

Sentimen Analisis Operasi Tangkap Tangan KPK Menurut Masyarakat Menggunakan Algoritma Support Vector Machine, Naive Bayes Berbasis Particle Swarm Optimization

HERNAWATI
WINDU GATA

Magister Ilmu Komputer

STMIK Nusa Mandiri Jakarta

Jl. Kramat Raya No.18, RT.1/RW.7, Kwitang, Kec. Senen, Kota Jakarta Pusat

Email: 14002083@nusamandiri.ac.id, wgata@nusamandiri.ac.id

Abstract. *It is known from various public sentiments conveyed through comments on social media twitter against the capture operations carried out by the corruption eradication commission (KPK) that currently it does not meet the expectations of the community, where officials who are only officials have small corruption rates, not corruption As for the classification algorithms that have strong accuracy at this time are Support Vector Machine and Naïve Bayes algorithms, calculation of Support Vector Machine method for tweet data from 78 positive tweet data and 78 negative tweet data, resulting in an accuracy of 80.77% and AUC 0.867. Whereas the results of accuracy with the Naïve Bayes method are 76.92% and AUC 0.729. Having a difference in accuracy of 3.3%, and after optimizing with the Operator Vector Machine (PSO) weight Particle Swarm Optimization the accuracy is 83.79% and AUC 0.910, while for Naïve Bayes (PSO) produces an accuracy of 80.13% and AUC 0.771 Has a difference in accuracy of 3.6%.*

Keywords: *KPK, Support Vector Machine, Naïve Bayes, Sentiment Analysis, Particle Swarm Optimization.*

Abstrak. Diketahui dari berbagai sentimen masyarakat yang disampaikan melalui komentar di media sosial *twitter* terhadap operasi tangkap tangan yang dilakukan oleh Komisi Pemberantasan Korupsi (KPK) nyatanya saat ini belum memenuhi harapan masyarakat, dimana pejabat yang di ott hanya pejabat yang mempunyai angka korupsi kecil, bukan korupsi yang besar adapun algoritma klasifikasi yang kuat akurasi saat ini adalah algoritma *Support Vector Machine* untuk data *tweet* dari 78 data *tweet* positif dan 78 data *tweet* negatif, menghasilkan akurasi sebesar 80.77% dan AUC 0.867. Sedangkan hasil akurasi dengan metode *Naïve Bayes* adalah 76.92% dan AUC 0.729. Memiliki selisih akurasi sebesar 3.3%, dan setelah di optimalisasi dengan oprator *Weight Partical Swarm Optimization* untuk *Support Vector Machine (PSO)* menghasilkan akurasi 83.79% dan AUC 0.910, sedangkan untuk *Naïve Bayes (PSO)* menghasilkan akurasi sebesar 80.13% dan AUC 0.771 memiliki selisih akurasi sebesar 3.6%.

Kata Kunci: *KPK, Support Vector Machine, Naïve Bayes, Sentimen Analisis, Particle Swarm Optimization.*

PENDAHULUAN

Hukum sebagai kumpulan peraturan atau kaedah mempunyai isi yang bersifat umum dan normatif, hukum tidak mungkin ada tanpa adanya lembaga yang merumuskan, melaksanakan dan menegakkannya, yaitu lembaga legislatif, eksekutif dan yudikatif.

KPK adalah komisi di Indonesia yang dibentuk pada tahun 2003 untuk mengatasi, menanggulangi dan memberantas korupsi di Indonesia. Komisi ini didirikan berdasarkan Undang-Undang Republik Indonesia Nomor 30 Tahun 2002 mengenai Komisi Pemberantasan Tindak Pidana Korupsi. Demi tercapainya pelayanan publik yang optimal bagi masyarakat, maka pemerintah memandang perlu untuk membentuk lembaga baru (Priambada, Bintara Sura. 2008).

Pasal 11 Undang-Undang No. 30 Tahun 2002 Tentang Komisi Pemberantasan Tindak Pidana Korupsi memberikan mandat kepada Komisi Pemberantasan Korupsi (KPK) untuk melakukan pemberantasan korupsi terhadap tiga hal: pejabat publik, penegak hukum, dan pihak-pihak yang terkait dengan keduanya. Dengan landasan yuridis yang jelas, KPK dianggap berhasil memenuhi harapan publik dalam melakukan pemberantasan korupsi di Indonesia.

Kesadaran kolektif bangsa akan bahaya perilaku korupsi muncul seiring dengan tumbanganya Orde Baru dan lahirnya reformasi, sehingga sudah menjadi pengetahuan umum bahwa salah satu amanat reformasi adalah pemberantasan terhadap Korupsi, Kolusi dan Nepotisme (KKN). Namun pertanyaannya, apakah betul era di reformasi seperti sekarang ini ada pemberantasan korupsi yang sesungguhnya dari segenap bangsa? Karena yang terlihat jelas bukanlah pemberantasan korupsi, namun sebaliknya, yakni “memberantas pemberantas korupsi”, yakni KPK (Muttaqin, dkk. 2018).

Sementara itu, OTT tidak disebut dalam KUHAP dan UU Tipikor, serta UU KPK. OTT adalah istilah KPK untuk “menangkap basah” seseorang yang diduga sebagai pelaku korupsi. Sebuah operasi yang rahasia, terukur dan jarang korbannya dapat selamat dari tuduhan karena didasari dengan proses yang panjang ketika KPK “mengendus” adanya korupsi.

