

7 731-2433-1-ED.pdf

Pengukuran Kinerja Optimasi Algoritma Bat Pada Algoritma Naive Bayes, KNN Dan Decision Tree Untuk Sentimen Analisis Di Lini Masa Twitter

Candra Adipradana

¹⁾Teknik Informatika, Universitas Kahuripan Kediri
candra@kahuripan.ac.id

ABSTRACT

Social media is a very effective communication medium in today's digital era. Twitter is one of the social media that can be used as data mining-based research, namely sentiment analysis. This study aims to compare the performance levels of sentiment analysis accuracy on 3 classification algorithms namely Naïve Bayes, Decision Tree and K-NN as well as optimize the accuracy results for all three. Furthermore, the results of accuracy before optimization and after optimization are compared. The optimization algorithm used in this research is the Bat Algorithm. In this study, two research scenarios were made: first, calculating the accuracy of the Naïve Bayes, K-NN algorithm, and Decision Tree. Second, optimizing the classification results of the 3 algorithms with the Bat algorithm method, which then re-tested the accuracy value. In the first scenario the percentage is generated from the accuracy value of Naïve Bayes of 33,58, K-NN of 33,61 and Decision Tree of 32,82. In the second scenario, using one of the objective functions, namely $f(x) = x^2$, the Naïve Bayes value is obtained 39,01, K-NN 66,15 and Decision Tree 76,63. From the results of 3 the optimization test of classification Algorithm, it was found that the overall objective functions of the Bat algorithm were all able to increase the data accuracy value from before optimization. From all the tests, it was found that the Decision Tree algorithm has the highest average value of optimization increment, namely 43,81 %.

Keywords: Naïve Bayes, Decision Tree, K-NN, Bat Algorithm, Optimation

I. PENDAHULUAN

Pengguna internet di Indonesia tercatat pada tiap tahun terus mengalami peningkatan dibandingkan pada tahun sebelumnya[1][2][3]. Adanya suatu media sosial seperti Twitter[4][5] mendorong orang melakukan penelitian untuk mengetahui apa yang sedang dipikirkan oleh orang kebanyakan. Oleh sebab itu dikembangkan suatu teknik klasifikasi untuk mengelompokkan suatu opini yang sedang berkembang. Sebuah analisa sentimen atas opini yang berkembang di media twitter dapat dikelompokkan menjadi kategori opini positif, negatif ataupun netral[6].

Manfaat dari pengklasifikasian ini adalah mampu meneliti tren dan memperkirakan produk yang dihasilkan. Hal ini juga bisa digunakan untuk mempelajari pandangan individu, perilaku, perasaan terhadap orang lain, diri sendiri, masalah yang berkembang serta kegiatan yang selama ini sudah dilakukan. Informasi ini menjadi hal yang menarik bagi para pengambil keputusan sebab mampu memastikan apa yang sedang dipikirkan oleh orang lain[7].

Teknik klasifikasi ini kemudian berkembang menjadi suatu algoritma klasifikasi. Dari sekian algoritma yang sering

digunakan tersebut adalah Naive Bayes Classifier dan Decision Tree [8][9]. Pada penelitian ke media Twitter sebelumnya, Naive Bayes Classifier menghasilkan Nilai Precision 74,5 persen, Recall 73,5 persen, Accuracy 73,9 persen [10].

Pada penelitian di media sosial Twitter, algoritma Decision Tree[11][10] menghasilkan Nilai Precision 52,6 persen, Recall 53 persen, Accuracy 52,9 persen. Pada penelitian sebelumnya di media Twitter, K-NN mampu menghasilkan Nilai Precision 62,8 persen, Recall 62,9 persen, Accuracy 62,9 persen[10].

