Akreditasi Ristekdikti,
No: 30/E/KPT/2019 (Sinta 4)
DOI: https://doi.org/10.31294/paradigma.v24i1.968

Paradigma,
Vol. 24, No. 1 Maret 2022, Hal. 67-74
P-ISSN 1410-5063, E-ISSN: 2579-3500

Perbandingan Algoritma Dengan Particle Swarm Optimization Untuk Analisis Sentimen Pada Peraturan PSBB di Indonesia

Mugi Raharjo[1], Jordy Lasmana Putra[2], Tommi Alfian Armawan Sandi[3], Musriatun Napiah[4]

[1,2] Universitas Nusa Mandiri
Mugi.mou@nusamandiri.ac.id
Jordy.jlp@nusamandiri.ac.id

[3,4] Universitas Bina Sarana Informatika
Tommi.taf@bsi.ac.id
Musriatun.mph@bsi.ac.id

Abstrak - Pandemi telah memunculkan aturan dan istilah baru di masyarakat. Berbagai negara memiliki regulasinya masing-masing, termasuk Indonesia dengan nama PSBB untuk itu penulis mencoba melakukan penelitian terkait dengan kondisi PSBB di Indonesia dengan maksud dan tujuan untuk mengetahui sentimen masyarakat terhadap hal tersebut, penulis melakukan pemodelan ini secara positif dan negatif. . model dalam tweet di *Twitter*. Kami menangkap informasi melalui media *Twitter* yang kemudian kami olah datanya sehingga siap untuk diuji pada algoritma yang digunakan. Dalam pengumpulan dan pemrosesan data, kami menggunakan aplikasi penambang cepat. Dalam penelitian ini digunakan *Naive Bayes*, *KNN*, dan *SVM*. Kami juga melakukan perbandingan model dengan Particle Swarm Optimization. model 1 menguji tiga algoritma menggunakan validasi rasio 0,7-0,8 dan validasi silang 10 kali lipat, Pada Model 2 penulis menggunakan fitur seleksi yaitu *Particle swarm Optimization* dimana *PSO* digunakan sebagai optimasi. Dari model kedua, akurasinya adalah 88,00%. untuk *SVM + PSO*, 88,54%% untuk *NB + PSO* dan 81,58% untuk *K -NN + PSO*. Dan setelah dilakukan pengujian terhadap dua metode tersebut, ternyata *Naive Bayes + PSO* memiliki tingkat akurasi dan presisi yang paling tinggi. Dengan hasil penelitian ini diharapkan dapat bermanfaat bagi penelitian dan pengembangan terkait sentiment masyarakat terhadap kebijakan-kebijakan yang lakukan kedepannya, dengan hasil ini kedepannya dapat dikembangkan lagi menjadi model aplikasi yang bisa langsung melakukan labeling terhadap sentiment masyarakat,

Kata Kunci: Teks Mining, Sentimen, Algoritma

Abstract -*. The pandemic gave rise to new rules and terms in society. Various countries have their own regulations, including Indonesia under the name PSBB. For this reason, the author tries to conduct research related to the PSBB condition in Indonesia with the aim and purpose of knowing people's sentiments towards this, the authors carry out this modeling positively and negatively. . model in a tweet on Twitter. We capture information through Twitter media which then we process the data so that it is ready to be tested on the algorithm used. In data collection and collection, we use fast miner application. In this study, Naive Bayes, KNN, and SVM were used. We also did a model comparison with Particle Swarm Optimization. model 1 tested three algorithms using a validation ratio of 0.7-0.8 and cross-validation 10 times, In Model 2 the author uses a selection feature, namely Particle swarm Optimization where PSO is used as optimization. From the second model, the accuracy is 88.00%. for SVM + PSO, 88.54%% for NB + PSO and 81.58% for K -NN + PSO. And after testing the two methods, it turns out that Naive Bayes + PSO has the highest level of accuracy and precision. With the results of this study, it is hoped that it will be useful for research and development related to public sentiment towards policies carried out in the future, with these results in the future it can be further developed into an application model that can directly label public sentiment.*

Keywords: Text Mining, Sentiment, Algorithm

PENDAHULUAN

Penggunaan media sosial semakin meningkat pada masyarakat Indonesia. Penggunaan jaringan *internet* sendiri saat ini berkembang dengan pesat, Media sosial adalah sebuah wadah berbasis daring yang memungkinkan penggunanya berinteraksi dengan orang lain tanpa adanya batasan waktu, batasan area seperti wilayah,kota dan bahkan Negara. Indonesia merupakan salah satu negara dengan angka pengguna media sosial tertinggi di dunia. Kementerian Komunikasi dan Informatika

Naskah diterima: 14 Januari 2022, direvisi: 14 Maret 2022, disetujui: 16 Maret 2022

(Kemenkominfo) menyatakan 95% dari sekitar 63 juta pengguna internet adalah pengguna media sosial. Twitter menjadi salah satu media sosial yang sering digunakan oleh masyarakat Indonesia. Country Industry Head Twitter Indonesia mengklaim bahwa Indonesia merupakan negara dengan pertumbuhan pengguna aktif harian Twitter-nya paling besar (Abdulloh & Hidayatullah, 2019).

