

**ANALISIS SENTIMEN TERHADAP PEMINDAHAN IBU KOTA PADA MEDIA SOSIAL
TWITTER MENGGUNAKAN ALGORITMA NAIVE
BAYES BERBASIS PARTICLE SWARM OPTIMIZATION**

Karsito¹, Ahmad Taufiq²
Program Studi Teknik Informatika, Universitas Pelita Bangsa
karsito@pelitabangsa.ac.id

Disetujui, 31 Agustus 2019

ABSTRAKSI

Analisis Sentimen *Twitter* merupakan teknik untuk mengidentifikasi sentimen atau pendapat dalam *tweet* dan kemudian mengkategorikannya ke dalam *tweet* positif atau *tweet* negatif. Salah satu topik yang sering dibahas pada media sosial *twitter* adalah pemindahan Ibu Kota Jakarta ke Kalimantan Timur. Algoritma *naive bayes* merupakan metode pengklasifikasian probabilistik sederhana untuk klasifikasi teks. Masalah dalam pengklasifikasian ialah kurang optimalnya akurasi yang dihasilkan sehingga perlu seleksi fitur tambahan untuk meningkatkan akurasi menjadi lebih optimal. Tujuan dari penelitian ini adalah untuk mengoptimasi algoritma *naive bayes* dengan seleksi fitur *particle swarm optimization*. Dari hasil pengujian yang dilakukan dengan mengubah nilai *population size* dan *inertia weight* pada parameter *particle swarm optimization* hasil akurasi terbaik yang diperoleh adalah 91,50% sedangkan pengujian *naive bayes* tunggal diperoleh hasil akurasi sebesar 78,88%. Dengan demikian dapat disimpulkan bahwa penerapan seleksi fitur berbasis *particle swarm optimization* dapat mengatasi masalah seleksi fitur pada algoritma *naive bayes* sehingga dapat meningkatkan akurasi dalam menganalisis sentimen pengguna *Twitter* terhadap pemindahan Ibu Kota. Dari hasil akurasi tersebut dapat diketahui bahwa 35% pengguna *twitter* memberikan sentimen positif sedangkan 65% memberikan sentimen negatif terhadap pemindahan Ibu Kota Jakarta ke Kalimantan.

Kata Kunci : BPJS, Prediksi, *Data Mining*, *Decision Tree*, Algoritma C4.5, Algoritma Genetika.

Abstrct

Twitter's sentiment analysis is a technique to identify a sentiment or opinion in a tweet and then categorize it into a positive tweet or a negative tweet. One of the topics that are often discussed on social media is the transfer of the capital of Jakarta to East Kalimantan. The Naive Bayes algorithm is a simple probabilistic classifying method for text classification. The problem in the calibrating is the lack of optimal accuracy that is generated so that it needs a selection of additional features to improve the accuracy to be more optimal. The purpose of this research is to optimize the naive Bayes algorithm with the selection of particle swarm optimization features. From the results of the test done by changing the value of population size and inertia weight in the particle swarm optimization parameter The best accuracy obtained is 91.50% while the Naive Bayes single test obtained the accuracy of the results 78.88%. Thus it can be concluded that the implementation of a selection of features assistants particle swarm optimization can overcome the problem of selection of features in the Naive Bayes algorithm so as to improve accuracy in analyzing the sentiment of Twitter users against Transfer of the capital city. From the results of such accuracy can be known that 35% of Twitter users give positive sentiment while the 65% give negative sentiment to the transfer of the capital of Jakarta to Kalimantan.

Keywords : BPJS, Prediction, *Data Maining*, *Decission Tree*, *Algorithm C4.5*, *Genetic Algorithm*

1. Pendahuluan

Rencana pemindahan Ibu Kota baru ke Kalimantan Timur menuai banyak pendapat di berbagai kalangan. Sehingga memunculkan jejak pendapat di dunia maya, khususnya di media sosial *Twitter*. Pengguna media sosial *Twitter* di Indonesia terbilang cukup banyak tidak sedikit pula para pejabat yang memiliki akun *Twitter*. *Twitter* sebagai media sosial memiliki keunggulan dibandingkan dengan media sosial lainnya yaitu selalu menampilkan topik yang sedang hangat diperbincangkan (*Trending Topic*).

