

ANALISIS SENTIMEN APLIKASI RUANG GURU DI TWITTER MENGUNAKAN ALGORITMA KLASIFIKASI

Angelina Puput Giovani¹⁾, Ardiansyah²⁾, Tuti Haryanti³⁾, Laela Kurniawati⁴⁾, Windu Gata⁵⁾

^{1,2,5}Ilmu Komputer, STMIK Nusa Mandiri

^{3,4}Sistem Informasi, STMIK Nusa Mandiri

^{1,2,3,4,5}Jl. Kramat Raya No. 18, Jakarta

Email: ¹14002338@nusamandiri.ac.id, ²14002404@nusamandiri.ac.id, ³tuti@nusamandiri.ac.id,
⁴laela@nusamandiri.ac.id, ⁵windu@nusamandiri.ac.id

Abstract

E-learning is electronic-based learning using computers or computer-based. One e-learning application that is widely known today is Ruang Guru. One way to find out the success of an application is to do a sentiment analysis of the application. In this study, sentiment analysis was taken from Twitter social media user comments on Ruang Guru of 513 tweets, after data cleaning, with 338 tweets of positive sentiment and 175 tweets of negative sentiment. The data was extracted using the Naive Bayes (NB) algorithm, Support Vector Machine (SVM), K-Nearest Neighbor (K-NN), and feature selection with the Particle Swarm Optimization (PSO) algorithm. This study compares the NB, SVM, K-NN methods without using feature selection with the NB, SVM, K-NN methods that use feature selection and compares the Area Under Curve (AUC) values of these methods to find the most optimal algorithm. The test results get the results that the best optimization application in this model is the SVO-based PSO algorithm with an accuracy value of 78.55% and AUC of 0.853. This research succeeded in getting the most effective and best algorithm in classifying positive and negative comments related to Ruang Guru.

Keyword: Ruang Guru, Sentiment analysis, Twitter, Classification Algorithms.

Abstrak

E-learning merupakan pembelajaran berbasis elektronik dengan menggunakan komputer atau berbasis komputer. Salah satu aplikasi e-learning yang banyak dikenal saat ini adalah aplikasi Ruang Guru. Salah satu cara untuk mengetahui keberhasilan suatu aplikasi adalah dengan melakukan analisis sentimen terhadap aplikasi tersebut. Pada penelitian ini, analisis sentimen diambil dari komentar pengguna media sosial Twitter terhadap aplikasi Ruang Guru sebanyak 513 tweet, setelah dilakukan data cleaning, dengan sentimen positif sebanyak 338 tweet dan sentimen negatif sebanyak 175 tweet. Data tersebut diekstraksi menggunakan algoritma Naive Bayes (NB), Support Vector Machine (SVM), K-Nearest Neighbour (K-NN), dan feature selection dengan algoritma Particle Swarm Optimization (PSO). Penelitian ini membandingkan metode NB, SVM, K-NN tanpa menggunakan feature selection dengan metode NB, SVM, K-NN yang menggunakan feature selection serta membandingkan nilai Area Under Curve (AUC) dari metode-metode tersebut untuk mengetahui algoritma yang paling optimal. Hasil pengujian mendapatkan hasil bahwa aplikasi optimasi terbaik dalam model ini adalah algoritma PSO berbasis SVM dengan nilai akurasi sebesar 78,55% dan AUC sebesar 0,853. Penelitian ini berhasil mendapatkan algoritma yang efektif dan terbaik dalam mengklasifikasikan komentar positif dan komentar negatif terkait dengan aplikasi Ruang Guru.

Kata Kunci: Ruang Guru, Analisis sentimen, Twitter, Algoritma Klasifikasi.

1. Pendahuluan

Semakin berkembangnya teknologi informasi membuat metode pembelajaran dituntut untuk mengikuti perkembangan teknologi. Metode pembelajaran yang digunakan seharusnya dapat memanfaatkan berbagai media untuk meningkatkan kualitas hasil pembelajaran. Salah satunya, media pembelajaran dapat menggunakan e-learning atau pembelajaran berbasis elektronik dengan menggunakan komputer atau berbasis komputer. E-learning pertama kali diperkenalkan oleh Universitas

Illinois di Urbana-Champaign dengan menggunakan sistem instruksi berbasis komputer (*computer assisted instruction*) dan komputer bernama PLATO. Sejak saat itu e-learning, pembelajaran berbasis internet, berkembang sejalan dengan perkembangan dan kemajuan teknologi [1]. Sampai saat ini aplikasi atau situs e-learning berkembang sangat pesat. Terdapat beberapa aplikasi e-learning di Indonesia, dari mulai yang berbayar hingga tak berbayar, salah satunya adalah aplikasi Ruang Guru.

Dikutip dari laman ruangguru.com, Ruang Guru merupakan perusahaan teknologi terbesar di Indonesia

yang berfokus pada layanan berbasis pendidikan, yang telah memiliki lebih dari 15 juta pengguna serta mengelola 300.000 guru yang menawarkan jasa di lebih dari 100 bidang pelajaran. Ruang Guru mengembangkan berbagai layanan belajar berbasis teknologi, termasuk layanan kelas virtual, platform ujian online, video belajar berlangganan, marketplace, les privat, serta konten-konten pendidikan lainnya yang bisa diakses melalui web dan aplikasi Ruang Guru. Aplikasi ini memudahkan siswa untuk mengakses materi pembelajaran yang berkualitas. Di mana pun mereka berada, dapat membantu proses belajar siswa tanpa batasan ruang dan waktu, dapat membantu siswa, guru, dan orang tua untuk menjalankan aktivitasnya menjadi lebih efektif dan efisien [2].

Suatu aplikasi selalu memiliki kekurangan dan kelebihan masing-masing, dimana hal tersebut dapat menimbulkan berbagai respon dari pengguna aplikasi seperti kepuasan dan kekecewaan terhadap aplikasi tersebut. Media sosial menjadi salah satu tempat untuk melontarkan kepuasan dan kekecewaan pengguna atau opini terhadap aplikasi tersebut. Hal tersebut dapat dijadikan bahan analisis sentimen terhadap aplikasi Ruang Guru.