Pandemi yang kita kenal dengan seutan COVID19 mengganggu anyak aktivitas di eragai elahan dunia yang menjadi tantangan agi semua negara yang terkena virus Corona yang mengancam seagian esar dunia. Serikat dan Italia memiliki sekitar 2 5% tingkat kematian untuk virus baru. (Pastor, 2020)

Text mining adalah penerapan konsep dari teknik data mining untuk menemukan pola dalam teks dengan tujuan menemukan informasi yang erguna untuk tujuan tertentu. (Hasan, 2018). *Text mining* dapat diolah untuk berbagai macam keperluan diantaranya adalah untuk *summarization*, pencarian dokumen teks dan sentiment analisis (Aaputra, Didi Rosiyadi, Windu Gata, & Syepry Maulana Husain, 2019)

Analisis sentimen ertujuan untuk mengklasifikasikan ulasan pengguna menjadi opini positif atau negatif erdasarkan perasaan emosi opini dan sikap. Klasifikasi analisis sentimen diagi menjadi tiga level yaitu level dokumen level kalimat dan level aspek. Dalam penelitian ini analisis sentimen hanya memiliki satu level kalimat. Metode yang digunakan adalah metode Support Vector Machine (SVM) dan metode Herd Optimization (PSO). Analisis sentimen adalah proses menentukan apakah seuah kalimat cenderung positif negatif atau netral. (Warjiyono et al., 2019).

Text mining adalah langkah analisis teks yang dilakukan secara otomatis oleh komputer untuk mengekstrak informasi erkualitas dari serangkaian teks yang dirangkum dalam seuah dokumen. .(Adhe, Rachman, Goejantoro, & Tisna, 2020)

Opinion mining atau dikenal sebagai analisa sentiment adalah proses yang bertujuan untuk menentukan apakah polaritas kumpulan teks tulisan (dokumen, kalimat, paragrap, dan lain-lain) cenderung ke arah positive, negative, atau netral (Taufik, 2018). Menurut Ipmawati dalam jurnal Wardhani Analisis sentimen adalah sebuah proses untuk menentukan sentimen atau opini dari seseorang yang diwujudkan dalam bentuk teks dan bisa dikategorikan sebagai sentimen posisif atau negative (Hadna, Santosa, & Winarno, 2016). Analisis sentimen juga merupakan riset komputasional dari opini, sentimen dan emosi yang diekspresikan secara tekstual. (Wardhani et al., 2018)

Classification adalah model data mining dimana *classifier* dibangun untuk memprediksi label klasifikasi seperti “aman” atau “berisiko” untuk data aplikasi pinjaman, “ya” atau “tidak” untuk data pemasaran atau “label A”, “label B ", "label C" untuk data. Kategori-kategori tersebut dapat diwakili oleh nilai-nilai sesuai dengan kebutuhannya (Hasan, 2018).

Support vector machine (SVM) adalah sistem pemelajaran classifier yang menggunakan ruang hipotetis seagai fungsi linier dalam ruang fitur erdimensi tinggi dilatih dengan metode optimasi algoritma pemelajaran erasis teori. . teori statistik (Widiastuti Informasi dan Gunadarma 2007)

NBC merupakan salah satu algoritma teknik data mining yang menerapkan teori Bayesian untuk klasifikasi. Teorema keputusan Bayesian adalah metode statistik dasar untuk pengenalan pola. Naive Bayes didasarkan pada asumsi sederhana ahwa nilai atriut independen ersyarat jika mereka menerima nilai output. Dengan kata lain mengingat nilai output proailitas yang diamati secara kolektif adalah produk dari proailitas individu. (Mustafa, Ramadhan, & Thenata, 2018).

Naive Bayes merupakan pengklasifikasian dengan metode probabilitas dan statistik yang dikemukan oleh ilmuwan Inggris *Thomas Bayes*, yaitu memprediksi peluang di masa depan berdasarkan pengalaman di masa sebelumnya.(Manalu, Sianturi, & Manalu, 2017). Algoritma Nearst Neighbor merupakan Algoritma yang melakukan klasifikasi berdasarkan kedekatan lokasi (jarak) satu data dengan data yang lain. Pengenalan pola wilayah dengan menggunakan algoritma "k-Nearest Neighbor" (k-NN) merupakan metode klasifikasi, dimana objek baru diberi label berdasarkan objek yang terdekat.(Susanto & Riana, 2016)

Particle Swarm Optimization (PSO) yang sediakan satu set kelas primitif berguna untuk memecahkan masalah optimasi kombinatorial dan kontinu. (James V. Miranda, 2018). Algoritma Particle Swarm Optimization (PSO) adalah meta-heuristik terinspirasi dari alam yang telah digunakan sebagai alat pengoptimalan dalam berbagai aplikasi sejak tahun 1995. (Sedighizadeh, Masehian, Sedighizadeh, & Akbaripour, 2021)

Algoritma PARTICLE swarm optimization (PSO) terinspirasi oleh perilaku kawanan urung diusulkan oleh Kennedy dan Eerhart pada tahun 1995 dalam aturan PSO setiap partikel diwakili oleh titik dalam kerangka Cartesian dengan memilih solusi keseluruhan teraik. serta pengalaman teraik priadinya seagai model pemelajaran. Bahkan jika seuah partikel memiliki kecerdasan yang sangat rendah perilaku kolektif dari partikel yang ereda dapat menunjukkan kemampuan yang kuat untuk memecahkan masalah yang kompleks. (Xia et al., 2020).