Istilah OTT baru muncul dalam Peraturan Presiden Republik Indonesia (Perpres) Nomor 87 Tahun 2016 tentang Satuan Tugas Sapu Bersih Pungutan Liar yang selanjutnya disebut Satgas Saber Pungli, yang berkedudukan di bawah dan bertanggung jawab kepada Presiden (Hikmawati, Puteri. 2018).

Menurut Lembaga Survei Indonesia (LSI) dan Indonesia Corruption Watch (ICW) merilis hasil survei nasional yang betajuk “*Tren Persepsi Publik tentang Korupsi di Indonesia*”. Dari hasil survei tersebut mayoritas responden menilai bahwa tingkat korupsi mengalami peningkatan dengan hasil 52%. (sumber: www.cnnindonesia.com, 10 Desember 2018).

Dalam laporan kinerja 2018 Komisi Pemberantasan Korupsi (KPK) meliris sejumlah capaian selama satu tahun, mulai dari serapan anggaran yang mendekati 90% sampai rekor Operasi Tangkap Tangan (OTT) terbanyak sepanjang tahun. (sumber: www.tempo.co, 20 Desember 2018).

Di era digital sebagian besar kita memiliki akun media sosial seperti *Facebook*, *Twitter*, *Liknedin*, *Instagram* dan media sosial lainnya. Dikehidupan sehari-harinya banyak orang menuangkan ide dari pikirannya yang dibagikan kedalam berbagai status media sosial yang dimiliki, atau sekedar memberikan komentar atas status orang lain.

Di *Twitter* setiap orang membuat *tweet* apapun jenis *tweetnya*, terkadang dalam *tweet* seseorang terdapat *tweet* yang hanya terdiri dari penggalan kata yang maknanya sulit diketahui, tidak jarang seseorang membuat *tweet* sampah yang memang tidak ada artinya dalam referensi besar seperti dalam Kamus Besar Bahasa Indonesia jika mengkhususkan hanya pada *tweet* yang berbahasa Indonesia. Jadi data yang dihasilkan

banyak yang berharga tetapi sedikit pengetahuan yang didapat dari data-data tersebut, maka dari itu peran sebuah ilmu yang dinamakan *Data Mining*.

Tinjauan Pustaka

Penelitian dengan judul *Sentiment Analysis Article News Coordinator Minister Of Maritime Affairs Using Agorithm Naïve Naves And Support Vector Machine With Particle Swarm Optimization* (Wardhani, dkk. 2018). Dalam penelitiannya menghasilkan akurasi sebesar 89.50% dengan AUC sebesar 0.500 untuk *Naïve Bayes*, sedangkan *Naïve bayes (PSO)* menghasilkan akurasi 92.00% dengan AUC sebesar 0.550. Untuk hasil yang didapat dengan menggunakan metode *Support Vector Machine* menghasilkan akurasi 87.50% dengan AUC sebesar 0.979, dan untuk *Support Vector Machine (PSO)* menghasilkan akurasi sebesar 90.50% dan AUC sebesar 0.975.

Penelitian dengan judul *And the Winner is...: Bayesian Twitter-Based Prediction on 2016 U.S. Presidential Election* (Tunggawan, E. 2016). Dalam penelitiannya menjelaskan penggunaan algoritma model prediksi *Naïve Bayes* untuk pemilihan presiden Amerika Serikat tahun 2016 berdasarkan data *Twitter*, peneliti menggunakan *hashtag* resmi dan metode preprocessing sederhana untuk menyiapkan data tanpa memodifikasi maksud atau artinya, pemodelan dapat mencapai akurasi 95,8%. Selama pengetesan model dan memprediksi jajak pendapat dengan akurasi 54,8%. Pemodelan ini memprediksi bahwa *Barnie Sanders* dan *Ted Cruz* akan menjadi nominasi Partai Demokrat dan Republik.

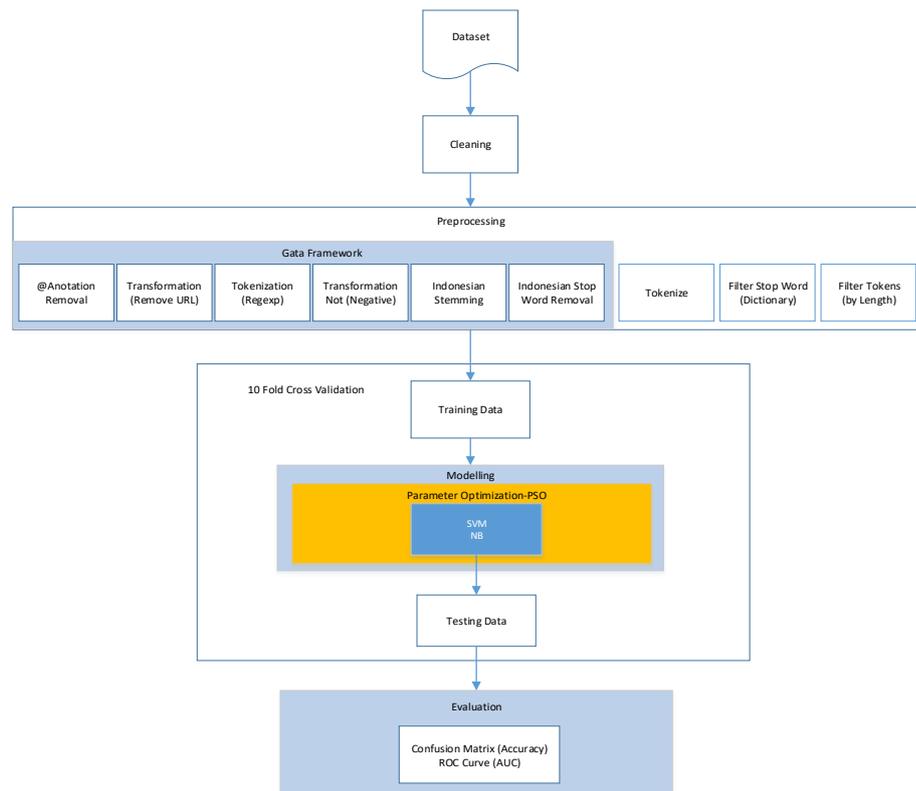
Penelitian dengan judul *Sentiment Analysis Untuk Menilai Kepuasan Masyarakat Terhadap Kinerja Pemerintah Daerah Menggunakan Naïve Bayes Classifier (studi kasus: Walikota Bandung Periode 2013 – 2018)* (Soepriadi, dkk. 2018). Dalam penelitiannya yaitu untuk mengetahui kinerja pejabat daerah melalui opini publik yang disampaikan melalui media sosial, nyatanya saat ini belum memenuhi harapan masyarakat, sehingga menimbulkan citra yang kurang baik. Penerapan metode *Naïve Bayes Classifier* ini dapat membantu walikota sebagai bahan pertimbangan menentukan kebijakan dari hasil pengklasifikasian sentimen berdasarkan bidang yang diuji.