Analisis Sentimen digunakan untuk menemukan informasi berharga yang dibutuhkan dari data yang tidak terstruktur, sehingga diharapkan pada penelitian ini dapat diketahui sentimen pengguna twitter terhadap aplikasi Ruang Guru. Penelitian ini juga diharapkan dapat mengetahui pengaruh *feature selection* terhadap akurasi proses analisis sentimen yang dilakukan.

Twitter merupakan salah satu media sosial yang sangat populer dikalangan pengguna internet, hal ini dikarenakan kesederhanaan dan kemudahan dalam penggunaannya, serta pengguna dapat dengan bebas mengeluarkan pendapat atau opini mereka. Beberapa peneliti yang menggunakan bahan penelitian dari twitter antara lain terkait teks berbahasa Indonesia, calon gubernur DKI Jakarta 2017, calon Presiden Indonesia 2014, dan penggunaan transportasi umum darat dalam kota [3]–[6]. Beberapa peneliti sebelumnya telah melakukan penelitian untuk mengukur sentimen analisis dengan berbagai algoritma, antara lain algoritma *Naive Bayes* (NB), *Support Vector Machine* (SVM), *K-Nearest Neighbour* (K-NN), *Decision Tree*, *K-Medoid*, dan *Backpropagation Neural Network* (BNN). Beberapa penelitian menambahkan *feature selection* seperti *Particle Swarm Optimization* (PSO). Diantara penelitian tersebut terdapat beberapa peneliti yang membandingkan dua atau lebih algoritma untuk mengetahui algoritma terbaik dari algoritma yang dibandingkan [7]–[15].

Metode yang digunakan dalam penelitian ini adalah NB, SVM, dan K-NN, dengan tambahan *feature selection* dengan PSO. Penelitian akan membandingkan metode NB, SVM, K-NN tanpa PSO dibandingkan dengan metode NB, SVM, K-NN dengan PSO. Hal ini untuk menguji pengaruh penggunaan *feature selection* PSO dengan membandingkan nilai AUC (*Area Under Curve*) dari metode tersebut.

Algoritma klasifikasi NB merupakan salah satu algoritma teknik klasifikasi, algoritma ini memanfaatkan

metode probabilitas dan statistik. Algoritma ini merupakan salah satu metode klasifikasi yang populer dan termasuk dalam sepuluh algoritma teratas dalam *data mining* menurut *IEEE International Conference on Data Mining* (ICDM'06) di Hongkong [16]. *Naive Bayes* merupakan cabang matematika yang dikenal sebagai teori probabilitas untuk mencari peluang terbesar dari kemungkinan klasifikasi, dengan melihat frekuensi masing-masing klasifikasi dalam data pelatihan [17]. Pada algoritma klasifikasi NB terdapat dua tahapan dalam proses klasifikasi teks. Pertama, proses analisis sampel dokumen seperti penyeleksian kosa kata, yaitu kosa kata yang mungkin muncul dalam koleksi dokumen sampel, yang kemudian dilakukan proses penentuan probabilitas untuk setiap kategori berdasarkan dokumen sampel. Kedua, tahap klasifikasi, dilakukan penentuan nilai kategori suatu dokumen berdasarkan kosa kata yang muncul dalam dokumen yang diklasifikasi.

SVM merupakan seperangkat metode pembelajaran yang menganalisis data dan mengenali pola. Algoritma ini dapat digunakan untuk klasifikasi dan analisis regresi. Namun tidak hanya itu, algoritma ini juga dapat melakukan prediksi dan penilaian tentang sebuah sistem [18]. Tujuan dari SVM adalah memberikan nilai dari banyaknya kemunculan sebuah kata dan dapat mengklasifikasi kalimat dengan label positif dan negatif. Keuntungan SVM yaitu Ruang Input Dimensi Tinggi dan Ruang Dokumen Vektor [19].

K-NN adalah salah satu algoritma paling sederhana untuk memecahkan masalah klasifikasi. Algoritma ini sering digunakan untuk klasifikasi teks dan data. Pada metode ini dilakukan klasifikasi terhadap obyek berdasarkan data yang jaraknya paling dekat dengan obyek tersebut [20]. Tujuan dari algoritma ini adalah untuk mengklasifikasikan obyek berdasarkan atribut dan *training sample*. Algoritma ini menggunakan klasifikasi ketetanggaaan sebagai nilai prediksi dari *query instance* yang baru.

PSO banyak digunakan untuk memecahkan masalah optimasi serta masalah seleksi fitur (*feature selection*). Dalam teknik PSO terdapat beberapa cara untuk melakukan pengoptimasian, diantaranya meningkatkan bobot atribut (*attribute weight*) terhadap semua atribut atau variabel yang dipakai, menyeleksi atribut (*attribute selection*) dan *feature selection*. Algoritma ini adalah suatu teknik optimasi yang sangat sederhana untuk menerapkan dan memodifikasi beberapa parameter [21].

