 6. Perbandingan *Accuracy* dan AUC

Algoritma	<i>Accuracy</i>	AUC
SVM	84,95%	0,958
NB	87,54%	0,817
SVM+PSO	86,52%	0,948
NB+PSO	90,12%	0,816

Perbandingan nilai *accuracy* dan AUC ke empat algoritma disajikan dalam bentuk grafik dapat dilihat pada gambar 13 dan gambar 14.



Gambar 13. Grafik Perbandingan *Accuracy*



Gambar 14. Grafik Perbandingan AUC

Dari hasil rekapitulasi tersebut menunjukkan bahwa evaluasi menggunakan algoritma NB dengan PSO mendapatkan performa terbaik dan dapat menjadi solusi untuk mendapatkan nilai *accuracy* tertinggi, sedangkan untuk hasil AUC performa terbaik menggunakan algoritma SVM tanpa PSO.

KESIMPULAN

Pengujian terhadap data hasil *crawling* dari media sosial *twitter* dengan *query Omnibus Law* dengan algoritma klasifikasi telah berhasil dilakukan. Berdasarkan hasil analisis sentimen terhadap *Omnibus Law* di media sosial *twitter* dengan menggunakan algoritma klasifikasi berbasis *Particle Swarm Optimization*, dapat ditarik kesimpulan bahwa penggunaan algoritma NB+PSO menghasilkan nilai akurasi lebih baik dibandingkan dengan algoritma SVM, NB dan SVM+PSO. Sehingga algoritma NB yang dioptimasi dengan *Particle Swarm Optimization* menjadi solusi untuk melakukan klasifikasi sentimen masyarakat Indonesia terhadap *Omnibus Law*. Saran untuk penelitian selanjutnya agar proses *crawling* tidak hanya dilakukan pada media sosial *twitter* dan optimalisasi algoritma dengan menggunakan operator *feature selection by correlation* agar dapat dibandingkan dan mendapat akurasi yang paling baik.

REFERENSI

- Achyani, Y. E. (2018). *Penerapan Metode Particle Swarm Optimization Pada Optimasi Prediksi Pemasaran Langsung*. 5(1), 1–11.
- Alkhalifi, Y., Gata, W., Prasetyo, A., & Budiawan, I. (2020). Analisis Sentimen Penghapusan Ujian Nasional pada *Twitter* Menggunakan Support Vector Machine dan Naïve Bayes berbasis Particle Swarm Optimization. *CoreIT*, 6(2).
- Ariadi, D., & Fithriasari, K. (2015). Klasifikasi Berita Indonesia Menggunakan Metode Naive Bayesian Classification dan Support Vector Machine dengan Confix Stripping Stemmer. *JURNAL SAINS DAN SENI ITS Vol. 4, No.2, 4(2)*, 248–253.
- Cahyaningrum, N. I., Yoshida Fatima, D. W., Kusuma, W. A., Ramadhani, S. A., Destanto, M. R., & Nooraeni, R. (2020). Analysis of User Sentiment of *Twitter* to Draft KUHP. *Jurnal Matematika, Statistika Dan Komputasi*, 16(3), 273. <https://doi.org/10.20956/jmsk.v16i3.8239>
- Faisal, A., Alkhalifi, Y., Rifai, A., & Gata, W. (2020). Analisis Sentimen Dewan Perwakilan Rakyat Dengan Algoritma Klasifikasi Berbasis Particle Swarm Optimization. *JOINTECS (Journal of Information Technology and Computer Science)*, 5(2), 61. <https://doi.org/10.31328/jointecs.v5i2.1362>
- Fitryantica, A. (2019). Harmonisasi Peraturan Perundang-Undangan Indonesia melalui Konsep Omnibus Law. *Gema Keadilan*, 6(3), 300–316. <https://ejournal2.undip.ac.id/index.php/gk/article/view/6751>
- Gunawan, F., Fauzi, M. A., & Adikara, P. P. (2017). Analisis Sentimen Pada Ulasan Aplikasi Mobile Menggunakan Naive Bayes dan Normalisasi Kata Berbasis Levenshtein Distance (Studi Kasus Aplikasi BCA Mobile). *Systemic: Information System and Informatics Journal*, 3(2), 1–6. <https://doi.org/10.29080/systemic.v3i2.234>
- Haddi, E., Liu, X., & Shi, Y. (2013). The Role of Text Pre-processing in Sentiment Analysis The Role of Text Pre-processing in Sentiment Analysis. *Procedia Computer Science*, 17(December), 26–32. <https://doi.org/10.1016/j.procs.2013.05.005>
- Han, J., Kamber, M., & Pei, J. (2012). Data mining: Data mining concepts and techniques. In *Proceedings - 2013 International Conference on Machine Intelligence Research and Advancement, ICMIRA 2013*. <https://doi.org/10.1109/ICMIRA.2013.45>
- Hayatin, N., Marthasari, G. I., Nuraini, L., & Info, A. (2020). Optimization of Sentiment Analysis for Indonesian Presidential Election using Naïve Bayes and Particle Swarm Optimization. *JOIN (Jurnal Online Informatika)*, 5(1), 81–88. <https://doi.org/10.15575/join.v5i1.558>
- Heranova, O. (2019). Synthetic Minority Oversampling Technique pada Averaged One Dependence Estimators untuk Klasifikasi Credit Scoring. *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, 3(3), 443–450. <https://doi.org/10.29207/resti.v3i3.1275>
- Hermawan, L., & Bellanar Ismiati, M. (2020). Pembelajaran Text Preprocessing berbasis Simulator Untuk Mata Kuliah Information Retrieval. *Jurnal Transformatika*, 17(2), 188. <https://doi.org/10.26623/transformatika.v17i2.1705>
- Istighfarin, N. F., Rahmastati, R. A., & Nugroho, H. (2020). Penerapan Metode Particle Swarm Optimization (PSO) Dan Genetic Algorithm (GA) Pada Sistem Optimasi Visible Light Communication (VLC) Untuk Menentukan Posisi Robot. *Simetris: Jurnal Teknik Mesin, Elektro Dan Ilmu Komputer*, 11(1), 279–286. <https://doi.org/10.24176/simet.v11i1.4052>
- Kusmarna, I., Wardhani, L. K., & Safrizal, M. (2015). Aplikasi Penjadwalan Mata Kuliah Menggunakan Algoritma Particle Swarm Optimization (Pso). *Jurnal Teknik Informatika*, 8(2), 1–8. <https://doi.org/10.15408/jti.v8i2.2441>
- Mahardhika, A. A., Saptono, R., & Anggrainingsih, R. (2015). Sistem Klasifikasi Feedback Pelanggan Dan Rekomendasi Solusi Atas Keluhan Di Upt Puskom Uns Dengan Algoritma. *Jurnal ITSMART*, 4(1), 36–42.
- Maskoen, T. T., & Purnama, D. (2018). Area Under the Curve dan Akurasi Cystatin C untuk Diagnosis Acute Kidney Injury pada Pasien Politrauma. *Majalah Kedokteran Bandung*, 50(4), 259–264. <https://doi.org/10.15395/mkb.v50n4.1342>
- Medhat, W., Hassan, A., & Korashy, H. (2014). Sentiment analysis algorithms and applications: A survey. *Ain Shams Engineering Journal*, 5(4), 1093–1113. <https://doi.org/10.1016/j.asej.2014.04.011>
- Mihuandayani, Utami, E., & Luthfi, E. T. (2018). Text mining based on tax comments as big data analysis using SVM and feature selection. *2018 International Conference on Information and Communications Technology, ICOIACT 2018, 2018-Janua*, 537–542. <https://doi.org/10.1109/ICOIACT.2018.8350743>
- Putra, A. (2020). Penerapan Omnibus Law Dalam Upaya Reformasi Regulasi. *Jurnal Legislasi Indonesia, Vol 17(12)*, 1–10.