Penelitian dengan judul *Analisis Sentimen Untuk Menilai Kepuasan Masyarakat Terhadap Kinerja Puan Maharani Di Kabinet Kerja Menggunakan Support Vector Machine Dan Naïve Bayes* (Yang, J., & Honavar, V. 1998). Dalam penelitiannya menghasilkan akurasi 90.38% dan AUC sebesar 0.968 untuk Algoritma *Support Vector Machine*, sedangkan untuk *Naïve Bayes* menghasilkan akurasi 90.87% dengan AUC sebesar 0.892.

Kerangka Berpikir

Mengenai klasifikasi tweet sentimen analisis Operasi Tangkap Tangan KPK, kerangka penulisan dapat dijelaskan sebagai berikut. Penelitian ini terdiri dari dua tahap *Text Preprocessing*, yaitu dari *Gataframework* dan *RapidMiner Text Preprocessing*. Dalam penelitian ini alat *textmining* pada *Gataframework* untuk melakukan preprocessing versi bahasa Indonesia, salah satunya adalah bahasa Indonesia stemming. Berdasarkan review dari review studi di atas, dapat dilihat bahwa *Support Vector Machine* adalah klasifikasi yang sering digunakan untuk menyelesaikan masalah analisis sentimen. Dalam penelitian ini, klasifikasi algoritma *Support Vector Machine*, *Naïve Bayes* digunakan sebagai algoritma yang dibandingkan, *Particle Swarm Optimization* digunakan untuk teknik optimasi yang sangat sederhana untuk menerapkan dan memodifikasikan beberapa parameter diantaranya meningkatkan atribut atau variable yang dipakai untuk menseleksi atribut (*attribute selection*) dan *feature selection*. Membuat model klasifikasi tweet untuk mendapatkan tingkat akurasi menggunakan algoritma *Support Vector Machine (SVM)*, *Naïve Bayes (NB)* berbasis *Particle Swarm*

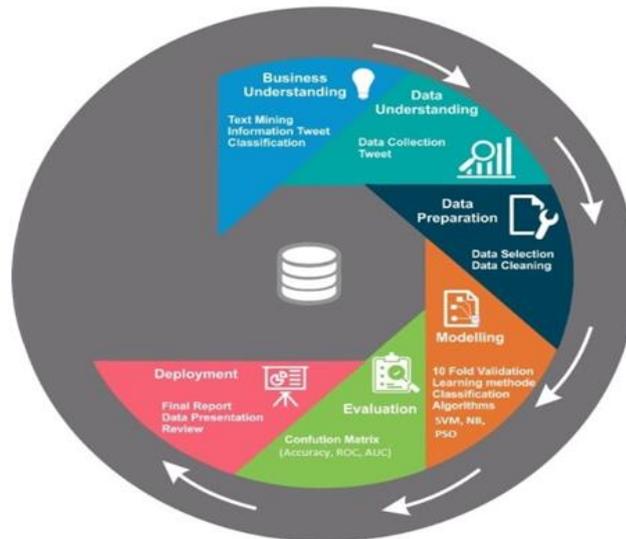
Optimization (PSO). Dataset akan digunakan untuk mengevaluasi model yang terbentuk akan diuji menggunakan *Confusion matrix* yang akan menggambarkan hasil akurasi mulai dari prediksi positif yang benar, prediksi positif yang salah, prediksi negative yang benar, dan prediksi negative yang salah. Akurasi akan dihitung dari seluruh hasil prediksi yang benar (baik prediksi positif dan negatif) dibandingkan dengan seluruh data *testing*. Semakin tinggi nilai akurasi, semakin baik pula model yang dihasilkan. Pengujian juga diukur dengan menggunakan *ROC Curve*. *ROC Curve* akan menggambarkan kelas positif dalam bentuk kurva. Pengujian dilakukan dengan menghitung nilai *AUC* (*Area Under Curve*), semakin tinggi nilai *AUC* dan *ROC Curve*, maka semakin baik pula model klasifikasi yang terbentuk.