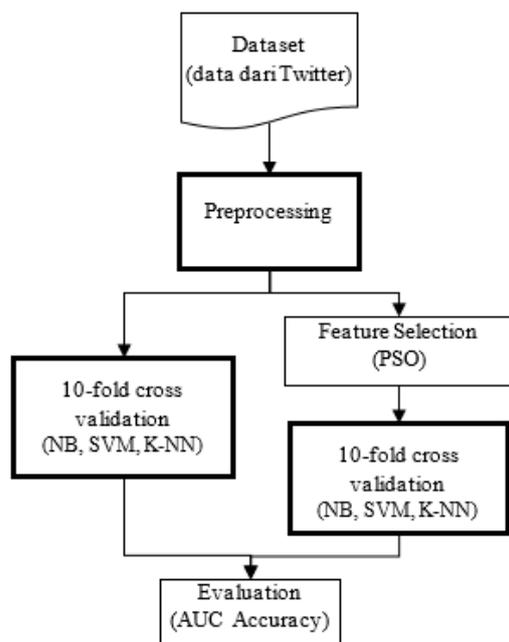
Beberapa penelitian sebelumnya telah membandingkan beberapa algoritma klasifikasi, seperti membandingkan algoritma SVM dengan algoritma NB yang dilakukan pada penelitian dengan judul *Sentiment Analysis Article News Coordinator Minister Of Maritime Affairs Using Algorithm Naive Bayes And Support Vector Machine With Particle Swarm Optimization*, yang membandingkan algoritma NB, NB (PSO), SVM dan SVM (PSO) mendapatkan hasil akurasi NB (PSO) memiliki akurasi yang lebih tinggi dibandingkan dengan SVM, SVM (PSO) dan NB, dan optimisasi terbaik adalah PSO berbasis NB [17].

Penelitian lain yang dilakukan dengan judul *Sentimen Analisis Operasi Tangkap Tangan KPK Menurut Masyarakat Menggunakan Algoritma Support Vector Machine, Naive Bayes Berbasis Particle Swarm Optimization*, menyimpulkan bahwa metode PSO untuk SVM (PSO) menghasilkan akurasi 83.79% dan AUC 0.910 adalah yang paling tinggi dibandingkan algoritma NB, NB (PSO), dan SVM [22].

Dengan adanya perbedaan beberapa penelitian sebelumnya, mendorong penulis untuk melakukan penelitian dengan membandingkan beberapa metode klasifikasi dengan kasus yang berbeda. Di dalam penelitian ini, akan dibahas tahapan yang dilalui untuk melakukan proses analisis sentimen terhadap komentar tentang aplikasi Ruang Guru pada media sosial Twitter. Dimulai dari tahap *preprocessing* sampai tahap analisis sentimen dengan NB, SVM, K-NN, dan *feature selection* PSO serta mengukur kualitas hasil analisis algoritma tersebut tanpa *feature selection* dan menggunakan *feature selection*.

2. Bahan dan Metode

Penelitian ini menggunakan data yang bersumber dari komentar berbahasa Indonesia pada media sosial Twitter mengenai aplikasi Ruang Guru. Metode pada penelitian ini diadopsi dari model *Cross-Industry Standard Process for Data Mining* (CRISP-DM). Kerangka berpikir penelitian ini sebagaimana dapat dilihat pada gambar 1 berikut.



Gambar 1. Kerangka berpikir dalam penelitian

Langkah-langkah kerangka berpikir yang diadopsi dari *Cross-Industry Standard Process for Data Mining* (CRISP-DM) adalah sebagai berikut [23]:

1) Business Understanding

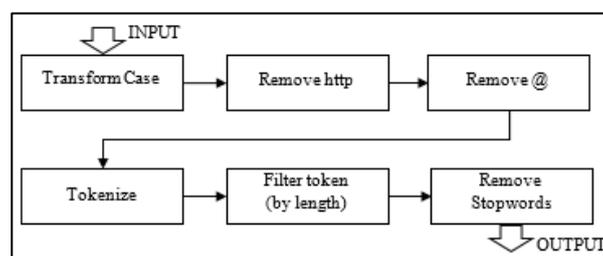
Proses memahami objek penelitian yang akan dilakukan untuk menghasilkan hasil sesuai dengan yang diharapkan. Pemahaman terhadap objek penelitian dapat dilakukan dengan mencari mencari dan menggali informasi dari sumber objek.

2) Data Understanding

Proses yang dilakukan pada tahap ini yaitu pengambilan data mentah yang dilakukan sesuai objek penelitian dengan atribut yang diperlukan Data di peroleh dari media sosial twitter dari tanggal 02 April 2020 hingga 03 April 2020 dengan dengan kata kunci “Ruang Guru” dan *filter* bahasa yang digunakan adalah Bahasa Indonesia. Setelah data diperoleh, kemudian dilakukan proses pembersihan data (*data cleaning*). Data yang diperoleh dari hasil *data cleaning* tersebut akan digunakan sebagai bahan penelitian.

3) Data Preparation (Preprocessing)

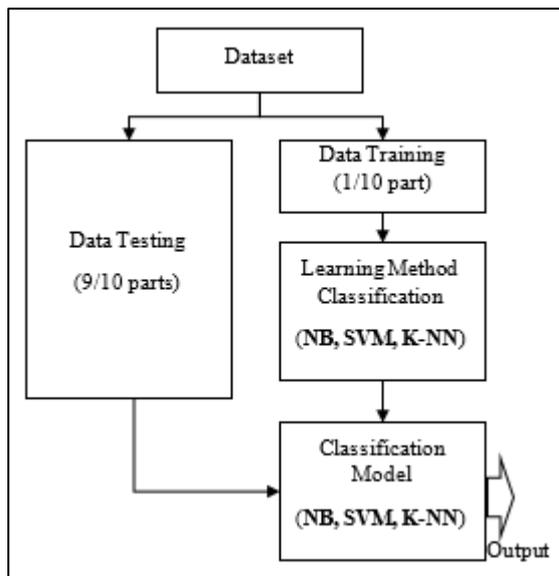
Data preparation dilakukan sebelum *dataset* dimasukkan kedalam model. Proses tersebut meliputi beberapa tahap, yaitu: (1) *transform case* untuk mengubah huruf pada *text* menjadi huruf kecil semua guna menyeragamkan bentuk huruf, (2) *remove http* untuk menghapus tautan yang terdapat di dalam tweet, (3) *remove @* untuk menghapus *mention* di dalam tweet, (4) *tokenize* untuk memisahkan kata-kata pada setiap kalimat menjadi kata tersendiri sekaligus menghapus karakter yang tidak diperlukan, (5) *Filter tokens by length* untuk menghapus kata dengan jumlah karakter yang kurang dari nilai yang ditentukan, dalam hal ini 4 karakter, dan (6) *remove stopwords* untuk menghapus kata-kata yang dapat diabaikan, biasanya berupa kata keterangan dan kata sambung, dalam hal ini digunakan kamus *stopwords* Bahasa Indonesia. Alur kerja *data preparation* dapat dilihat pada gambar 2 berikut.