Dalam penerapannya, algoritma klasifikasi ini seringkali kurang menghasilkan solusi yang baik bagi masalah yang dihadapi. Untuk itu adanya suatu optimasi atas algoritma klasifikasi diharapkan mampu menghasilkan tingkat solusi yang lebih baik[12]. Ada beberapa algoritma optimasi yang memungkinkan dapat digunakan dalam mengoptimasi algoritma klasifikasi salah satunya yaitu Algoritma Bat atau Bat Algorithm (BA)[8][13][14].

II. TINJAUAN PUSTAKA

2.1 Penelitian-penelitian terdahulu

Penelitian Rozlini, Munirah dan Noorhaniza, Melakukan pengujian tingkat efektivitas pengolahan data untuk Algoritma Bat melalui tipe data dan ukuran atribut dalam berbagai dataset menggunakan algoritma Naive Bayes, Decision Tree dan K-NN[8].

Penelitian Khurana dan Sanjib Kumar Sahu, mendalami tentang sentimen analisis pada data twitter menggunakan optimasi Algoritma Bat dengan algoritma klasifikasi yang digunakannya adalah *Support Vector Machine*[15].

Penelitian Bilal, Shahid dan Amin Khan, tentang klasifikasi sentimen berdasarkan algoritma Naive Bayes, Decision Tree dan K-NN [16].

2.2 Text Mining

Text mining juga dikenal sebagai *data mining* teks[17] atau penemuan pengetahuan dari database tekstual[18]. Sesuai dengan buku *The Text Mining Handbook*[19], Tujuan dari *text mining* adalah untuk mendapatkan informasi yang berguna dari sekumpulan dokumen[20].

2.3 Term Weighting TF-IDF / Pembobotan Kata

Pada data yang kelasnya besar, skema yang paling sukses dan secara luas digunakan untuk pemberian bobot term adalah skema pembobotan atau **Term Weighting TF-IDF**. Kelemahan scoring dengan Jaccard coefficient adalah tidak disertakannya frekuensi suatu term dalam suatu dokumen, maka diperlukan scoring dengan kombinasi **Term Weighting TF-IDF**.

1. Term Frequency (TF)

TF (Term Frequency) adalah frekuensi dari kemunculan sebuah term dalam dokumen yang bersangkutan. Semakin besar jumlah kemunculan suatu term (TF tinggi) dalam dokumen, semakin besar pula bobotnya atau akan memberikan nilai kesesuaian yang semakin besar.

$$TF = \begin{cases} 1 + \log_{10}(f_{t,d}), & f_{t,d} > 0 \\ 0, & f_{t,d} = 0 \end{cases} \quad (1)$$

$$TF = 0.5 + 0.5 \times \left[\frac{f_{t,d}}{\max\{f_{i,d}: i \in d\}} \right] \quad (2)$$

2. Inverse Document Frequency (IDF)

Sebuah perhitungan dari bagaimana term didistribusikan secara luas pada koleksi dokumen yang bersangkutan.

$$IDF_j = \log\left(\frac{D}{df_j}\right) \quad (3)$$

$$w_{ij} = tf_{ij} \times idf_j \quad (4)$$

$$w_{ij} = tf_{ij} \times \log\left(\frac{D}{df_j}\right) \quad (5)$$

$$w_{ij} = tf_{ij} \times \log\left(\frac{D}{df_j}\right) + 1 \quad (6)$$

2.4 Algoritma Optimasi

Algoritma[21] Optimasi[22] (Optimization Algorithms) dapat didefinisikan sebagai algoritma untuk menemukan nilai x sedemikian hingga nilai f(x) bernilai sekecil atau sebesar mungkin untuk suatu fungsi f yang diberikan, yang dalam hal ini mungkin dapat disertai dengan beberapa batasan pada nilai x. Nilai x sendiri dapat berupa skalar atau vektor dari nilai-nilai kontinu maupun diskrit[23].

2.5 Sentimen Analisis

Analisis sentimen adalah alat untuk memproses koleksi hasil pencarian yang bertujuan dengan mencari atribut suatu produk (kualitas, fitur, dll) dan proses memperoleh hasil pendapatnya. Tugas dasar dalam analisis sentimen adalah mengelompokkan polaritas dari teks yang ada dalam dokumen, apakah pendapat yang dikemukakan dalam dokumen bersifat positif, negatif atau netral[24].