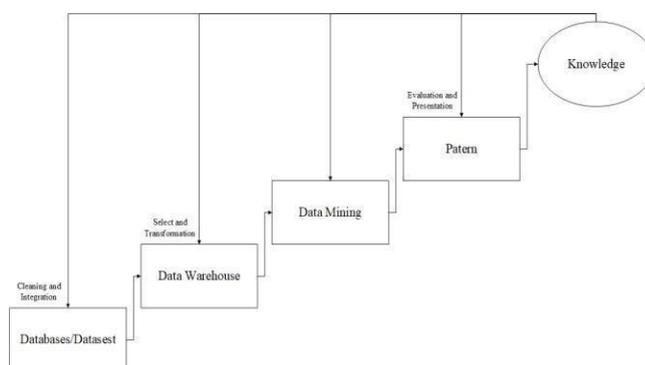
Kecenderungan masyarakat Indonesia dalam menilai kebijakan serta kinerja pemerintah di *Twitter* dapat menjadi acuan untuk mengetahui sentimen masyarakat terhadap pemindahan Ibu Kota Negara. Salah satu cara untuk mengetahui sentimen masyarakat terhadap pemindahan Ibu Kota adalah dengan menggunakan teknik *data mining*. Dengan teknik *data mining* dapat memprediksi, mengklasifikasikan, memfilter, dan mengelompokkan data [1].

Text mining sebagai salah satu jenis dari *data mining* merupakan metode untuk mencari informasi dari data tekstual. *Text mining* juga dapat membantu menentukan apa yang ada dalam penulisan subjek tertentu. *Text mining* dalam menganalisis sentimen mampu mengevaluasi perasaan terhadap suatu pernyataan [2]. Salah satu metode untuk mengetahui hasil sentimen adalah menggunakan algoritma *Naive Bayes*.

2. Tinjauan Pustaka

Algoritma *Naive Bayes* merupakan klasifikasi sederhana dan efektif. Namun *Naive Bayes* sebagai klasifikasi memiliki kelemahan yaitu sangat sensitif dalam seleksi fitur [3]. *Particle Swarm Optimization* merupakan metode untuk mengoptimasi parameter yang ada pada algoritma. Penggunaan *Particle Swarm Optimization* sangat sederhana dan sering dipakai untuk mengatasi masalah seleksi pada fitur [4].

Data mining adalah metode pembelajaran komputer untuk menganalisis dan mengekstraksi pengetahuan secara otomatis [5]. *Data mining* merupakan serangkaian proses untuk mengeksplorasi data menggunakan komputer dengan proses ekstraksi dan menggali pola penting dari data. *Data mining* dibagi menjadi beberapa kelompok yaitu: Prediksi, Estimasi, Klasifikasi, Pengklusteran, dan Asosia [6]. *Data mining* sebagai teknik pembelajaran komputer memiliki beberapa tahapan agar bisa menghasilkan pengetahuan. Berikut ini merupakan proses pengolahan *data mining*.



Gambar 1. Tahapan *Data Mining*

Berikut adalah tahapan-tahapan *data mining* atau *Knowledge Discovery in Database (KDD)*:

1. Integrasi Data: Merupakan fase awal, *sample* data di susun dan digabungkan dari berbagai sumber.
2. Pemilihan Data: Fase ini terdiri dari membuat kumpulan data target, atau berfokus pada subset variabel dan sampel data, di mana penemuan akan dilakukan.
3. Pembersihan Data: Proses dimana data yang telah tersimpan dibersihkan terlebih dahulu agar mendapatkan data yang konsisten.
4. Transformasi Data: Tahap perubahan data agar menjadi data yang valid.
5. *Data Mining*: Fase ini terdiri atas pencarian pola minat dalam bentuk representasi tertentu, tergantung pada tujuan *data mining*.
6. Evaluasi: Tahap ini terdiri atas interpretasi dan evaluasi pola yang ditambang.
7. Pengetahuan/*Knowledge*: langkah terakhir yang dapat membantu pengguna untuk membuat keputusan tentang informasi yang telah didapatkan..