METODOLOGI PENELITIAN

Metode penelitian yang akan dilakukan terdiri dari beberapa tahapan seperti pengumpulan data, pengolahan data awal, metode yang diusulkan,

eksperimen dan hasil pengujian serta evaluasi dan validasi hasil. Berikut ini akan dijelaskan tahapan dari metode penelitian yang akan dilakukan antara lain.

1. Pengumpulan Data

Data yang digunakan adalah data yang diamil dari jejaring sosial Twitter terkait "PSBB" di Indonesia terkait covid19. Data yang diperoleh dari tweet yang diamil pada ulan Maret 2020 dan April 2020 memiliki total 20.083 item data. Data diperoleh pada minggu terakhir ulan Maret dan minggu ke-2 ulan April dan data terseut diperoleh dari hasil survei pada tanggal 30 Maret 2020 7 April 2020. Data diperoleh menggunakan aplikasi Rapidminer menggunakan aplikasi Rapidminer Twitter. search engine dengan Twitter API dengan menempatkan query "PSBB" pada filter language dan filter date. Proses pada aplikasi Rapid Miner erlanjut ke langkah selanjutnya yaitu data preprocessing dan alignment l. Setelah diproses seperti menghapus duplikat dan eerapa string pengganti data dieri lael oleh ahli anotasi atau ahasa hingga 2632 data. Data erlael positif dan negatif data erlael dan data yang digunakan dalam penelitian ini erjumlah 1776 positif dan 856 setelah itu dilakukan pengolahan seagai pengkodean untuk kata kunci dan ngram transfer case.

Tabel 1. Tabel Data Twitter
Positive dan Negative

No	Text	Status
1	percaya saja mas pemerintah pasti mengambil jalan terbaik makanya dihimbau untuk PSBB physic distancing kembali ke kita dg menjaga kesehatan diri maka kita sudah menjadi pahlawan kesehatan	positive
2	Negara punya kewajiban untuk nanggung semua biaya hidup rakyatnya selama lockdown yes itu benar tp negara pasti tidak sanggup Mknya dipilih opsi tidak lockdown tp PSBB Ya tugas kita nurut dong kalau gk penting ya jangan keluar	positive
3	Semoga pandemi ini cepat berakhir Ayo tetap lakukan PSBB bersama jauhi kerumunan orang banyak dan tetap jaga kebersihan sekitar Bersama	positive
4	lockdown tidak rapid test tidak PSBB tapi matidak APD terbatas negara sudah dibocori soal awal Januari tapi milih cengengesan perkara sego kucing sego rawon	negative
5	Disuruh PSBB tapi mudik diperkenankan Kalau gitu mah ga perlu ada gugus tugas Ga guna juga	negative
6	pak melihat ketidak tegasan anda terhadap penetidakan PSBB measures termasuk ketakutan anda untuk tidak populer dalam melarang mudik baik anda minta izin atasan anda di untuk memulai pembuatan kuburan massal korban	negative

Sumber : Peneliti (2020)

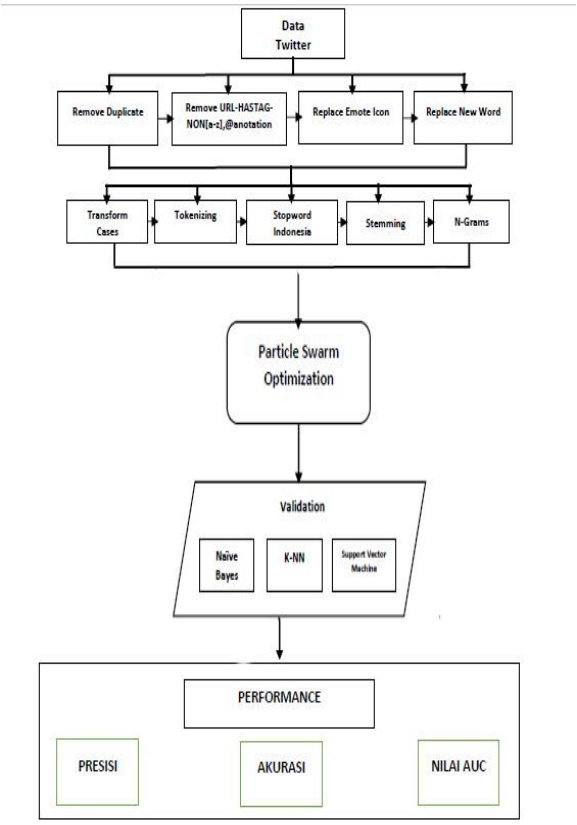
2. Pengolahan Data Awal

Pada tahap ini, klasifikasi teks atau sentimen dilakukan dengan langkah pra pemrosesan sehingga teks memiliki konten yang tidak sempurna seperti data yang hilang, data yang tidak valid atau bahkan kesalahan ketik sederhana. Selain itu, ada atribut data asing. Data paling baik ditolak karena keberadaannya dapat mengurangi kualitas atau akurasi. Pada titik ini, data dalam bentuk tweet dibersihkan sehingga menjadi terstruktur dan data dengan baik. Langkah-langkah yang telah dilakukan antara lain:

a. Hapus URL
b. Hapus data duplikat
c. Penggantian ikon emot
d. Menghilangkan angka, simbol dan tanda baca
e. Menghapus hashtag, username, dan RT
f. Ganti dengan string yang dihasilkan secara otomatis relatif terhadap bahasa kekinian
g. Huruf kecil
h. Filter token karakter min dan maks
i. Berhenti bicara
j. Menggunakan alat Ngrams

3. Model Yang Diusulkan

Metode atau model yang diusulkan yaitu menggunakan tiga algoritma yaitu *Naive Bayes*, *Support Vector Machine* dan *K-Nearest neighbour*.. Sebagai bahan perbandingan digunakanlah pengujian dengan menggunakan *feature PSO*. Particle swarm optimization sendiri menjadi salah satu *feature* yang paling sering digunakan dalam penelitian *text mining* terutam jika menggunakan algoritma seperti *Naïve Bayes, K-NN dan SVM. PSO* melakukan pencarian menggunakan populasidisebut *swarm* individu disebut partikel yang diperbarui dari iterasi ke iterasi. Untuk menemukan solusi optimal, setiap partikel mengubah arah pencariannya sesuai dengan dua faktor, pengalaman terbaik sebelumnya.



Sumber : Peneliti (2020)

Gambar 1. Alur Penelitian

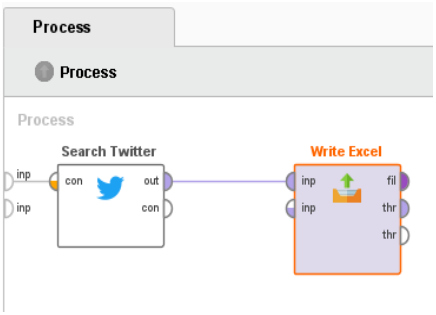
diatas dapat dijelaskan bahwa dataset yang dikumpulkan akan diolah dalam 4 tahapan, yaitu

1. Text Processing
2. Feature Selection
3. Validation
4. Evaluation

Dari keempat tahapan diatas maka akan didapatkan hasil komparasi antara algoritma *Naive Bayes*, *K-Nearest Neighbor* dan *Support Vector Machine* dengan dua model yaitu dengan fitur seleksi *PSO*.

HASIL DAN PEMBAHASAN

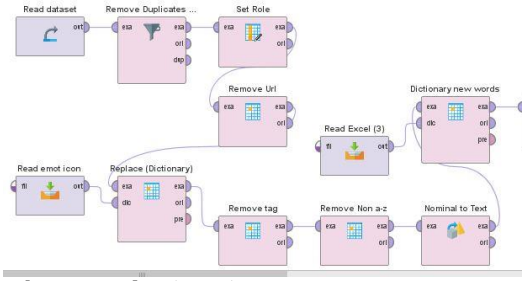
Pada tahap pengumpulan data, berbagai jenis perangkat untuk mendapatkan data, banyak cara yang bisa dilakukan dalam melakukan *crawling* data kepada situs *website* data yang diperlukan. Adapun dalam penelitian ini pengambilan data menggunakan aplikasi rapidminer. Berikut proses pengambilan data menggukan rapidminer. Pengumpulan data *twitter* menggunakan *plugin crawler* yang terdapat dalam *tools* rapidminer dengan bantuan menggunakan *access twitter* API sebagai konektor ke API *Twitter*. Kita harus mempunyai akun twitter terlebih dahulu untuk bisa mendapatkan akses dari rapidminer ke twitter.



Sumber : Peneliti (2020)

Gambar 2. Crawling Twitter Rapidminer

Pengambilan data *twitter* dilakukan dengan memasukan token yang didapatkan dari *twitter API* , serta memasukkan *query* yang akan dibutuhkan. Kali ini yang penulis cari datanya adalah terkait *PSBB* Dari pemilihan kata tersebut dibuatlah limtasi data yang ingin di *crawling* penulis melakukan sebanyak dua kali penarikan data jadi total data yang didapatkan sebanyak 20.289 data twit terkait *PSBB.* Tahap selanjutnya melakukan proses *removing*, dan *replacing* data seperti remove duplicates,remove url,string replace emot icon,remove non character seperti symbol-simbol, remove hastag dan annotation dan peneliti membuatkan sebuah string replace baru dengan membuat sebuah kamus yang isinya berbahasa istilah gaul atau kekinian setelah itu barulah masuk ke dalam labelisasi yang di lakukan oleh ahli bahasa atau bertindak sebagai annotator dengan tugas yaitu melakukan labelisasi terhadap twit tersebut dengan pelabelan positive dan negative. Setelah melalui remove duplicate data berkurang mejadi 10.147 dan setelah dilakukan pelabelan data menjadi 1176 yang terdiri dari 536positive sebanyak dan negative sebanyak 640 data.