2.6 Naive Bayes

Penggunaan metode *Naive Bayes Classifier* pada penelitian ini didasarkan pada banyaknya dataset yang dipakai sehingga membutuhkan suatu metode yang mempunyai performansi yang cepat dalam pengklasifikasian serta keakuratan yang cukup tinggi[25].

Keuntungan penggunaan *Naive Bayes Classifier* adalah metode ini hanya membutuhkan jumlah data pelatihan (*training data*) yang kecil untuk menentukan estimasi parameter yang diperlukan dalam proses pengklasifikasian[26]. Teorema *Naive Bayes* dapat dinyatakan dalam persamaan 7 dibawah ini[27] :

$$P(X_k|Y) = \frac{P(Y|X_k)}{\sum_i P(Y|X_i)} \quad (7)$$

2.7. Decision Tree

Decision Tree dibuat dengan memisahkan ruang vektor dari fitur dokumen ke dalam bagian-bagian terpisah secara interaktif[28]. *Decision Tree classifier* menghasilkan sebuah hirarki ruang ketidaktepatan dari data training dimana suatu kondisi dari nilai atribut

digunakan untuk memilahkan data. Pembagian data dikatakan selesai hingga titik daun berisi nilai minimum dari *record* sehingga dapat digunakan untuk menunjukkan maksud dari pengelompokan[29].

2.8. **K-Nearest Neighbor (K-NN)**
 Klasifikasi K-Nearest Neighbor (K-NN)

adalah suatu instan yang didasarkan pada pengelompokan pada label klas dari data training yang sesuai dengan data uji. Bobot jumlah dalam K-NN yang dikelompokkan dapat ditulis dengan persamaan yang ditunjukkan pada rumus 8 dibawah ini[30].

$$Score(d, c) = \sum_{d_j \in KNN(d)} sim(d, d_j) \delta(d_j, c_i) \quad (8)$$

2.9. **Algoritma Bat**

Salah satu algoritma heuristik baru yang diadaptasi dari perilaku echolocation kelelawar dalam mencari makanan. Kemampuan echolocation ini membuat kelelawar dapat membedakan rintangan dengan sumber makanan.

Dalam Algoritma Bat, ada dua hal penting yakni kebisingan dan *pulse rate*. Kebisingan untuk mendeteksi jauh dekatnya jarak dengan sumber makanan. Semakin dekat dengan sumber makanan maka tingkat kebisingan akan menurun. *Pulse rate* merupakan sinyal yang didapat dari pantulan emisi sonar yang dimiliki kelelawar terhadap rintangan. Semakin dekat kelelawar dengan rintangan maka *pulse rate* semakin meningkat. Solusi yang lebih minimum berusaha dicari melalui proses *movement* dan *local search*[31][15]. Berikut ini *pseudocode* dari Algoritma Bat [32] :

Kondisi awal:

Fungsi obyektif (x), $x = (x^1 \dots x^n)^t$.

Nilai awal jumlah kelelawar x_i dan v_i , $i = 1, 2 \dots n$.

While $t < T$

For each bat x_i , do

Solusi menggunakan Persamaan (1), (2) dan (3).

If $rand > r_i$, maka

Pilih solusi di antara solusi terbaik.

Solusi lokal di sekitar solusi terbaik.

If $rand < A_i$ dan $f(x_i) < (x^*)$, maka

Terima solusi baru.

Tambah r_i dan kurangi A_i .

Beri peringkat kelelawar dan temukan yang terbaik saat ini x^* .

Tentukan frekuensi pulsa f_i pada x_i , $\forall i = 1, 2 \dots n$.

Inisialisasi denyut nadi r_i dan kenyaringan A_i , $i = 1, 2 \dots n$.