Text mining adalah metode mencari atau mengekstraksi informasi bermanfaat dari data tekstual. *Text mining* juga dikenal sebagai *Text data mining* dan *knowledge Discovery in Textual Databases*. *Text mining* sama dengan *data mining* hanya saja *data mining* dirancang untuk menangani data terstruktur sedangkan *text mining* dirancang untuk menangani data yang tidak terstruktur atau semi-terstruktur [7]. *Text mining* berdasarkan karakteristiknya dibagi menjadi 7 bidang, meskipun memiliki perbedaan bidang ini saling terkait. Tujuh bidang *text mining* adalah sebagai berikut [8].

Gambar 2. 7 Bidang *Text Mining*

Text Preprocessing adalah proses sistematis dan fundamental yang digunakan dalam pengumpulan dokumen teks yang berkaitan dengan penghapusan karakter yang tidak berarti, kata-kata yang tidak bermakna dan penghapusan awal dan akhir dari suatu kata (Uysal & Gunal, 2014). *Text mining* mempunyai beberapa tahapan yaitu *tokenizing*, *stopword* dan *stemming* [9]:

1. Tahap *Tokenizing*

Tokenisasi membagi skema penulisan menjadi kata atau token pribadi. Dari segi terminologi, operasi ini bisa membawa berbagai macam. Penggunaan spasi dan tanda baca sebagai batas token adalah metode tokenisasi yang mudah dan efisien.

2. Tahap *Stopword*

Proses ini menghilangkan frasa dari dokumen yang tidak melakukan peran penting dalam menyediakan pola atau data. Misalnya, kata-kata seperti "Bagaimana", "Apa," "adalah" dll.

3. Tahap *Stemming*

Stemming adalah pengurangan token yang relevan menjadi satu jenis. Teknik stemming biasanya berisi awalan dan akhiran, pluralisasi yang tidak patut untuk dideteksi akan dihilangkan. Analisis sentimen adalah teknik pemrosesan bahasa alami untuk memperkirakan opini atau sentimen yang diungkapkan pada kalimat [10]. Analisis sentimen berperan untuk mengekstraksi, mendeteksi, meringkas opini, polaritas, dan emosi. Umumnya, analisis sentimen diperlukan untuk mengatasi masalah klasifikasi dengan kategori pendapat dan fakta, positif dan negatif [11].

Tujuan dari analisis sentimen adalah untuk menentukan pendapat dengan mempertimbangkan subjek tertentu. Perilaku dapat menunjukkan alasan, pendapat atau penilaian, kondisi kecenderungan [4]. Analisis sentimen juga bisa mengungkapkan perasaan emosional tentang kesedihan, kegembiraan, atau kemarahan [12].

Twitter adalah *microblogging* yang paling sering digunakan untuk melakukan penelitian analisis sentimen [13]. *Twitter* sebagai media komunikasi dengan membuat status yang di sebut *tweets* sangat digemari oleh para penggunanya [14]. *Twitter* Sentiment Analysis (TSA) menangani masalah menganalisis pesan yang diposting di *Twitter* dalam hal sentimen yang mereka ungkapkan [15]. *Twitter* memiliki keterbatasan dalam membuat postingan, tweets hanya dapat mencapai 140 karakter [16].

Naive Bayes merupakan metode pengklasifikasian probabilistik sederhana untuk klasifikasi teks. Algoritma *Naive Bayes* menerapkan teorema *Bayes* dengan asumsi independensi yang kuat, semua atribut independen atau tidak ketergantungan yang diberikan nilai variabel kelas [17]. Klasifikasi *Naive Bayes* adalah teknik yang bekerja untuk mengatasi masalah pada kelas tertentu, yaitu dengan mengasosiasikan objek dengan kategori diskrit [18]. Kelebihan dari algoritma *Naive Bayes* adalah: kesederhanaan implementasi, proses pembelajaran yang cukup cepat, juga

memberikan hasil yang cukup baik [19][20][21].

Particle Swarm Optimization adalah teknik optimisasi stokastik berbasis populasi yang digunakan untuk memecahkan berbagai masalah optimasi kontinu dan diskrit[22]. PSO berisi segerombolan partikel, dan setiap partikel yang sesuai dengan solusi potensial dapat terbang dengan kecepatannya di ruang pencarian. Karakteristik utamanya adalah bahwa semua partikel mengikuti pencarian kolaboratif di mana setiap partikel tertarik ke posisi global terbaik (*gbest*) di kawanan dan ke posisi terbaiknya sendiri (*pbest*) [23] *Particle Swarm Optimization* dirumuskan sebagai berikut:

$$V_i(t) = V_i(t - 1) + c1r1[X_{pbest_i} - X_i(t)] + c2r2[X_{gbest} - X_i(t)]$$

$$X_i(t) = X_i(t - 1) + V_i(t)$$

Rumus

$V_i(t)$ = Kecepatan partikel *i* saat iterasi *t*

$X_i(t)$ = Posisi partikel *i* saat iterasi *t*

$c1$ dan $c2$ = *Learning rates* untuk kemampuan individu

dan pengaruh sosial

$r1$ dan $r2$ = Bilangan random yang berdistribusi *uniformal* dalam *interval* 0 dan 1