Gambar 1. Kerangka Berpikir.

METODE

Metode penelitian yang dilakukan dengan text mining pada tweet untuk melakukan klasifikasi dari tweet yang berkaitan dengan sentimen analisis operasi tangkap tangan KPK. Metodologi penelitian yang digunakan dalam penelitian eksperimen ini dengan menggunakan metode Cross-Industry Standard Process for Data Mining (CRISP-DM) terdiri dari enam tahap yaitu Business Understanding, Data Understanding, Data Preparation, Modelling, Evaluation, dan Deployment.



Gambar 2. CRISP-DM Disesuaikan.

HASIL DAN PEMBAHASAN

Business Understanding.

Tahap ini diperlukan pemahaman mengenai objek penelitian dengan menggali informasi dari media sosial yaitu data tweet, yang bersumber dari twitter mengenai dengan sentimen analisis operasi tangkap tangan KPK, dari posting tweet bisa mengetahui komentar masyarakat, klasifikasi tweet dilakukan untuk melihat tweet yang bernilai positif dan negatif.

Berdasarkan ini, pendekatan model klasifikasi tweet yang akan digunakan algoritma Support Vector Machine (SVM), Naïve Bayes (NB) berbasis Particle Swarm Optimization.

Data Understanding.

Proses pengambilan data mentah (data twitter murni) sesuai dengan atribut yang dibutuhkan. Data tersebut dikumpulkan dari periode 01 September 2018 sampai 31 Desember 2018 mengenai operasi tangkap tangan Komisi Pemberantasan Korupsi (KPK) Data yang diambil sebanyak 759 data, proses yang dilakukan adalah *cleaning* dan yang dihasilkan akan digunakan sebagai bahan yang akan diteliti.

Data Preparation.

Pada tahap yang ini yaitu mempersiapkan data untuk melakukan langkah-langkah yang disebut dengan *text preprocessing*, dengan menggunakan dua aplikasi preprocessing, pertama menggunakan Gata Framework yang diakses melalui link <http://gataframework.com/textmining> dan preprocessing dari rapidminer berikut adalah tahapannya:



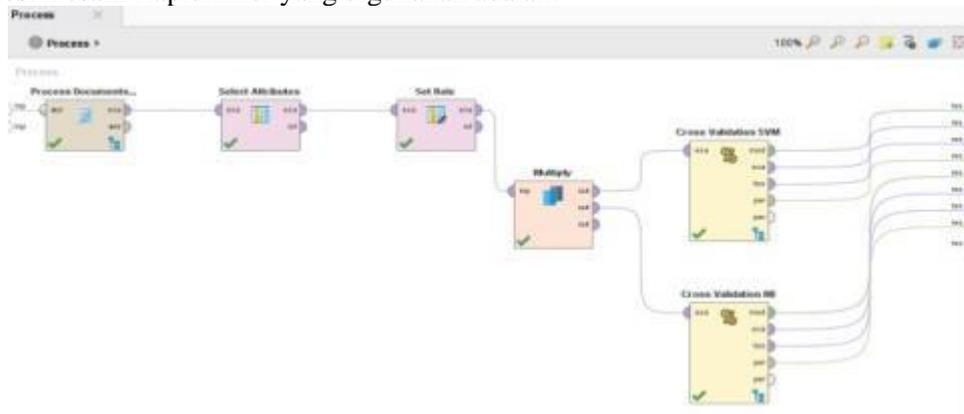
Gambar 3. Desain Model *Preprocessing Data Local* menggunakan *Gataframework*.

Pada Gambar 3, menjelaskan tentang alur proses preprocessing. Adapun tahapan-tahapan yang digunakan dalam proses *preprocessing* adalah *@Anotation Removal*, *Transformation Remove URL*, *Tokenization: Regexp*, *Transformation Not (Negative)*, *Indonesian Stemming*, *Indonesian StopWord Removal*.

Modelling

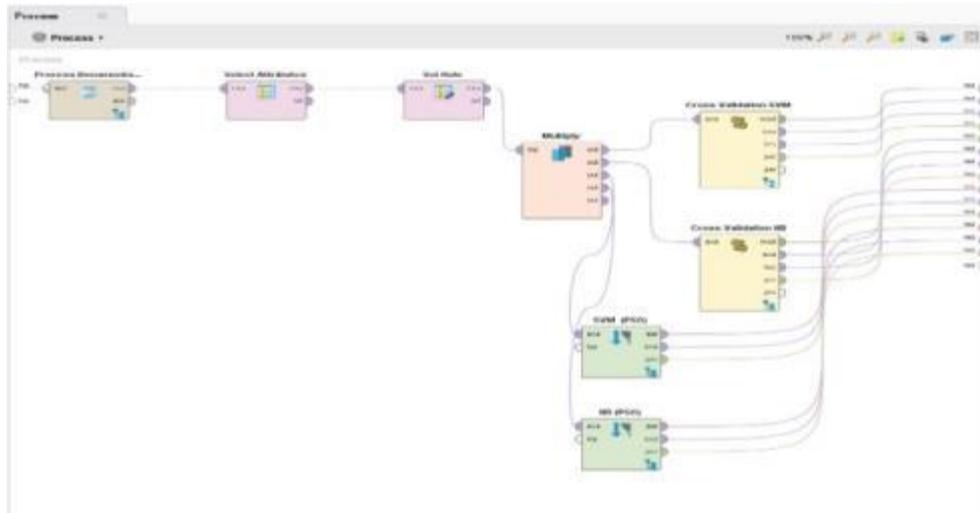
Pada fase ini secara langsung melibatkan teknik data mining yaitu dengan cara melakukan pemilihan teknik data mining dan menetapkan algoritma yang akan digunakan. *Tool* yang digunakan pada fase pemodelan ini adalah *RapidMiner* versi 7.3.0.