Sumber : Peneliti (2020)

Gambar 3. Desain pengolahan data awal

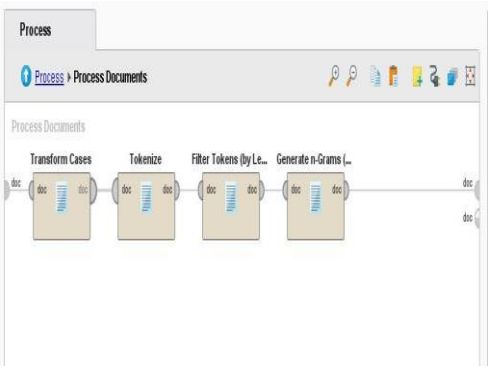
1. Fase Pengolahan Data

Tahap *data preparation* merupakan tahap dengan proses penyiapan data yang bertujuan untuk mendapatkan data yang bersih dan siap untuk digunakan dalam penelitian. Dalam *text mining* tahapan awal yang akan dilakukan adalah tahap *preprocessing*. Berikut merupakan tahapan yang dilakukan dalam *preprocessing*:

Berikut merupakan tahap desain model *preprocessing* yang digunakan dalam melakukan tahapan awal pengolahan data.

2. *Fase Preprocessing*

Dalam fase ini Penulis melakukan preprocessing data Adapun tujuan dari proses ini adalah menghilakan noise data guna mentransformasikan data ke suatu format yang dibutuhkan disini penulis menggunakan tools di rapidminer seoerti transform casas,tokenizing,filter length,stopwords dan n-grams. Namun dalam penerapannya Penulis juga menggunakan bantuan aplikasi dari Gataframeworks untuk melakukan proses stopwords agar lebih maksimal.

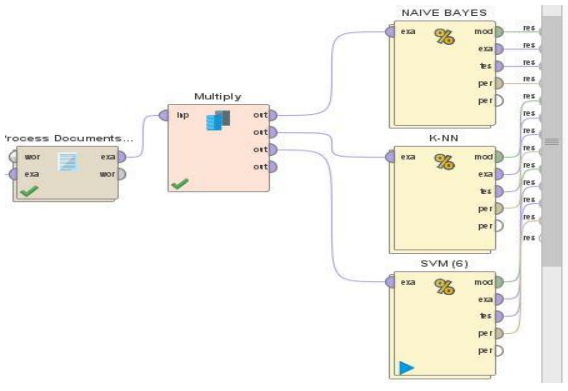


Sumber : Peneliti (2020)

Gambar 4. Model Preprocesing

3. Fase Pemodelan

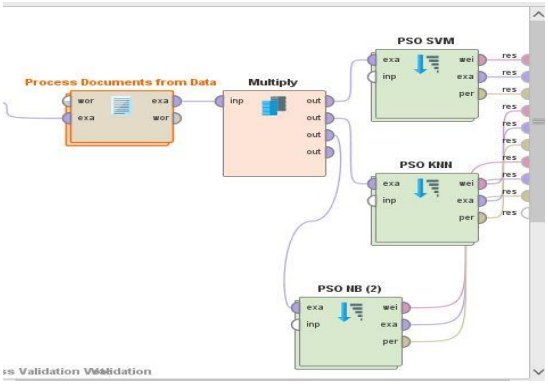
Dalam fase ini merupakan pemilihan teknik mining dengan menentukan algoritma atau model yang akan digunakan. *Tool* yang digunakan adalah *RapidMiner* versi 9.7.



Sumber : Peneliti (2020)

Gambar 5. Desain Model Algoritma

Berikut adalah desain model *Rapidminer* untuk pemodelan dengan menggunakan operator *multiply*, operator ini sendiri berfungsi untuk menjalankan beberapa algoritma sekaligus kemudian digunakan *cross validation* Parameter ini menentukan jumlah lipatan (jumlah himpunan bagian) dari *ExampleSet* yang harus dibagi menjadi. Setiap *subset* memiliki jumlah Contoh yang sama. Juga jumlah iterasi yang akan terjadi adalah sama dengan jumlah lipatan. Jika port output model terhubung, *subproses* Pelatihan diulangi sekali lagi dengan semua Contoh untuk membangun model akhir



Sumber : Penliti (2020)

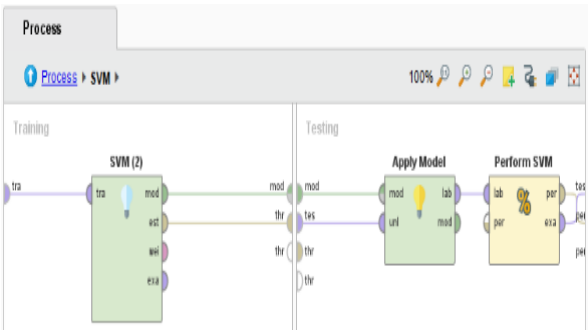
Gambar 6. Desain algoritma dengan PSO

Model ini digunakan sebagai perbandingan dengan model sebelumnya tanpa optimasi, seperti diketahui PSO merupakan fitur yang banyak sekali digunakan dalam beberapa algoritma dasar. Untuk itu Penulis memilih PSO sebagai model untuk perbandingan.