Gambar 2. Alur proses Data Preparation (Preprocessing)

4) Modelling

Proses ini dilakukan menggunakan metode *10-fold cross validation*, yaitu metode membagi *dataset* menjadi 10 bagian yang mana 1 bagian diantaranya menjadi data uji (*testing*) sedangkan bagian lainnya menjadi data latih (*training*). Kemudian data tersebut dimasukkan ke dalam model algoritma yang diuji yaitu NB, SVM, K-NN. Hal ini dilakukan bergantian pada setiap bagian data hingga didapatkan nilai terbaik dari model tersebut. Alur proses *10-fold cross validation* dapat dilihat pada gambar 3 berikut.



Gambar 3. Alur proses 10-fold cross validation

5) Evaluation

Tahap ini melakukan proses evaluasi terhadap hasil dari pemodelan dengan membandingkan hasil dari pemodelan yang berupa *accuracy*, *precision*, *recall* dan AUC antara model yang menggunakan *feature selection* PSO dengan model yang tidak menggunakan *feature selection* PSO. Tingkat akurasi dan *ROC Curve* dari masing-masing algoritma diklasifikasikan berdasarkan nilai AUC sebagai berikut: 0,90-1,00 = sangat baik. 0,80-0,90 = baik. 0,70-0,80 = cukup. 0,60-0,70 = buruk. 0,50-0,60 = gagal [17].

3. Hasil dan Pembahasan

Langkah-langkah yang diterapkan dalam penelitian ini terdiri dari beberapa proses, antara lain:

1) Business Understanding

Pada penelitian ini sumber informasi yang digunakan yaitu media sosial Twitter. Penelitian dilakukan untuk menggali informasi mengenai analisis sentimen mengenai aplikasi Ruang Guru berdasarkan data Twitter tersebut.

2) Data Understanding

Pengambilan data Twitter dilakukan menggunakan aplikasi *RapidMiner Studio Educational 9.6* dengan *query* “Ruang Guru” dan parameter-parameter sesuai dengan yang sudah disebutkan di dalam metodologi, kemudian data tersebut di-*export* ke dalam format *file* Microsoft Excel. Data yang diperoleh sebanyak 568 tweet. Data tersebut kemudian dilakukan pembersihan (*data cleaning*) untuk menghilangkan duplikasi data. Data bersih yang diperoleh dari hasil *data cleaning* sebanyak 513 tweet, yang terdiri dari 338 tweet positif dan 175 tweet negatif.

3) Data Preparation (Preprocessing)

Proses *data preparation* dilakukan terhadap data 513 tweet tersebut dalam 6 tahapan, sebagaimana disebutkan dalam metodologi, menggunakan

RapidMiner Studio Educational 9.6 dengan contoh hasil dapat dilihat pada tabel 1 sampai dengan tabel 6 berikut.

Tabel 1. Hasil dari Transform Case

Text	Transform Case
@octania_: Udah pake dunks dari DULU sebelum ada wabah²an' Bagus kook .. Ruang Guru jg bagus, ga NJELIMET juga meski kadang buat KIDDONYA aku agak susah paham bahasanya ??	@octania_: udah pake dunks dari dulu sebelum ada wabah²an' bagus kook .. ruang guru jg bagus, ga njelimet juga meski kadang buat kiddonya aku agak susah paham bahasanya ??
REKOMEN BANGET BUAT YANG PUNYA ANAK² BELAJAR DIRUMAH ????	rekomen banget buat yang punya anak² belajar dirumah ????
https://t.co/bjwpglcfa4	https://t.co/bjwpglcfa4

Tabel 2. Hasil dari Remove HTTP

Text	Hasil Remove HTTP
@octania_: udah pake dunks dari dulu sebelum ada wabah²an' bagus kook .. ruang guru jg bagus, ga njelimet juga meski kadang buat kiddonya aku agak susah paham bahasanya ??	@octania_: udah pake dunks dari dulu sebelum ada wabah²an' bagus kook .. ruang guru jg bagus, ga njelimet juga meski kadang buat kiddonya aku agak susah paham bahasanya ??
rekomen banget buat yang punya anak² belajar dirumah ????	rekomen banget buat yang punya anak² belajar dirumah ????
https://t.co/bjwpglcfa4	https://t.co/bjwpglcfa4

Tabel 3. Hasil dari Remove @

Text	Hasil Remove @
@octania_: udah pake dunks dari dulu sebelum ada wabah²an' bagus kook .. ruang guru jg bagus, ga njelimet juga meski kadang buat kiddonya aku agak susah paham bahasanya ??	udah pake dunks dari dulu sebelum ada wabah²an' bagus kook .. ruang guru jg bagus, ga njelimet juga meski kadang buat kiddonya aku agak susah paham bahasanya ??
rekomen banget buat yang punya anak² belajar dirumah ????	rekomen banget buat yang punya anak² belajar dirumah ????