Adapun hasil dalam pengujian model yang dilakukan adalah mengklasifikasi *tweet* positif dan *tweet* negatif menggunakan algoritma *Support Vector Machine* dan *Naïve Bayes*. Desain *Rapidminer* yang digunakan adalah:



Gambar 4. Desain Model Komparasi *Support Vector Machine* dan *Naïve Bayes*.

Model ini terdapat operator *Proses Document from file* untuk mengimport data file *dataset training* dalam format *txt*, operator *Select Attributes* untuk mempermudah pemilihan atribut, operator *Set Role* untuk *ExampleSet* dapat memiliki atribut khusus, operator *Multiply* untuk mengambil objek dari *port input Rapidminer* dan mengirimkannya ke *port output*, operator *Cross Validation* untuk melakukan validasi.



Gambar 5. Desain model Komparasi *Support Vector Machine*, *Naïve Bayes* dengan *Support Vector Machine (PSO)* dan *Naïve Bayes (PSO)*.

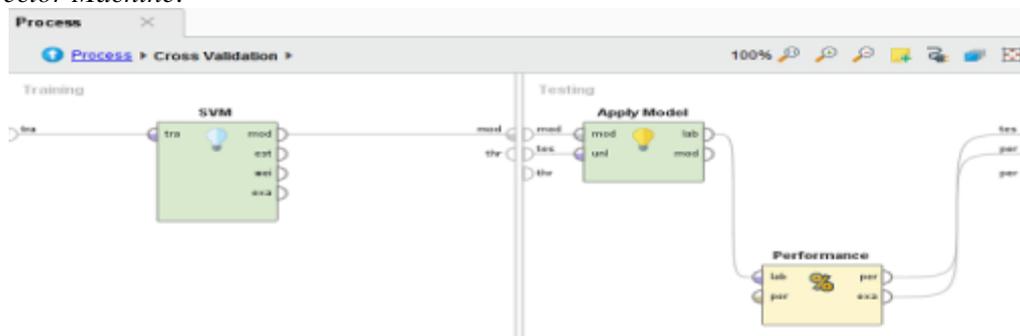
Model ini terdapat operator *Proses Document from file* untuk mengimport data file *dataset training* dalam format TXT, operator *Select Attributes* untuk mempermudah pemilihan atribut, operator *Set Role* untuk *ExampleSet* dapat memiliki atribut khusus, operator *Multiply* untuk mengambil objek dari *port input Rapidminer* dan mengirimkannya ke *port output*, operator *Weight PSO* untuk melukan optimasi, operator *Cross Validation* untuk melakukan validasi.

Evaluation.

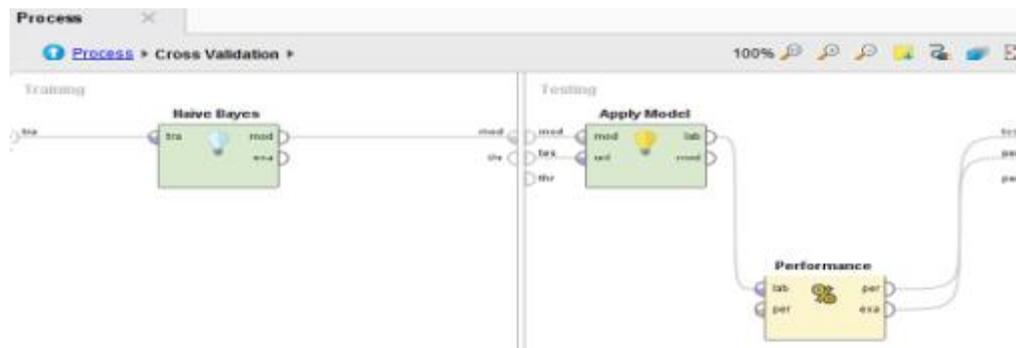
Pada evaluasi ini peneliti menggunakan 10 kali pemetaan data atau *10-fold cross validation*. Tujuan diadakan guna menentukan nilai fungsi dari model yang berhasil dilakukan pada tahapan sebelumnya dengan menggunakan satu atau lebih model.

Menurut [8], (dalam Jiang, Ping., 2017) *K-Fold Cross Validation* adalah salah satu dari jenis pengujian cross validation yang berfungsi untuk menilai kinerja proses sebuah metode algoritma dengan membagi sampel data secara acak dan mengelompokkan data tersebut sebanyak *K k-fold*. Kemudian salah satu kelompok *k-fold* tersebut akan dijadikan sebagai data uji sedangkan sisa kelompok yang lain akan dijadikan sebagai data latih.

Berikut ini adalah *Layout* desain dalam proses uji model dari metode *Support Vector Machine*:



Gambar 6. Layout Proses *Support VectorMachine* dengan 10-Fold Cross Validation.