III. METODE PENELITIAN

Berikut ini bagan dari alur penelitian seperti yang ditunjukkan pada Gambar 1.



Gambar 1 Alur Penelitian

IV. HASIL DAN PEMBAHASAN

4.1. Crawling Data Tweet (Twitter search API)

Berikut ini contoh hasil crawling pada RapidMiner ditunjukkan pada Gambar 2.

Id	User-Id	To-User	To-User-Id	Language	Source	Text	Geo-Local	Geo-Local Re
8129942				-1	in	<a href="T" ini Tawaran Jokowi agar Purnomo Legowo Gibran Maju Pilkada Solo https://		
6882261				-1	in	<a href="T" ikut Pilkada adalah hak warga negara. Tapi kalo anak presiden ikut pilkada		
17128979				-1	in	<a href="T" Gibran Berpotensi Lawan Kotak Kesong di Pilkada Kota Solo https://t.co/DI		
17000994	BarathaA1	290744031		in		<a href="T" @BarathaA1mdjia @SuburBur bukan berlega ga mau tapi emg itu TATA K		
295897062				-1	in	<a href="T" RT @Geloraco: Don Adam Usulkan Jokowi Keluarkan Kepres untuk Tetap		
225205043				-1	in	<a href="T" RT @jansen, Jip: Maa malu ya sikhank sja. Tapi, tak usahlah eranya istan		
4418108712	1743745			-1	in	<a href="T" KPU Surabaya mulai mendata pemilih pada Pilwali. <a href="T" <u>Pulwali 2023</u>. Rendat		
816319113				-1	in	<a href="T" RT @Geloraco: Don Adam Usulkan Jokowi Keluarkan Kepres untuk Tetap		
119301632597682496				-1	in	<a href="T" RT @Hilmi28: Saya sdg menunggu komentar para pendekar demokrasi, par		
119213778717810688				-1	in	<a href="T" RT @Geloraco: Don Adam Usulkan Jokowi Keluarkan Kepres untuk Tetap		
11203801454085529	DonAdam	111251471		in		<a href="T" @DonAdam68 Biarkan P Poer ikut Pilkada, rakyat solo yg akan memilih pe		
1119044712				-1	in	<a href="T" RT @Hilmi28: Saya sdg menunggu komentar para pendekar demokrasi, par		
119301632597682496				-1	in	<a href="T" RT @Geloraco: Gibran ikut Pilkada, SBY Terbukti Lebih Baik Ketimbang Joko		
34888890703037517				-1	in	<a href="T" Ayo kita sukseskan Pilkada 2020 dengan aman, damai dan sejuk dan tanpa		
112396623232929665				-1	in	<a href="T" RT @Geloraco: Don Adam Usulkan Jokowi Keluarkan Kepres untuk Tetap		
82661259				-1	in	<a href="T" RT @rensandabachar: Jika guru kencong berlari, bagaimana muridnya lagi?		
24888890703037517				-1	in	<a href="T" Ayo kita sukseskan Pilkada 2020 dengan aman, damai dan sejuk dan tanpa		
17179484	bungtree	7007047		in		<a href="T" @bungtree DKI ga gelar pilkada 2022... ngelanjut pilkada serentak 2024klo		
69561471930449321				-1	in	<a href="T" RT @BennyHarmanid: @Geloraco Gibran maju itu haknya seperti warganeg		
523375041				-1	in	<a href="T" RT @demosoCRAZY: Ini mgkn salah satu saran terbaik uk Pak Jokowi tebak		
1127079177590725122				-1	in	<a href="T" RT @kafradikalik: ...Gw begadang semalaman smpsa saat ini belum tidur t		
6185165969400490				-1	in	<a href="T" Dia juga beresin kepada para kader PAN di Pesawaran, agar berusaha dir		
062148535729601888				-1	in	<a href="T" RT @hafizulharyung: Rekomendasi itu datang dari partai bukan @jokowi ji		

Gambar 2 Hasil Crawling data Twitter dari RapidMiner

4.2. Analisis Sumber Data

Pada tahap ini akan dilakukan penghapusan data kembar dan pemilahan kolom data yang relevan dengan kebutuhan penelitian ini.