Tabel 4. Hasil dari Tokenize

Text	Hasil Tokenize
udah pake dunks dari dulu sebelum ada wabah²an' bagus kook .. ruang guru jg bagus, ga	udah pake dunks dari

njelimet juga meski
kadang buat kiddonya
aku agak susah paham
bahasanya ??
rekomendasi banget buat
yang punya anak² belajar
dirumah ????

dulu
sebelum
wabah
an
bagus
kook
ruang
guru
jg
ga
njelimet
juga
meski
kadang
buat
kiddonya
aku
agak
susah
paham
bahasanya
rekomendasi
banget
yang
punya
anak
belajar
dirumah

Tabel 5. Hasil dari *Filter by Length*

Text	Hasil Filter by Length
udah	udah
pake	pake
dunks	dunks
dari	dari
dulu	dulu
sebelum	sebelum
ada	
wabah	wabah
an	kook
bagus	ruang
kook	guru
ruang	njelimet
guru	juga
jg	meski
bagus	kadang
ga	buat
njelimet	kiddonya
juga	agak
meski	susah
kadang	paham
buat	bahasanya
kiddonya	rekomendasi
aku	banget
agak	yang
susah	punya
paham	
bahasanya	
rekomendasi	

banget
buat
yang
punya
anak
belajar
dirumah

Tabel 6. Hasil dari *Remove Stopwords*

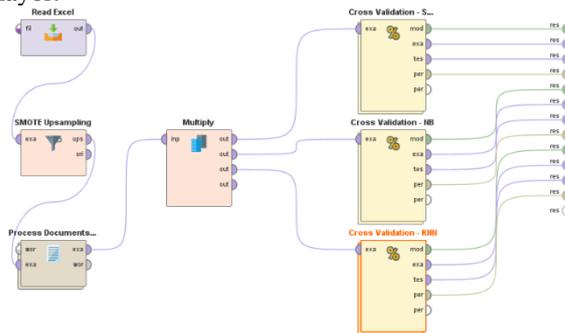
Text	Hasil Remove Stopwords
udah	udah
pake	pake
dunks	dunks
dari	wabah
dulu	bagus
sebelum	kook
wabah	ruang
bagus	guru
kook	njelimet
ruang	kadang
guru	kiddonya
bagus	susah
njelimet	paham
juga	bahasanya
meski	rekomendasi
kadang	banget
buat	anak
kiddonya	belajar
agak	dirumah
susah	
paham	
bahasanya	
rekomendasi	
banget	
buat	
yang	
punya	
anak	
belajar	
dirumah	

4) *Modelling*

Pada tahap ini melibatkan teknik *data mining* dengan algoritma klasifikasi menggunakan *RapidMiner Studio Educational 9.6*. Hasil pengujian model yang dilakukan adalah mengklasifikasi tweet positif dan tweet negatif menggunakan algoritma NB, SVM, KNN dan *feature selection* PSO.

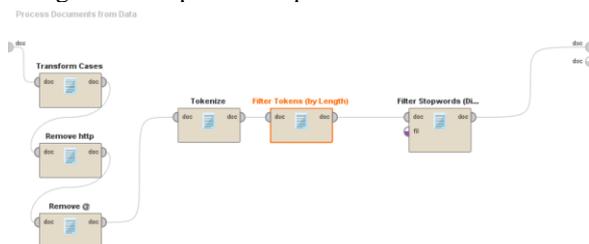
Pertama, dilakukan perbandingan algoritma algoritma NB, SVM, dan KNN tanpa *feature selection* PSO. Desain model tersebut pada aplikasi yang digunakan dapat dilihat pada Gambar 4. Pada gambar tersebut dapat dijelaskan bahwa bahwa data yang berhasil diperoleh telah disimpan dalam sebuah *file Microsoft Excel* yang akan dibaca oleh *operator Read Excel*. Sebelum diproses lebih lanjut, terhadap data tersebut dilakukan operasi *SMOTE Upsampling* untuk menyeimbangkan kelas data dengan menambah

jumlah data kelas minor agar setara dengan kelas mayor.



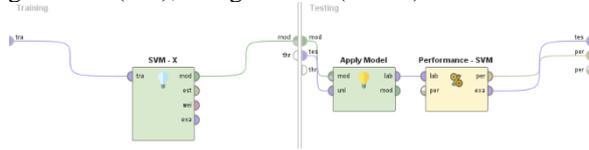
Gambar 4. Desain model perbandingan NB, SVM, dan K-NN

Setelah data diseimbangkan, selanjutnya dilakukan operasi *Process Document*. Proses ini adalah tahap *data preparation* yang di dalamnya terdiri dari *operator transform cases, remove http, remove @, tokenize, filter tokens by length, dan Stopword Filter*, sebagaimana dapat dilihat pada Gambar 5 berikut.

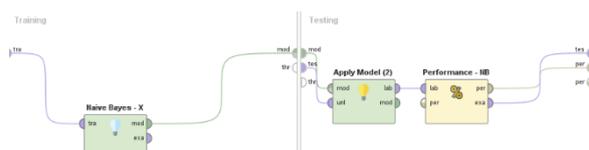


Gambar 5. Proses dokumen

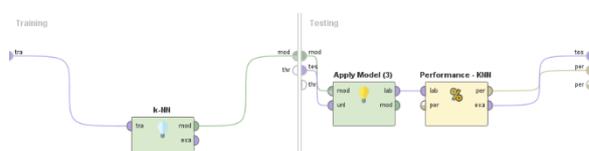
Selanjutnya, hasil proses tersebut digandakan menggunakan *operator Multiply* agar dapat digunakan dalam 3 algoritma pemodelan. Masing-masing pemodelan dilakukan menggunakan metode *10 fold cross validation* dengan *operator Cross Validation* sebagaimana dapat dilihat pada gambar 6 (SVM), gambar 7 (NB), dan gambar 8 (K-NN).



Gambar 6. Proses didalam Cross Validation-SVM

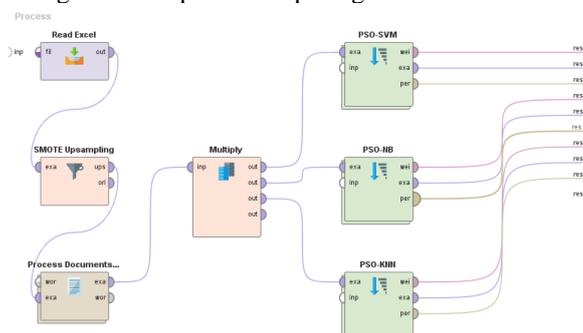


Gambar 7. Proses didalam Cross Validation-NB



Gambar 8. Proses didalam Cross Validation-K-NN

Berikutnya, dilakukan perbandingan algoritma NB, SVM, K-NN yang menggunakan *feature selection* PSO. Proses diawali dengan membaca data dari file dengan *operator read Excel* hingga penggandaan data dengan *operator Multiply*, memiliki proses yang sama seperti model desain sebelumnya. Sementara *operator* PSO ditempatkan setelah *operator Multiply* sebagaimana dapat dilihat pada gambar 9 berikut.