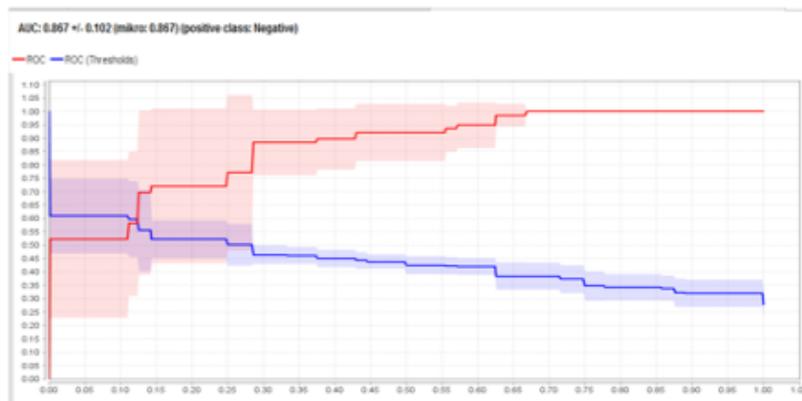


Gambar 7. Layout Proses *Naive Bayes* dengan 10-Fold Cross Validation.

Hasil Uji Data.

Kurva ROC *Support Vector Machine*.

Kurva ROC yang didapatkan dari pemodelan dengan *Support Vector Machine*.



Gambar 7. Kurva ROC *Support Vector Machine*.

Berdasarkan pengujian data dengan menggunakan *tool Rapidminer*, mendapatkan hasil sentiment masyarakat terhadap ott Komisi Pemberantasan Korupsi (KPK). Dari hasil uji model yang dilakukan untuk memperoleh hasil akurasi dan *Area Under Curve (AUC)*, mendapatkan hasil grafik ROC dengan *Area Under Curve (AUC)* sebesar 0.867 dengan *performance* akurasi adalah *Good Classification*.

Confusion Matrix *Support Vector Machine*.

Confusion matrix yang didapatkan dari pemodelan dengan *Support Vector Machine*.

Tabel 1. Confusion Matrix *Support Vector Machine*.

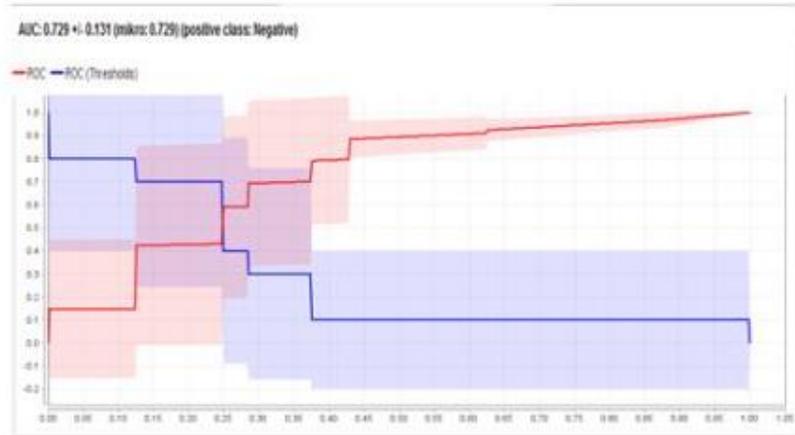
accuracy: 80.75% +/- 8.72% (mikro: 80.77%)			
	True Positif	True Negatif	Class Precision
Pred Positif	69	21	76.67%
Pred Negatif	9	57	86.36%
Class Recall	88.46%	73.08%	

Hasil akurasi yang didapat adalah 80.77% dari data *tweet* positif dan 78 *tweet* negatif tentang operasi tangkap tangan KPK. Data *tweet* positif yang sesuai dengan prediksi positif adalah 69 data. Sedangkan data negatif yang masuk ke dalam prediksi

positif adalah 21 data. *Data tweet* positif yang masuk ke dalam prediksi negatif adalah 9 data dan data *tweet* negatif yang sesuai prediksi adalah 57 data.

Kurva ROC *Naïve Bayes*.

Ini adalah hasil kurva ROC *Naïve Bayes*:



Gambar 8. Kurva ROC *Naïve Bayes*.

Berdasarkan pengujian data dengan menggunakan *tool Rapidminer*, mendapatkan hasil sentiment masyarakat terhadap ott Komisi Pemberantasan Korupsi. Dari hasil uji model yang dilakukan untuk memperoleh hasil akurasi dan *Area Under Curve (AUC)*, mendapatkan hasil grafik ROC dengan nilai *Area Under Curve (AUC)* sebesar 0.729 dengan *performance* akurasi adalah *Fair Classification*.

Confusion Matrix *Naïve Bayes*.

Confusion matrix yang didapatkan dari pemodelan dengan *Naïve Bayes*.

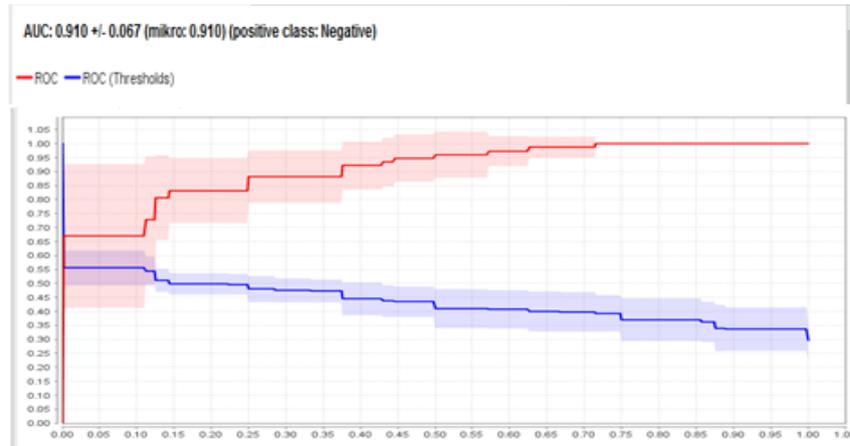
Tabel 2. Confusion Matrik *Naïve Bayes*.