X_{pbest_i} = Posisi terbaik partikel *i*

X_{gbest} = Posisi terbaik global

Confusion matrix merupakan suatu metode untuk melakukan proses perhitungan akurasi pada konsep *data mining*. *Confusion matrix* mempunyai perhitungan dengan 4 (empat) keluaran yaitu *recall*, *precision*, *accuracy* dan *error rate*. Konsep ini untuk melakukan proses evaluasi terhadap model klasifikasi berdasarkan proses perhitungan objek *testing* yang memiliki nilai prediksi benar atau salah. Proses perhitungan ini diimplementasikan kedalam tabel yang disebut *confusion matrix*. *Confusion matrix* adalah *dataset* yang terdiri dari 2 (dua) kelas yang bersifat positif dan negatif. Berikut adalah tabel *confusion matrix* yang terdiri dari 4 (empat) sel yaitu *true positif*, *false positif*, *true negative*, *false negative*.

Tabel 1. *Confusion Matrix*

		<i>True Class</i>	
		<i>Class 1</i>	<i>Class 2</i>
<i>Predicted Class</i>	<i>Class 1</i>	<i>True Positive</i>	<i>False Negative</i>
	<i>Class 2</i>	<i>False Positive</i>	<i>True Negative</i>

Adapun dalam menghitung akurasi menggunakan rumus sebagai berikut:

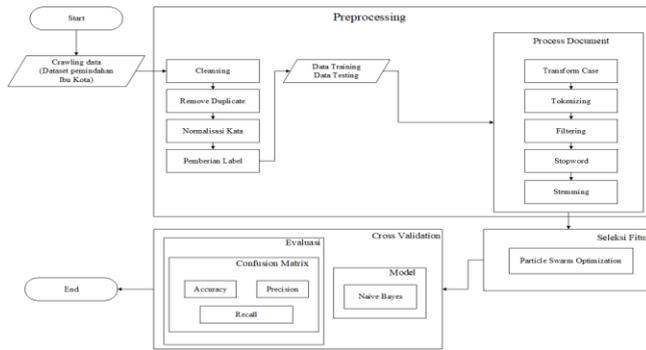
$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \times 100\%$$

$$Precision = \frac{TP}{TP+FP} \times 100\%$$

$$Recall = \frac{TP}{TP+FN} \times 100\%$$

3. Metodologi

Penelitian ini menggunakan algoritma *Naive Bayes* dengan menambahkan metode seleksi fitur yaitu *Particle Swarm Optimization* Berikut adalah gambar metode penelitian:



Gambar 4. Metode Penelitian

Uraian dari gambar di atas dapat dilihat sebagai mana pada penjelasan berikut ini:

A. Pengumpulan Data

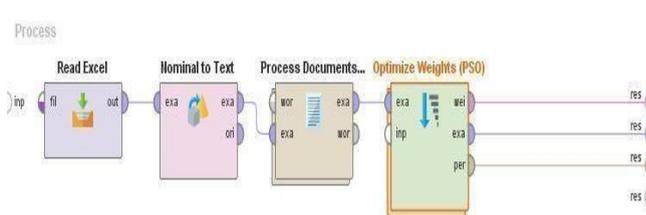
Pengambilan data *twitter* menggunakan operator search *twitter* di perangkat lunak RapidMiner. Data tweet diambil dari pendapat-pendapat masyarakat perihal pemindahan Ibu Kota yang terhubung dengan akun *twitter* Presiden Jokowi @jokowi. Dibawah ini adalah gambar proses pengambilan data *twitter* dengan perangkat lunak RapidMiner.