4. Fase Evaluasi *(Evaluation Phase)*

Tahapan evaluasi bertujuan untuk menentukan nilai kegunaan dari model yang telah berhasil dibuat pada langkah sebelumnya. Untuk evaluasi digunakan *10-fold cross validation*. Dalam 10 *fold Cross Validation*, data dibagi menjadi 10 fold berukuran kira-kira sama, sehingga kita memiliki 10 *subset* data untuk mengevaluasi kinerja model atau algoritma. Misalkan kita memiliki 10 data dimana kita akan melakukan K-fold Cross Validation pada data tersebut, dimana data akan dibagi menjadi data testing (untuk pengujian model) dan data training (untuk melatih model). Pada percobaan kita akan menentukan nilai K = 10 dimana data nantinya akan ada 10 lipatan. Tujuannya adalah untuk menemukan kombinasi data yang terbaik, bisa saja dalam akurasi, presisi, eror dan lain-lain.

Berikut ini desain proses pengujian model metode *Support Vector Machine* yang digunakan yaitu :



Sumber : Peneliti (2020)

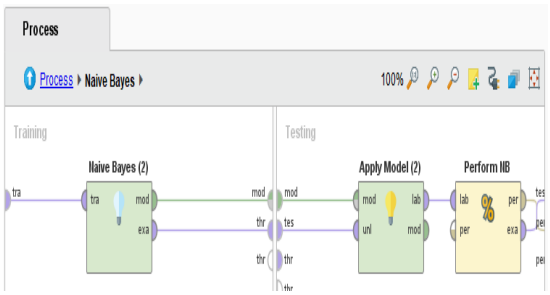
Gambar 7. Desain algoritma SVM

model klasifikasi penulis menerapkan beberapa kali uji dengan menggunakan 10 folds cross validation dan juga pengaturan ratio untuk data training dan testing sebesar 0.7 – 0.9 ratio dan juga dilakukan penambahan fitur *Particle swarm optimization*. Berikut adalah hasil yang didapatkan berupa Akurasi, Presisi dan Nilai AUC.

Tabel 2. Hasil Algoritma SVM

SVM	Akurasi	Presisi	AUC
0.9 Ratio	86.92	85,29	0.959
0.8 Ratio	84.88	84,62	0.958
0.7 Ratio	86.86	85,57	0.976
10folds	87.38	83,72	0.967

Hasil Pengujian Model *Naive Bayes Classifier*
Berikut ini desain proses pengujian model metode *Naive Bayes Classifier* yang digunakan yaitu :



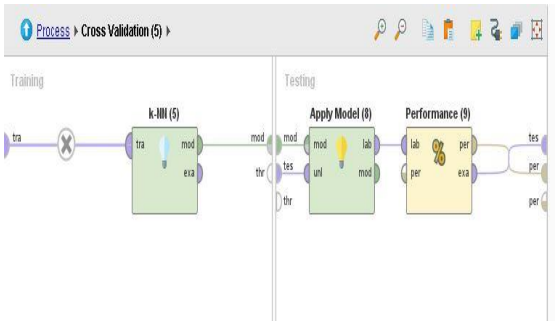
Sumber : Peneliti (2020)
Gambar 8 Proses *10-Fold Cross Validation NB*

Pada prosesnya disetiap model klasifikasi penulis menerapkan beberapa kali uji dengan menggunakan 10 folds cross validation dan juga pengaturan ratio untuk data training dan testing sebesar 0.7 – 0.9 ratio dan juga dilakukan penambahan fitur *Particle swarm optimization*. Berikut adalah hasil yang didapatkan berupa Akurasi, Presisi dan Nilai AUC.

Tabel.3 Hasil Algoritma Naïve Bayes

Naïve BAYES	Akurasi	Presisi	AUC
0.9 Ratio	89.23	88,06%	0.5
0.8 Ratio	86.43	85,61%	0.79
0.7 Ratio	86.60	86,22%	0.8
10folds	86.84	85,56%	0.588

Hasil Pengujian Model Metode *K-Nearest Neighbor*
Berikut ini desain proses pengujian model metode *K-Nearest Neighbor* yang digunakan yaitu :



Sumber : Peneliti (2020)
Gambar.9 Proses *10-Fold Cross Validation* K-NN

Gambar 9 pada prosesnya disetiap model klasifikasi penulis menerapkan beberapa kali uji dengan menggunakan 10 folds cross validation dan juga pengaturan ratio untuk data training dan testing sebesar 0.7 – 0.9 ratio dan juga dilakukan penambahan fitur *Particle swarm optimization*. Berikut adalah hasil yang didapatkan berupa Akurasi, Presisi dan Nilai AUC.