Accuracy 76.83% +/- 10.40% (Mikro:76.92%)			
	True Positif	True Negatif	Class Precision
Pred Positif	63	21	75.00%
Pred Negatif	15	57	79.17%
Class Recall	80.77%	73.08%	

Hasil akurasi yang didapatkan adalah 76.92% dari 78 data *tweet* positif dan 78 data *tweet* negatif tentang sentiment masyarakat terhadap operasi tangkap tangan KPK yang menyatakan OTT terbanyak sepanjang sejarah di tahun 2018. *Data tweet* positif yang sesuai dengan prediksi positif adalah 63 data. Sedangkan *data tweet* negatif yang masuk ke dalam prediksi positif adalah 21 data. *Data tweet* positif yang masuk ke dalam prediksi negatif adalah 15 data dan *data tweet* negatif yang sesuai prediksi adalah 57 data.

Kurva ROC *Support Vector Machine (PSO)*.

Kurva ROC yang didapatkan dari pemodelan dengan *Support Vector Machine* dan PSO.



Gambar 9. Kurva ROC *Support Vector Machine* dan PSO.

Berdasarkan Gambar 9, pengujian data dengan menggunakan *tool Rapidminer*, mendapatkan hasil sentiment masyarakat terhadap operasi tangkap tangan Komisi Pemberantasan Korupsi (KPK). Dari hasil uji model yang dilakukan untuk memperoleh hasil akurasi dan *Area Under Curve (AUC)*, mendapatkan hasil grafik ROC dengan nilai *Area Under Curve (AUC)* sebesar 0.910 dengan *performance* akurasi adalah *Excellent Classification*.

Confusion Matrix *Support Vector Machine (PSO)*.

Confusion matrix yang didapatkan dari pemodelan dengan *Support Vector Machine (PSO)*.

Tabel 3. Confusion Matrix *Support Vector Machine (PSO)*.

Accuracy 83.97% +/- 10.63% (Mikro 83.97%)			
	True Positif	True Negatif	Class Precision
Pred Positif	68	15	81.93%
Pred Negatif	10	63	86.30%
Class Recall	87.18%	80.77%	

Hasil akurasi yang didapatkan adalah 83.97% dari 78 data *tweet* positif dan 78 data *tweet* negatif tentang sentiment masyarakat terhadap operasi tangkap tangan Komisi Pemberantasan Korupsi (KPK). Data *tweet* positif yang sesuai dengan prediksi positif adalah 68 data. Sedangkan data *tweet* negatif yang masuk kedalam prediksi positif adalah 15 data. Data *tweet* positif yang masuk kedalam prediksi negatif adalah 10 data dan data *tweet* negatif yang sesuai prediksi adalah 63 data.

Kurva ROC *Naïve Bayes (PSO)*.

Kurva ROC yang didapatkan dari pemodelan dengan *Naïve Bayes* dan PSO.



Gambar 10. Kurva ROC *Naive Bayes* (PSO).

Berdasarkan pengujian data dengan menggunakan *tool Rapidminer*, mendapatkan hasil sentiment masyarakat terhadap operasi tangkap tangan Komisi Pemberantasan Korupsi (KPK). Dari hasil uji model yang dilakukan untuk memperoleh hasil akurasi dan *Area Under Curve* (AUC), mendapatkan hasil grafik ROC dengan nilai *Area Under Curve* sebesar 0.771 dengan performance akurasi adalah *Fair Classification*.

Confusion Matrix *Naive Bayes* (PSO).

Confusion matrix yang didapatkan dari pemodelan dengan *Naive Bayes* dan PSO.

Tabel 4. Confusion Matrix *Naive Bayes* (PSO).

Accuracy 80.21% +/- 7.98% (Mikro 80.50%)			
	True Positif	True Negatif	Class Precision
Pred Positif	66	19	77.65%
Pred Negatif	12	59	83.10%
Class Recall	84.62%	75.64%	

Hasil akurasi yang didapatkan adalah 80.21% dari hasil 78 data *tweet* positif dan 78 data *tweet* negatif tentang operasi tangkap tangan Komisi Pemberantasan Korupsi (KPK). Data *tweet* positif yang sesuai dengan prediksi positif adalah 66 data. Sedangkan data *tweet* negatif yang masuk kedalam prediksi positif adalah 19 data. Data *tweet* positif yang masuk kedalam prediksi negatif adalah 12 data dan data *tweet* negatif yang sesuai prediksi adalah 59 data.

Tabel 5. Perbandingan Akurasi dan AUC.

Algorithm	Accuracy	AUC
SVM	80.77%	0.867
NB	76.92%	0.729
SVM + PSO	83.97%	0.910
NB + PSO	80.21%	0.7771

Dalam penelitian ini, menghasilkan perhitungan metode *Support Vector Machine* untuk data *tweet* dari 78 data *tweet* positif dan data *tweet* negatif, menghasilkan akurasi sebesar 80.77% dan AUC 0.867. Sedangkan hasil akurasi dengan metode *Naive Bayes* adalah 76.92% dan AUC 0.729. Memiliki selisih akurasi sebesar 3.3%, dan setelah di optimalisasi dengan operator *Weight Particle Swarm Optimize* untuk *Support Vector*

Machine (PSO) menghasilkan akurasi 83.79% dan AUC 0.910, sedangkan untuk *Naïve Bayes* (PSO) menghasilkan akurasi sebesar 80.21% dan AUC 0.771 memiliki selisih akurasi 3.58%.