Gambar 9. Desain Model Perbandingan NB (PSO), SVM (PSO), dan K-NN (PSO)

Setelah dilakukan *feature selection* PSO, langkah selanjutnya adalah menjalankan masing-masing algoritma pemodelan menggunakan metode *10 fold cross validation* dengan *operator Cross Validation* sebagaimana proses sebelumnya pada gambar 6, gambar 7, dan gambar 8.

5) Evaluation

Tahap evaluasi bertujuan untuk menentukan nilai kegunaan model yang telah berhasil dibuat pada langkah sebelumnya. Metode evaluasi yang digunakan dalam penelitian ini adalah metode *10-fold cross validation*. Tingkat akurasi dan *ROC Curve* dari masing-masing algoritma diklasifikasikan berdasarkan nilai AUC sebagai berikut: 0,90-1,00 = sangat baik. 0,80-0,90 = baik. 0,70-0,80 = cukup. 0,60-0,70 = buruk. 0,50-0,60 = gagal.

Hasil pengujian akurasi dari algoritma NB, SVM, K-NN tanpa menggunakan *feature selection* PSO dapat dilihat pada gambar 10 sampai dengan gambar 12 berikut.

accuracy: 65.40% +/- 4.83% (micro average: 65.38%)

	true Positif	true Negatif	class precision
pred. Positif	157	53	74.76%
pred. Negatif	181	285	61.16%
class recall	46.45%	84.32%	

Gambar 10. Tingkat accuracy algoritma NB

accuracy: 76.93% +/- 5.63% (micro average: 76.92%)

	true Positif	true Negatif	class precision
pred. Positif	296	114	72.20%
pred. Negatif	42	224	84.21%
class recall	87.57%	66.27%	

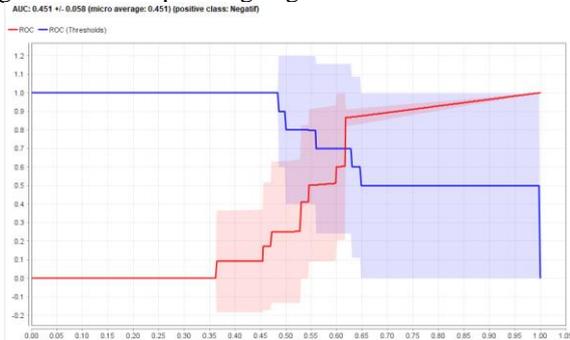
Gambar 11. Tingkat accuracy algoritma SVM

accuracy: 70.42% +/- 5.65% (micro average: 70.41%)

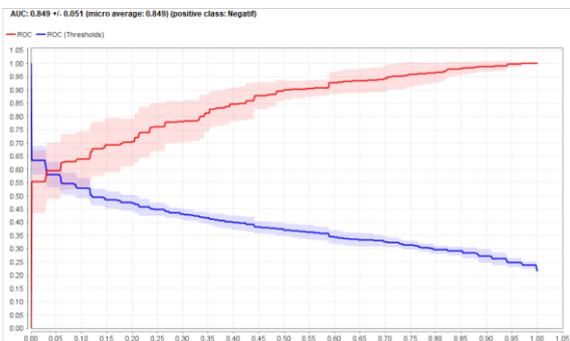
	true Positif	true Negatif	class precision
pred Positif	274	136	66.83%
pred Negatif	64	202	75.94%
class recall	81.07%	59.76%	

Gambar 12. Tingkat *accuracy* algoritma KNN

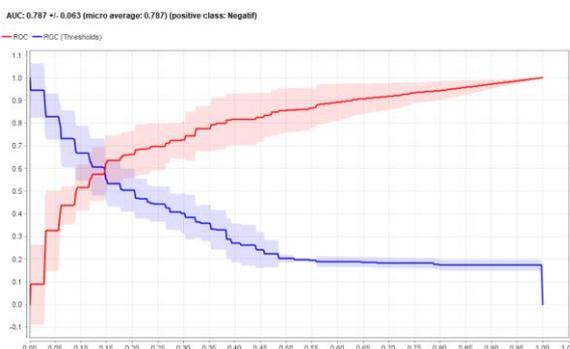
Hasil pengujian performa dalam bentuk *ROC Curve* dari algoritma NB, SVM, K-NN tanpa menggunakan *feature selection* PSO dapat dilihat pada gambar 13 sampai dengan gambar 15 berikut.



Gambar 13. *ROC Curve* algoritma NB



Gambar 14. *ROC Curve* algoritma SVM



Gambar 15. *ROC Curve* algoritma KNN

Dengan demikian, hasil akurasi dan nilai AUC dari algoritma NB, SVM, K-NN tanpa menggunakan *feature selection* PSO dapat dilihat pada tabel 7 berikut.

Tabel 7. Akurasi dan AUC algoritma NB, SVM, dan K-NN tanpa PSO

Algoritma	Accuracy	AUC	Klasifikasi
SVM	76,93%	0,849	Baik
NB	65,40%	0,451	Gagal
K-NN	70,42%	0,787	Cukup

Berdasarkan hasil tersebut, dapat diketahui bahwa tanpa menggunakan *feature selection* PSO, algoritma SVM memiliki performa yang terbaik jika dibandingkan algoritma NB dan K-NN dalam penelitian ini.