Tabel.4 Hasil Algoritma K-NN

K-NN	Akurasi	Presisi	AUC
0.9 Ratio	73.20	69,74	0.848
0.8 Ratio	73.64	71,03	0.835
0.7 Ratio	79.23	78,79	0.835
10folds	76.39	73,59	0.874

Hasil Pengujian Model Particle Swarm Optimization
Particle Swarm Optimization (PSO) algoritma berbasis populasi yang mengeksploitasi individu dalam pencarian. Dalam model ini penulis menggunakan population size = 5 dan maximum number of generations=30.
Hasil Pengujian Dengan *Particle Swarm Optimization*, Dari modeling pengujian menggunakan *particle swarm optimization* ini maka dihasilkan nilai akurasi,Presisi dan Kurva *ROC* yang akan dipaparkan berikut ini. Adapun perbandingan hasil komparasi akurasi, presisi dan AUC Algoritma telah digunakan sebagai berikut :

Tabel 5 Hasil Pengujian Tiga Algoritma Dengan PSO

	Akurasi	Presisi	AUC
Naïve Bayes +PSO	88.54	86,37	0.703
KNN+PSO	81.58	84,50	0.868
SVM + PSO	88.00	85,72	0.966

Dari hasil yang terlihat bahwa tingkat akurasi,presisi dan *AUC* masing-masing algoritma naik *SVM* naik sebesar 0,50% dan kenaikan tertinggi terjadi di algoritma *K-NN* naik sebesar 5%, kecuali pada algoritma *naïve bayes* terjadi penurunan di akurasi

tetapi untuk presisi dan *AUC* semua algortima menujukan kenaikan.

KESIMPULAN

Akhirnya kesimpulan yang penulis rangkum dari bab-bab sebelumnya mengenai penelitian ini yaitu Analisis Sentiment terkait konsisi PSBB di Indonesia ,maka penulis dapat mengambil kesimpulan dari hasil pengujian yang dilakukan dengan tiga model.

Model ke-1 penulis mencoba melakukan pengujian dengan tiga algoritma yang digunakan yaitu SVM,Naïve Bayes dan KNN penulis sengaja memilih 3 algoritma ini dikarenakan ketiga inilah yang paling banyak digunakan dalam kasus sentiment publik. Dari hasi pengujian itu sendiri dapat dilihat bahwa tingkat akurasi yang diperoleh sebenarnya sudah cukup tinggi yaitu SVM=87,38% , Naïve Bayes=86,84% dan K-NN=76,89%. Dari hasil pengujian mode ke-1 ini dapat diketahu ketiga algoritma ini cocok untuk digunakan dengan dataset terkait PSBB ini dikarenakan hasil akurasinya sudah mencapai diatas 80% yang dikatagorikan sudah cukup baik.

Model ke-2 penulis mencoba menambahkan fitur Particle Swarm Optimization sebagai model perbandingan dengan Model ke-1, dimana PSO dipilih sebagai perbandingan karena banyak penelitian terkait sentiment analisis menggunakannya dan beberapa menunjukan kenaikan akurasi.Dan benar saja Dari hasil pengujian dihasilkan 88.00% untuk SVM , 88,54% untuk Naïve Bayes yang justru turun akurasinya dari model sebelumnya yaitu model-1 dan KNN dengan nilai akurasi 81.58%.

Dengan demikian dapat dikatakan model *Naïve bayes* menggunakan *Partice Swarm* dapat digunakan dalam mengoptimalkan nilai akurasi,presisi dan AUC dan diharapkan model ini dapat digunakan untuk data-data ditahun berikutnya dan juga sentimen publik lainnya

REFERENSI

Aaputra, S. A., Didi Rosiyadi, Windu Gata, & Syepry Maulana Husain. (2019). Sentiment Analysis Analisis Sentimen E-Wallet Pada Google Play Menggunakan Algoritma Naive Bayes Berbasis Particle Swarm Optimization. *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, *3*(3), 377–382. https://doi.org/10.29207/resti.v3i3.1118

Abdulloh, N., & Hidayatullah, A. F. (2019). Deteksi Cyberbullying pada Cuitan Media Sosial Twitter. *Automata*, *Vol 1*(1), 1–5.

Adhe, D., Rachman, C., Goejantoro, R., & Tisna, D. (2020). Implementation Of Text Mining For Grouping Thesis Documents Using K-Means Clustering. *Jurnal EKSPONENSIAL*, *11*(2), 167–174.

Hadna, M. S., Santosa, P. I., & Winarno, W. W. (2016). Studi Literatur Tentang Perbandingan Metode Untuk Proses Analisis Sentimen Di Twitter. *Seminar Nasional Teknologi Informasi Dan Komunikasi*, *2016*(Sentika), 57–64. Retrieved from https://fti.uajy.ac.id/sentika/publikasi/makalah/2016/95.pdf

Hasan, F. N. (2018). Fuad Nur Hasan, Mochamad Wahyudi, *3*(1), 430–439.