1. Penghapusan data kembar
2. Cleansing
3. Labeling

4.3. PreProcessing Data

Dalam tahap ini dilakukan proses *case folding*, *tokenizing*, *stop word removal*, *stemming*.

Pada penelitian ini proses Pembobotan Kata disusun dalam tahap sebagai berikut:

1. Menyiapkan data hasil proses *tokenizing*, *stopwords removal* dan *stemming*.
2. Menghitung Term Weighting TF-IDF / Pembobotan Kata

Setelah melakukan proses *tokenizing* pada dokumen kemudian lakukan proses perhitungan pembobotan kata dengan langkah-langkah sebagai berikut:

- a. Lakukan perhitungan jumlah frekuensi kata (misalkan kata pilkada) pada dokumen 1, 2 dan 3 sehingga didapatkan nilai d yaitu 1,1,0
- b. Lakukan penjumlahan pada frekuensi kata selain 0 untuk mendapatkan **df**
- c. Lakukan pembagian **D** dibagi dengan **df**
- d. Hitung nilai IDF dengan menghitung \log_{10} pada hasil pembagian D/df
- e. Tambahkan angka 1 pada hasil IDF
- f. Lakukan proses nilai **W** dengan cara mengkalikan nilai **tf** atau **D** dengan $IDF + 1$
- g. Jumlahkan **W** pada masing-masing dokumen
- h. Hasil dari proses Pembobotan kata ini dapat dilihat pada Tabel 1.

Tabel 1 Perhitungan Pembobotan Kata

Q	D				df	D/df	IDF	W=tf * (IDF+1)				
	1	2	3	4				d1	d2	d3	d4	
pilkada	1	1	0	1	3	1,333	0,125	1,125	1,125	1,125	-	1,125
jabat	0	0	0	0	-	-	-	1,000	-	-	-	-
kawal	0	1	0	0	1	4,000	0,602	1,602	-	1,125	-	-
kepala	0	0	1	0	1	4,000	0,602	1,602	-	-	1,125	-
daerah	0	0	1	0	1	4,000	0,602	1,602	-	-	1,125	-
kader	1	0	0	0	1	4,000	0,602	1,602	1,125	-	-	-
demokras	0	1	0	0	1	4,000	0,602	1,602	-	1,125	-	-
proses	0	1	0	0	1	4,000	0,602	1,602	-	1,125	-	-
politik	0	0	2	0	1	4,000	0,602	1,602	-	-	2,250	-
Total Nilai Bobot Kata setiap Dokumen =								2,250	4,500	4,500	1,125	

4.4 Klasifikasi dan Akurasi

Dalam melakukan proses Klasifikasi dan Akurasi ini, kita akan membagi menjadi 2 skenario percobaan. Skenario pertama adalah membuat suatu klasifikasi dan akurasi atas hasil preprocessing tanpa optimasi. Skenario kedua adalah membuat suatu klasifikasi dan akurasi setelah dilakukan suatu optimasi dengan algoritma Bat. Berikut ini hasil dari 2 skenario tersebut.

4.4.1. Skenario Pertama

Pada skenario pertama ini, kita akan melakukan pengklasifikasian data menggunakan Algoritma Naive Bayes, Decision Tree dan K-NN tanpa menggunakan proses optimasi.

1. Klasifikasi pada Naïve Bayes Pada proses pengklasifikasian text mining menggunakan Algoritma Naïve Bayes ini dibutuhkan langkah-langkah sebagai berikut :

- a) Persiapan dokumen yang telah berlabel (Positif, Negatif, Neutral)
- b) Rekap tabel berlabel dengan nilai pembobotan kata.
- c) Hitung persentase tiap kolom ²⁶ klasifikasi Hasil klasifikasi Naïve Bayes ditunjukkan pada Tabel 2.