B. Preprocessing

Preprocessing merupakan tahap awal untuk mendapatkan data yang bersih supaya dapat lanjut ke tahap berikutnya, tahap preprocessing pada penelitian ini terdiri dari seleksi data, remove duplicate, cleansing, normalisasi kata, pemberian label, pembagian data, dan proses dokumen.

C. Seleksi Fitur

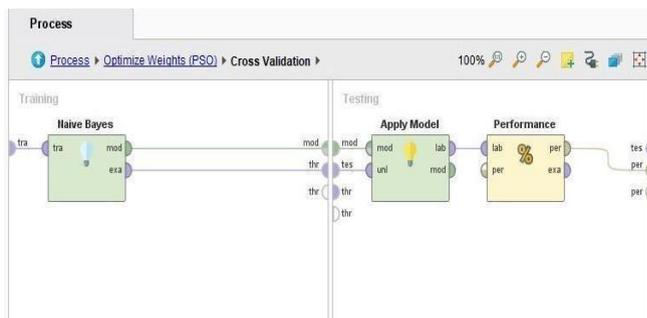
Tahap seleksi fitur pada penelitian ini menggunakan *Particle Swarm Optimization* agar pengelompokan lebih efektif.



Gambar 5. Penerapan Seleksi Fitur PSO

D. Penerapan Algoritma

Algoritma yang dipakai dalam penelitian ini adalah algoritma *Naive Bayes*, pada gambar dibawah ini merupakan contoh penerapan algoritma *Naive Bayes* pada perangkat lunak RapidMiner.



Gambar 6. Penerapan Algoritma *Naive Bayes*

E. Evaluasi

Untuk mengetahui hasil evaluasi pada penelitian ini menggunakan metode cross validation yang nantinya akan diukur dengan Confusion Matrix untuk mengetahui hasil akurasi, presisi, dan pengulangannya.

4. Hasil dan Pembahasan

A. Dataset

Pengambilan data dilakukan 2 kali yakni pada tanggal 18 September 2019 dan 08 Oktober 2019 yang terhubung dengan akun Presiden Joko Widodo @jokowi. Proses pengambilan data menggunakan operator Search *Twitter* yang telah disediakan oleh perangkat lunak RapidMiner data yang berhasil di crawling ialah 5000 data. Setelah melewati proses seleksi data yang didapat sebanyak 1000 data (Tweet) yang terdiri dari 522 data (Tweet) dengan respon positif dan 478 data (Tweet) dengan respon negatif. 1000 data (Tweet) tersebut nantinya akan di bagi menjadi 9:1, 900 data (Tweet) untuk data training dan 100 data (Tweet) untuk data testing.

Tabel 2. Sampel Dataset

Akun	Data Crawling	Data Bersih	Positif	Negatif
@jokowi	5000	1000	522	478

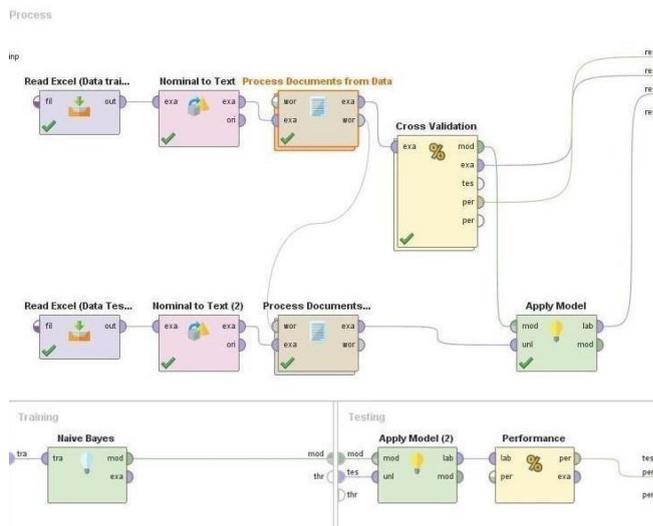
Dari hasil data yang diperoleh kemudian dikategorikan dengan variabel yang ada dengan nilai rendah, sedang dan tinggi dan pengujian berdasarkan data training dan data testing.

B. Proses Pengujian

Proses pengujian penelitian ini menggunakan metode *Naive Bayes* dengan seleksi fitur *Particle Swarm Optimization*. Pengujian dilakukan sebanyak 2 kali, pada pengujian pertama menggunakan *Naive Bayes* tanpa seleksi fitur *Particle Swarm Optimization*. Lalu pengujian kedua metode *Naive Bayes* digabungkan dengan seleksi fitur *Particle Swarm Optimization*.