Hasil pengujian akurasi dari algoritma NB, SVM, K-NN dengan menggunakan *feature selection* PSO dapat dilihat pada gambar 16 sampai dengan gambar 18 berikut.

accuracy: 67.32% +/- 5.06% (micro average: 67.31%)

	true Positif	true Negatif	class precision
pred Positif	163	46	77.99%
pred Negatif	175	292	62.53%
class recall	48.22%	86.39%	

Gambar 16. Tingkat *accuracy* algoritma NB (PSO)

accuracy: 78.55% +/- 3.78% (micro average: 78.55%)

	true Positif	true Negatif	class precision
pred Positif	291	98	74.81%
pred Negatif	47	240	83.62%
class recall	86.09%	71.01%	

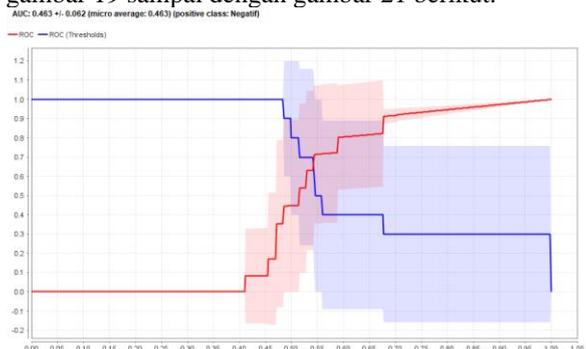
Gambar 17. Tingkat *accuracy* algoritma SVM (PSO)

accuracy: 77.21% +/- 4.59% (micro average: 77.22%)

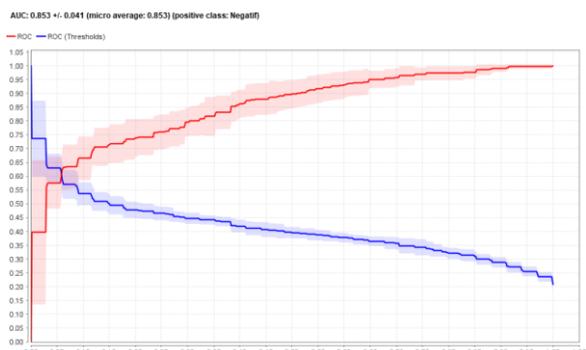
	true Positif	true Negatif	class precision
pred Positif	318	134	70.35%
pred Negatif	20	204	91.07%
class recall	94.08%	60.36%	

Gambar 18. Tingkat *accuracy* algoritma KNN (PSO)

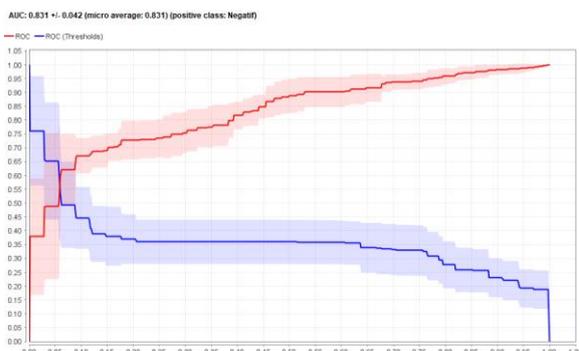
Hasil pengujian performa dalam bentuk *ROC Curve* dari algoritma NB, SVM, K-NN tanpa menggunakan *feature selection* PSO dapat dilihat pada gambar 19 sampai dengan gambar 21 berikut.



Gambar 19. *ROC Curve* algoritma NB (PSO)



Gambar 20. *ROC Curve* algoritma SVM (PSO)



Gambar 21. ROC Curve algoritma KNN (PSO)

Dengan demikian, hasil akurasi dan nilai AUC dari algoritma NB, SVM, K-NN dengan menggunakan *feature selection* PSO dapat dilihat pada tabel 8 berikut.

Tabel 8. Akurasi dan AUC algoritma NB, SVM, dan K-NN menggunakan PSO

Algoritma	Accuracy	AUC	Klasifikasi
SVM (PSO)	78,55%	0,853	Baik
NB (PSO)	67,32%	0,463	Gagal
K-NN (PSO)	77,21%	0,831	Baik

Berdasarkan hasil tersebut, dapat diketahui bahwa jika menggunakan *feature selection* PSO, maka algoritma SVM memiliki performa yang terbaik jika dibandingkan algoritma NB dan K-NN dalam penelitian ini.

4. Kesimpulan

Berdasarkan penelitian yang telah dilakukan maka dapat disimpulkan bahwa penggunaan *feature selection* PSO dalam algoritma klasifikasi dapat meningkatkan performa dan akurasi. Hasil dari pengujian secara keseluruhan algoritma SVM (PSO) memiliki nilai akurasi dan performa yang paling tinggi jika dibandingkan dengan NB, SVM, K-NN, NB (PSO), SVM (PSO), dan K-NN (PSO). Sedangkan algoritma NB, tanpa PSO maupun dengan PSO, mendapatkan nilai UAC dengan klasifikasi gagal. Dalam penelitian ini sumber data yang digunakan hanya dari satu sumber media sosial, yaitu Twitter. Hal ini dapat menjadi rujukan dalam penelitian berikutnya dengan menggunakan sumber data yang berbeda.

Daftar Pustaka

[1] M. A. Bora, “Analisa Kepuasan Penggunaan E-Learning Cloud Sekolah Tinggi Teknik (STT) Ibnu Sina Batam,” vol. 1, no. 1, pp. 55–62, 2017.

[2] RuangGuru, “Tentang Ruangguru.” <https://ruangguru.com/general/about> (accessed May 08, 2020).

[3] N. D. Putranti and E. Winarko, “Analisis Sentimen Twitter untuk Teks Berbahasa Indonesia dengan Maximum Entropy dan Support Vector Machine,” vol. 8, no. 1, pp. 91–100, 2014.

[4] G. A. Buntoro, “Analisis Sentimen Calon

Gubernur DKI Jakarta 2017 Di Twitter,” vol. 2, no. 1, pp. 32–41, 2017.