Sehingga didapatkan akurasi SVM (PSO) lebih tinggi dibandingkan lagoritma yang lainnya.

Deployment.

Pada tahap ini dilakukan perancangan dan pembuatan aplikasi dengan menggunakan data set sentiment *tweet* positif dan data *tweet* negatif mengenai operasi tangkap tangan KPK. Aplikasi ini berfungsi untuk memberikan bobot suatu kalimat apakah memuat sentiment negatif atau positif, dengan menggunakan *database* pembobotan kata-kata yang dihasilkan dari *Rapidminer*.

Berikut tampilan aplikasi pembobotan sentiment sebuah kalimat seperti dibawah ini:



Gambar 2. Masukkan Text Input.

Hasil Result Algoritma Svm

Kalimat Aslinya
Baru wali ini ada eksekutif saterbuka ini soal OTT @KPK_RI : Sekalian disema kementerian gitu pak Cahyo? https://t.co/bFza419cwW
Remove Annotation
baru wali ini ada eksekutif saterbuka ini soal ott : sekalian disema kementerian gitu pak cahyo https://t.co/bfza419cww _____
Remove URL
baru wali ini ada eksekutif saterbuka ini soal ott : sekalian disema kementerian gitu pak cahyo _____
Tokenize
baru wali ini ada eksekutif saterbuka ini soal ott sekalian disema kementerian gitu pak cahyo
Not Negative
baru wali ini ada eksekutif saterbuka ini soal ott sekalian disema kementerian gitu pak cahyo _____
Indonesian Stemming
baru wali ini ada eksekutif saterbuka ini soal ott sekalian semua menteri gitu pak cahyo _____
Indonesian Stop Word
baru wali eksekutif saterbuka ott menteri gitu cahyo _____

Gambar 12. Hasil Result Algoritma SVM



Gambar 13. Hasil Keseluruhan.

PENUTUP

Simpulan

Dari penelitian yang dilakukan mengenai Analisa Sentimen di Twitter mengenai operasi tangkap tangan Komisi Pemberantasan Korupsi (KPK) masing-masing dataset memiliki 78 data tweet positif dan 78 data tweet negative. Maka dapat disimpulkan untuk data *tweet* dengan metode *Particle Swarm Optimize* untuk *Support Vector Machine (PSO)* menghasilkan akurasi 83.79% dan AUC 0.910 adalah yang paling tinggi dibandingkan algoritma *Naïve Bayes*, *Naïve Bayes (PSO)*, *Support Vector Machine*. Dari hasil pengujian data *tweet* dengan penggunaan metode *Support Vector Machine* baik yang belum dioptimasi maupun yang sudah dioptimasi dengan *Particle Swarm Optimization* selalu menghasilkan akurasi yang lebih baik dari metode *Naïve Bayes*.

Saran

Dapat dilakukan penelitian lanjutan dengan menggunakan metode penentuan sentiment yaitu metode *Analysis Coarse-Grained* dan *Fined-Grained Sentiment* sebelum melakukan proses *preprocessing*. Dimana metode *Analysis Coarse-Grained* menganalisa seluruh isi dokumen dan *Fined-Grained* menganalisa pada level kalimat, untuk penentuan positif dan negatif.

DAFTAR PUSTAKA

- Priambada, Bintara Sura. 2008. Eksistensi KPK Dalam Memberantas Tindak Pidana Korupsi.
- Muttaqin, Labib dan Susanto, Muhammad Edy. 2018. Mengkaji Serangan Balik Koruptor Terhadap KPK dan Strategi Menghadapinya. INTEGRITAS. Volume 4 Nomor 1 - Juni 2018 , p-ISSN: 2477-118X e-ISSN: 2615-7977.
- Hikmawati, Puteri. 2018. Operasi Tangkap Tangan Dalam Penanganan Kasus Korupsi *Hand Arrest Operation In Handling Corruption Case*. NEGARA HUKUM: Vol. 9, No. 1, Juni 2018.
- Wardhani, Nia Kusuma., Rezkiani, Kurniawan, Sigit., Setiawan, Hendra., Gata, Grace., Tohari, Siswanto., Gata, Windu., Wahyudi, Mochamad. (2018). Sentiment Analysis Article News Coordinator Mainister Of Maritime Affairs Using Algorithm Naive Bayes And Support Vector Machine With Particle Swarm Optimization. JATIT & LLS .31s December 2018. Vol.96. No 24. E-ISSN: 1817-

3195.

Tunggawan, E. (2016). And the Winner is ...: Bayesian Twitter-based Prediction on 2016 U . S . Presidential Election, (1), 1–5.

Soepriadi, a., permata, m., informatika, j. T., & bayes, n. (2018). Sentiment analysis untuk menilai kepuasan masyarakat terhadap kinerja pemerintah daerah menggunakan naive bayes classifier (studi kasus : walikota, 4(1),1–7.

Yang, J., & Honavar, V. (1998). Feature subset selection using a genetic algorithm. *IEEE Intelligent Systems and Their Applications*, 13(2), 44–49.

Rohani, Abbas., et al., 2017. A novel soft computing model (Gaussian process regression with K-fold cross validation) for daily and monthly solar radiation forecasting (Part: I). *Renewable Energy*, 115, 411-422.