C. Proses Pengujian Pertama

Pengujian pertama dilakukan dengan menggunakan model Naive yang mana nantinya hasil akurasi pada pengujian pertama akan dibandingkan dengan pengujian kedua. Dibawah ini merupakan gambar model pengujian pertama:



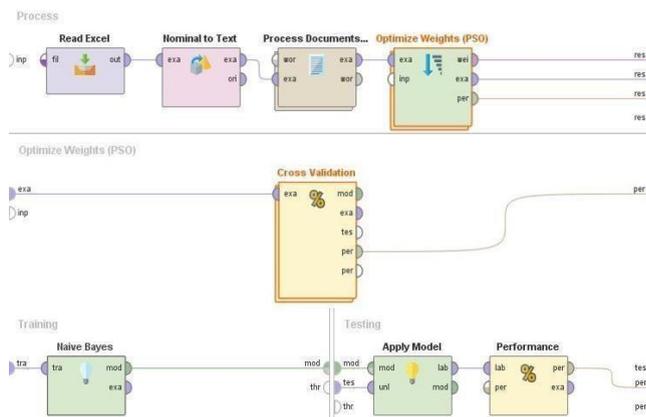
Gambar 7. Pengujian Algoritma *Naive Bayes*

Tabel 3. Hasil *Confusion Matrix* Pengujian Algoritma *Naive Bayes*

Pengujian	Akurasi	Recall	Precision
NB	78,88%	87,37%	78,02%

D. Proses Pengujian Kedua

Pengujian kedua dilakukan dengan menggunakan model *Naive Bayes* berbasis *Particle Swarm Optimization*. Berikut adalah gambar model dari pengujian kedua:



Gambar 8. Pengujian Algoritma *Naive Bayes* Berbasis PSO

Pada proses pengujian kedua peneliti melakukan eksperimen terhadap pengujian pertama dengan mengubah nilai pada parameter population size dan nilai inertia weight bertujuan untuk mencari hasil akurasi terbaik. Berikut ini adalah hasil dari eksperimen yang telah dilakukan:

Tabel 4. Rencana Eksperimen

Population size	Inertia weight	<i>Naive Bayes + PSO</i>			
		<i>Accuracy</i>	<i>Recall</i>	<i>Precision</i>	AUC
5-12	0.1-1.0	-	-	-	-

Tabel 5. Hasil Pengujian Mengubah Nilai Population Size

<i>Population Size</i>	<i>Inertia Wight</i>	NB+PSO			
		<i>Accuracy</i>	<i>Recall</i>	<i>Precision</i>	AUC
5	0.1	88,50%	84,62%	97,62%	0,938
6	0.1	88,50%	83,08%	99,00%	0,954
7	0.1	88,50%	83,85%	98,46%	0,946
8	0.1	89,50%	86,15%	97,46%	0,935
9	0.1	90,00%	86,92%	97,52%	0,943
10	0.1	89,50%	86,15%	97,40%	0,938
11	0.1	90,50%	86,92%	98,32%	0,955
12	0.1	90,50%	86,92%	98,40%	0,965

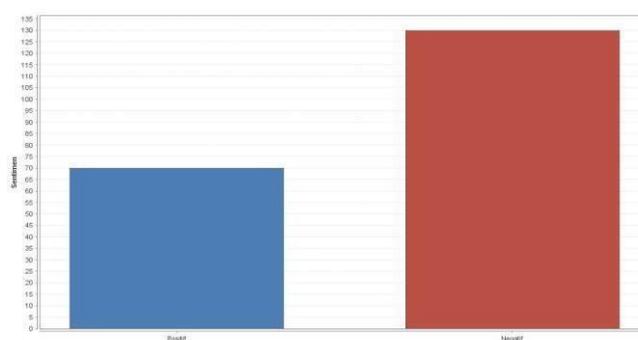
Tabel 6. Hasil Pengujian Mengubah Nilai Inertia Weight

Population Size	Inertia Weight	NB+PSO			
		Accuracy	Recall	Precision	AUC
10	0.2	90,00%	86,92%	97,54%	0,930
10	0.3	88,50%	84,62%	97,17%	0,924
10	0.4	88,50%	83,85%	98,40%	0,887
10	0.5	89,00%	83,85%	99,17%	0,949
10	0.6	89,50%	85,38%	98,40%	0,937
10	0.7	89,00%	85,38%	97,54%	0,938
10	0.8	88,50%	83,85%	98,38%	0,931
10	0.9	91,50%	87,69%	99,29%	0,957
10	1.0	89,50%	86,15%	97,69%	0,947

Berdasarkan pada hasil eksperimen dari kedua tabel diatas dapat diperoleh hasil akurasi terbaik dengan mengubah nilai population size dan nilai inertia weight menjadi 10 – 0.9.