[5] F. Nurhuda and S. W. Sihwi, “Analisis Sentimen Masyarakat terhadap Calon Presiden Indonesia 2014 berdasarkan Opini dari Twitter Menggunakan Metode Naive Bayes Classifier,” vol. 2, no. 2, 2014.

[6] A. Novantirani *et al.*, “Analisis Sentimen pada Twitter untuk Mengenai Penggunaan Transportasi Umum Darat Dalam Kota dengan Metode Support Vector Machine,” vol. 2, no. 1, pp. 1177–1183, 2015.

[7] S. A. Saputra, D. Rosiyadi, W. Gata, and S. M. Husain, “Analisis Sentimen E-Wallet Pada Google Play Menggunakan Algoritma Naive Bayes Berbasis Particle Swarm Optimization,” *J. RESTI*, vol. 1, no. 10, pp. 3–8, 2021.

[8] D. D. Saputra *et al.*, “Optimization Sentiments of Analysis from Tweets in myXLCare using Naive Bayes Algorithm and Synthetic Minority Over Sampling Technique Method,” *J. Phys. Conf. Ser.*, 2020, doi: 10.1088/1742-6596/1471/1/012014.

[9] D. A. Kristiyanti, A. H. Umam, M. Wahyudi, R. Amin, and L. Marlinda, “Comparison of SVM & Naive Bayes Algorithm for Sentiment Analysis Toward West Java Governor Candidate Period 2018-2023 Based on Public Opinion on Twitter,” *Int. Conf. Cyber IT Serv. Manag.*, no. Citsm, pp. 1–6, 2018, doi: 10.1109/CITSM.2018.8674352.

[10] L. N. Pradany and C. Faticah, “Analisa Sentimen Kebijakan Pemerintah Pada Konten Twitter Berbahasa Indonesia Menggunakan Svm Dan K-Medoid Clustering,” *SCAN-Jurnal Teknol. Inf. dan Komun.*, vol. 11, no. 1, pp. 59–66, 2016.

[11] N. Y. A. Faradhillah, R. P. Kusumawardani, and I. Hafidz, “Eksperimen Sistem Klasifikasi Analisa Sentimen Twitter pada Akun Resmi Pemerintah Kota Surabaya Berbasis Pembelajaran Mesin,” *Pros. Semin. Nas. Sist. Inf. Indones. 2016*, pp. 15–24, 2016.

[12] M. W. & N. N. Dwi Arum Ningtyas, “Klasifikasi Siswa Smk Berpotensi Putus Sekolah Menggunakan Algoritma Decision Tree , Support Vector Machine Dan Naive Bayes,” *J. Khatulistiwa Inform.*, vol. VII, no. 2, pp. 85–90, 2019.

[13] O. S. S. S. E. Syahfitri Kartika Lidya, “Sentiment Analysis Pada Teks Bahasa Indonesia Menggunakan Support Vector Machine (SVM) Dan K-Nearest Neighbor (K-NN),” *Semin. Nas. Teknol. Inf. dan Komun.*, vol. 2015, no. Sentika, pp. 1–8, 2015.

[14] S. W. Yudha and M. Wahyudi, “Komparasi Algoritma Klasifikasi Untuk Analisis Sentimen Review Film Berbahasa Asing,” in *Seminar Nasional Informatika, Sistem Informasi Dan Keamanan Siber (SEINASI-KESI)*, 2018, pp. 180–185.

[15] M. A. Assuja and S. Saniati, “Analisis Sentimen Tweet Menggunakan Backpropagation Neural

- Network,” *J. Teknoinfo*, vol. 10, no. 2, p. 48, 2016, doi: 10.33365/jti.v10i2.20.
- [16] X. Wu *et al.*, *Top 10 algorithms in data mining*, vol. 14, no. 1. 2008.
- [17] N. K. Wardhani *et al.*, “Sentiment analysis article news coordinator minister of maritime affairs using algorithm naive bayes and support vector machine with particle swarm optimization,” *J. Theor. Appl. Inf. Technol.*, vol. 96, no. 24, pp. 8365–8378, 2018.
- [18] M. A. Maulana, A. Setyanto, and M. P. Kurniawan, “Analisis Sentimen Media Sosial Universitas Amikom,” *Semin. Nas. Teknol. Inf. dan Multimed. 2018 Univ. AMIKOM Yogyakarta, 10 Februari 2018*, pp. 7–12, 2018.
- [19] E. Indrayuni, “Analisa Sentimen Review Hotel Menggunakan Algoritma Support Vector Machine Berbasis Particle Swarm Optimization Elly,” *J. Evolusi Vol.*, vol. 4, 2016.
- [20] W. E. Nurjanah, R. S. Perdana, and M. A. Fauzi, “Analisis Sentimen Terhadap Tayangan Televisi Berdasarkan Opini Masyarakat pada Media Sosial Twitter menggunakan Metode K-Nearest Neighbor dan Pembobotan Jumlah Retweet,” vol. 1, no. 12, pp. 1750–1757, 2017.
- [21] L. A. Utami, “Melalui Komparasi Algoritma Support Vector Machine Dan K-Nearest Neighbor Berbasis Particle Swarm Optimization,” vol. 13, no. 1, pp. 103–112, 2017.
- [22] Hernawati and W. Gata, “Sentimen Analisis Operasi Tangkap Tangan KPK Menurut Masyarakat Menggunakan Algoritma Support Vector Machine , Naive Bayes Berbasis Particle Swarm Optimization,” *faktir Exacta*, vol. 12, no. 3, pp. 230–243, 2019, doi: 10.30998/faktorexacta.v12i3.4992.
- [23] I. Santoso, W. Gata, and A. B. Paryanti, “Penggunaan Feature Selection di Algoritma Support Vector Machine untuk Sentimen Analisis Komisi Pemilihan Umum,” *J. RESTI*, vol. 3, no. 3, pp. 5–11, 2019.