James V. Miranda, L. (2018). PySwarms: a research toolkit for Particle Swarm Optimization in Python. *The Journal of Open Source Software*, *3*(21), 433. https://doi.org/10.21105/joss.00433

Manalu, E., Sianturi, F. A., & Manalu, M. R. (2017). Volume 1 No 2 Desember 2017 p-ISSN 2088-3943 e-ISSN 2580-9741 PENERAPAN ALGORITMA NAIVE BAYES UNTUK MEMPREDIKSI JUMLAH PRODUKSI BARANG BERDASARKAN DATA PERSEDIAAN DAN JUMLAH PEMESANAN PADA CV. PAPADAN MAMA PASTRIES. *Jurnal Mantik Penusa*, *1*(2), 16–21. Retrieved from https://ezp.lib.unimelb.edu.au/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=ffh&AN=2008-10-Aa4022&site=eds-live&scope=site

Mustafa, M. S., Ramadhan, M. R., & Thenata, A. P. (2018). Implementasi Data Mining untuk Evaluasi Kinerja Akademik Mahasiswa Menggunakan Algoritma Naive Bayes Classifier. *Creative Information Technology Journal*, *4*(2), 151. https://doi.org/10.24076/citec.2017v4i2.106

Pastor, C. K. L. (2020). Sentiment analysis of Filipinos and effects of extreme community quarantine due to coronavirus (COVID-19) Pandemic. *Journal of Critical Reviews*, *7*(7), 91–95. https://doi.org/10.31838/jcr.07.07.15

Sedighizadeh, D., Masehian, E., Sedighizadeh, M., & Akbaripour, H. (2021). GEPSO: A new generalized particle swarm optimization algorithm. *Mathematics and Computers in Simulation*, *179*, 194–212. https://doi.org/10.1016/j.matcom.2020.08.013

Susanto, W. E., & Riana, D. (2016). Komparasi Algoritma. *Jurnal Speed*, *8*(3), 18–27.

Taufik, A. (2018). Komparasi Algoritma Text Mining Untuk Klasifikasi Review Hotel. *Jurnal Teknik Komputer AMIK BSI (JTK)*, *IV*(2), 69–74. https://doi.org/10.31294/jtk.v4i2.3461

Wardhani, N. K., Rezkiani, Kurniawan, S., Setiawan, H., Gata, G., Tohari, S., … Wahyudi, M. (2018). Sentiment analysis

article news coordinator minister of maritime affairs using algorithm naive bayes and support vector machine with particle swarm optimization. *Journal of Theoretical and Applied Information Technology*, *96*(24), 8365–8378.

Warjiyono, Aji, S., Fandhilah, Hidayatun, N., Faqih, H., & Liesnaningsih. (2019). The Sentiment Analysis of Fintech Users Using Support Vector Machine and Particle Swarm Optimization Method. *2019 7th International Conference on Cyber and IT Service Management, CITSM 2019*. https://doi.org/10.1109/CITSM47753.2019.8965348

Widiastuti, D., Informasi, J. S., & Gunadarma, U. (2007). Analisa Perbandingan Algoritma Svm , Naive Bayes , Dan Decision Tree Dalam Mengklasifikasikan Serangan (Attacks), 1–8.

Xia, X., Gui, L., Yu, F., Wu, H., Wei, B., Zhang, Y. L., & Zhan, Z. H. (2020). Triple Archives Particle Swarm Optimization. *IEEE Transactions on Cybernetics*, *50*(12), 4862–4875. https://doi.org/10.1109/TCYB.2019.2943928

PROFIL PENULIS

Mugi Raharjo, Lahir di Jakarta, 29 Agustus 1995, berpendidikan terakhir meraih gelar Magister Komputer pada tahun 2020 di Program PASCA SARJANA STMIK Nusamandiri saat ini Mugi Raharjo mempunyai homebase di UNIVERSITAS NUSA MANDIRI pada Program Studi Teknik Informatika.

Jordy Lasmana Putra, Lahir di Lebak, 18 Oktober 1996, berpendidikan terakhir meraih gelar Magister Komputer pada tahun 2020 di Program PASCA SARJANA STMIK Nusamandiri saat ini Jordymempunyai homebase di UNIVERSITAS NUSA MANDIRI pada Program Studi Teknik Informatika.

Tommi Alfian Armawan Sandi, Lahir di Jakarta, 5 April 1996, berpendidikan terakhir meraih gelar Magister Komputer pada tahun 2020 di Program PASCA SARJANA STMIK Nusamandiri saat ini Tommi mempunyai homebase di UNIVERSITAS BINA SARANA INFORMATIKA.

Musriatun Napiah, Lahir di Brebes, 18 Mei 1996, berpendidikan terakhir meraih gelar Sarjana Komputer tahun 2019 di Program Studi Teknik dan Informatika STMIK Nusa Mandiri saat ini Musriatun Napiah mempunyai homebase di Universitas Bina Sarana Informatika pada program studi Teknologi Komputer .