E. Hasil Klasifikasi Sentimen

Hasil klasifikasi sentimen pada penelitian ini diperoleh dari hasil klasifikasi data testing dengan menggunakan model klasifikasi *Naive Bayes* berikut ini adalah gambar hasil klasifikasi sentimen.



Gambar 9. Hasil Klasifikasi Sentimen

Gambar diatas menunjukkan bahwa 70 data tweet bersifat positif dan 130 data tweet bersifat negatif. Dari grafik diatas dapat diketahui bahwa sentimen terhadap pemindahan Ibu Kota 35% bersifat positif dan 65% bersifat negatif.

F. Hasil Analisis Sentimen

Hasil analisis sentimen ini didapatkan dari hasil pengujian menggunakan model *Naive Bayes* tunggal dan *Naive Bayes* dengan Partical Swarm Optimization sehingga mendapatkan hasil sebagai berikut.

Tabel 7. Hasil Analisis Sentimen

Confusion Matrix	Pengujian Pertama	Pengujian Kedua	Keterangan	
	Naive Bayes	NB + PSO	Perbandingan Nilai	Hasil
Accuracy	78,88%	91,50%	16,11%	Naik
Recall	87,37%	87,69%	0,34%	Naik
Precision	78,02%	99,29%	27,17%	Naik
AUC	0,546	0,957	75,27%	Naik

5. Kesimpulan & Saran

Berdasarkan pada hasil penelitian yang telah dilakukan maka dapat disimpulkan bahwa penerapan seleksi fitur Partitcle Swarm Optimization dapat meningkatkan akurasi pada algoritma *Naive Bayes* dalam menganalisis sentimen pengguna *Twitter* terhadap pemindahan Ibu Kota. Hasil akurasi yang diperoleh yaitu 91,50% dan meningkatkan sebesar 16,11% dari hasil algoritma *Naive Bayes* tunggal yaitu 78,88%. Berasas dari hasil klasifikasi sentimen terhadap data testing dapat diketahui bahwa 35% pengguna *twitter* setuju dengan rencana pemindahan Ibu Kota dan 65% menolak rencana tersebut. Alasan banyaknya yang menolak rencana pemindahan Ibu Kota karena perekonomian Indonesia masih belum stabil dan pembangunan Ibu Kota baru akan berdampak buruk pada ekosistem di Kalimantan Timur.

Bertumpu pada kesimpulan diatas maka penulis menyarankan kepada penelitian selanjutnya untuk bereksperimen dengan teknik seleksi fitur yang lain. Saran tersebut bertujuan untuk menguji apakah performa yang diperoleh dari teknik seleksi fitur lain lebih baik dari teknik seleksi fitur *Particle Swarm Optimization*.