- Putranti, N. D., & Winarko, E. (2014). Analisis Sentimen *Twitter* untuk Teks Berbahasa Indonesia dengan Maximum Entropy dan Support Vector Machine. *IJCCS (Indonesian Journal of Computing and Cybernetics Systems)*, 8(1), 91. <https://doi.org/10.22146/ijccs.3499>
- Que, V. K. S., Iriani, A., & Purnomo, H. D. (2020). Analisis Sentimen Transportasi Online Menggunakan Support Vector Machine Berbasis Particle Swarm Optimization. *Jurnal Nasional Teknik Elektro Dan Teknologi Informasi*, 9(2), 162–170. <https://doi.org/10.22146/jnteti.v9i2.102>
- Razy, M. F., & Fedryansyah, M. (2020). Konflik Gerakan Masyarakat Sipil Dan Pemerintah Dalam Proses Penyusunan Rancangan Undang-Undang Omnibus Law. *Jurnal Kolaborasi Resolusi Konflik*, 2(2), 74. <https://doi.org/10.24198/jkrk.v2i2.28147>
- Santoso, V. I., Virginia, G., Lukito, Y., Kristen, U., & Wacana, D. (2017). *PENERAPAN SENTIMENT ANALYSIS PADA HASIL EVALUASI DOSEN DENGAN METODE SUPPORT VECTOR MACHINE*. 14(1), 79–83.
- Saputra, D. D., Gata, W., Wardhani, N. K., & Sakho, K. (2020). *Optimization Sentiments of Analysis from Tweets in myXLCare using Naïve Bayes Algorithm and Synthetic Minority Over Sampling Technique Method Optimization Sentiments of Analysis from Tweets in myXLCare using Naïve Bayes Algorithm and Synthetic Minority Ov.* <https://doi.org/10.1088/1742-6596/1471/1/012014>
- Styawati, & Mustofa, K. (2019). *A Support Vector Machine-Firefly Algorithm for Movie Opinion Data Classification*. 13(3), 219–230.
- Zamachsari, F., Saragih, G. V., Susafa'ati, & Gata, W. (2020). Analisis Sentimen Pemindahan Ibu Kota Negara dengan Feature Selection. *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, 4(3), 504–512.

PROFIL PENULIS

Annisa Elfina Augustia, Memperoleh gelar Strata 1 di STMIK Nusa Mandiri pada program studi Teknik Informatika tahun 2019 dan saat ini merupakan mahasiswa pasca sarjana di Universitas Nusa Mandiri pada program studi Ilmu Komputer.

Resi Taufan, Memperoleh gelar Strata 1 di STMIK Nusa Mandiri pada program studi Sistem Informasi tahun 2017 dan dan memperoleh gelar Strata 2 di Universitas Nusa Mandiri tahun 2021 pada program studi Ilmu Komputer.

Yuris Alkhalifi, Memperoleh gelar Strata 1 di STMIK Nusa Mandiri pada program studi Sistem Informasi tahun 2019 dan dan memperoleh gelar Strata 2 di Universitas Nusa Mandiri tahun 2021 pada program studi Ilmu Komputer.

Windu Gata, Dosen Pasca Sarjana Universitas Nusa Mandiri pada program studi Ilmu Komputer.