Daftar Pustaka

- [1] E. H. A. Rady and A. S. Anwar, "Prediction of kidney disease stages using data mining algorithms," *Informatics Med. Unlocked*, vol. 15, no. December 2018, p. 100178, 2019.
- [2] L. Zhang, R. Ghosh, M. Dekhil, M. Hsu, and B. Liu, "Combining lexicon-based and learning-based methods for twitter sentiment analysis," *HP Lab. Tech. Rep.*, no. 89, pp. 89–91, 2011.
- [3] W. Duan, Q. Cao, Y. Yu, and S. Levy, "Mining online user-generated content: Using sentiment analysis technique to study hotel service quality," *Proc. Annu. Hawaii Int. Conf. Syst. Sci.*, pp. 3119–3128, 2013.
- [4] A. S. H. Basari, B. Hussin, I. G. P. Ananta, and J. Zeniarja, "Opinion mining of movie review using hybrid method of support vector machine and particle swarm optimization," *Procedia Eng.*, vol. 53, pp. 453–462, 2013.
- [5] M. S. Mustafa, M. R. Ramadhan, and A. P. Thenata, "Implementasi Data Mining untuk Evaluasi Kinerja Akademik Mahasiswa Menggunakan Algoritma Naive Bayes Classifier," *Creat. Inf. Technol. J.*, vol. 4, no. 2, p. 151, 2018.
- [6] L. Sing'oei and J. Wang, "Data Mining Framework for Direct Marketing: A Case Study of Bank Marketing," *IJCSI Int. J. Comput. Sci. Issues*, vol. 10, no. 2, pp. 198–203, 2013.
- [7] S. Vijayarani, M. J. Ilamathi, M. Nithya, A. Professor, and M. P. Research Scholar, "Preprocessing Techniques for Text Mining -An Overview," vol. 5, no. 1, pp. 7–16, 2015.
- [8] M. Gary, D. Dursun, E. John, F. Andrew, H. Thomas, and R. A. Nisbet, *Practical Text Mining and Statistical Analysis for Non-structured Text Data Applications*. 2012.
- [9] A. N. Hidayat, "Analisis Sentimen Terhadap Wacana Politik Pada Media Masa Online Menggunakan Algoritma Support Vector Machine Dan Naive Bayes," *J. Elektron. Sistim Inf. Dan Komput.*, vol. 1, no. 1, pp. 1–7, 2015.
- [10] T. Carpenter and T. Way, "Tracking Sentiment Analysis through Twitter," *Proc. 2012 Int. Conf. Inf. Knowl. Eng.*, no. Figure 1, 2013.
- [11] M. Marie-Francine, L. Juanzi, and C. Tat-Seng, *Mining User Generated Content*, vol. 53, no. 9. 2014.
- [12] A. Idrus, H. Brawijaya, and Maruloh, "Sentiment Analysis of State Officials News on Online Media Based on Public Opinion Using Naive Bayes Classifier Algorithm and Particle Swarm Optimization," *2018 6th Int. Conf. Cyber IT Serv. Manag. CITSM 2018*, no. Citsm, pp. 1–7, 2019.
- [13] P. Alexander and P. Paroubek, "Twitter as a Corpus for Sentiment Analysis and Opinion Mining," *Ces. Dermatologie*, vol. 46, no. 5, pp. 223–229, 2010.
- [14] J. Spencer and G. Uchyigit, "Sentimentor: Sentiment analysis of twitter data," *CEUR Workshop Proc.*, vol. 917, pp. 56–66, 2012.
- [15] A. Giachanou and F. Crestani, "Like it or not: A survey of Twitter sentiment analysis methods," *ACM Comput. Surv.*, vol. 49, no. 2, 2016.
- [16] 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," *2018 6th Int. Conf. Cyber IT Serv. Manag. CITSM 2018*, no. Citsm, pp. 1–6, 2018.
- [17] L. Hande and G. Patnaik, "Review of sentiment analysis using naive bayes and neural network classifier," *Int. J. Sci. Eng. Technol. Res.*, vol. 03, no. 07, pp. 1110–1113, 2014.

- [18] T. Chris, P. Nishant, P. Hrishikesh, L. T Y, and T. SooTee, "Classifying Twitter Data with Naive Bayes Classifier," 2012.
- [19] J. Rothfels and J. Tibshirani, "Unsupervised sentiment classification of English movie reviews using automatic selection of positive and negative sentiment items," *CS224N-Final Proj.*, vol. 43, no. 2, pp. 52–56, 2010.
- [20] F. Bütow, F. Schultze, and L. Strauch, "Semantic Search: Sentiment Analysis with Machine Learning Algorithms on German News Articles," 2011.
- [21] H. Thakkar and D. Patel, "Approaches for Sentiment Analysis on Twitter: A State-of-Art study," 2015.
- [22] X. Liang, W. Li, Y. Zhang, and M. C. Zhou, "An adaptive particle swarm optimization method based on clustering," *Soft Comput.*, vol. 19, no. 2, pp. 431– 448, 2014.
- [23] Y. Li, Z. H. Zhan, S. Lin, J. Zhang, and X. Luo, "Competitive and cooperative particle swarm optimization with information sharing mechanism for global optimization problems," *Inf. Sci. (Ny)*, vol. 293, pp. 370–382, 2015.