Tabel 2. Naive Bayes Model

Sentimen	POSITIF	NEGATIF	NEUTRAL
P(POSITIF/NEGATIF/NEUTRAL)	20%	10%	70%
TOTAL	100%		
Kata	POSITIF	NEGATIF	NEUTRAL
pilkada	50%	100%	14%
jabat	0%	0%	0%
kawal	0%	0%	14%
kepala	0%	0%	14%
daerah	0%	0%	14%
kader	50%	0%	0%
demokrasi	0%	0%	14%
proses	0%	0%	14%
politik	0%	0%	14%
TOTAL	100%	100%	100%
Dokumen	POSITIF	NEGATIF	NEUTRAL
D1	100%	0%	0%
D2	0%	0%	57%
D3	0%	0%	43%
D4	0%	100%	0%
TOTAL	100%	100%	100%
Bobot	POSITIF	NEGATIF	NEUTRAL
1	100%	100%	86%
2	0%	0%	14%
TOTAL	100%	100%	100%

- d) *Testing Data Model* ²³ Contoh dari tabel data uji/Testing Data yang ditunjukkan pada Tabel 3.

Tabel 3 Training Data Model

Testing data	Kata	Dokumen	Bobot	Sentimen
	kader	D1	1	POSITIF
	kawal	D2	1	NEUTRAL
	proses	D2	1	NEGATIF
	pilkada	D2	1	NEUTRAL
pilkada	D4	1	NEGATIF	

- e) *Prediction Model* Hasil dari *Prediction Model* ini ditunjukkan pada Tabel 4.

Tabel 4 Prediction Model

Prediction	CLASS PREDICTION	POSITIF	NEGATIF	NEUTRAL
	POSITIF	10,0%	0,0%	0,0%
	NEUTRAL	0,0%	0,0%	4,9%
	NEUTRAL	0,0%	0,0%	4,9%
	NEUTRAL	0,0%	0,0%	4,9%
NEGATIF	0,0%	10,0%	0,0%	

22

f) **Confusion Matrix**

Hasil dari *confusion matrix* ini ditunjukkan pada Tabel 5.

Tabel 5 Confusion Matrix Tabel

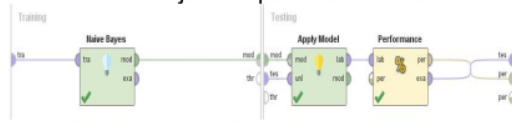
Confusion Table	Predicted	CLASS		
		POSITIF	NEGATIF	NEUTRAL
	POSITIF	1	0	0
	NEGATIF	0	1	0
	NEUTRAL	0	1	2
Accuracy =		80%		

g) **Hitung Akurasi**

Dalam simulasi ini didapatkan hasil akurasi yaitu 80 % seperti yang ditunjukkan pada tabel 5 di atas.

h) **Hasil RapidMiner**

Untuk akurasi pada penelitian ini, menggunakan aplikasi RapidMiner dengan operator *Cross Validation*. Proses *Cross Validation* ditunjukkan pada Gambar 3.



Gambar 3 Klasifikasi Naive Bayes dan akurasi

Hasil akurasi dari Naive Bayes senilai 33,58 % ditunjukkan pada Gambar 4.

accuracy: 33.58% +/- 1.19% (micro average: 33.58%)

	true pos	true neu	true neg	class precision
pred_pos	1575	1600	1594	33.03%
pred_neu	1462	1458	1445	33.40%
pred_neg	1419	1374	1463	34.38%
class recall	35.35%	32.90%	32.50%	

Gambar 4 Hasil akurasi dari Naive Bayes

2. **Klasifikasi pada Decision Tree**

Pada proses pengklasifikasian *text mining* menggunakan Algoritma *Decision Tree* ini dibutuhkan langkah-langkah sebagai berikut :

a) **Persiapan Training Data**

Pada tahap awal implementasi *Decision Tree* adalah mempersiapkan *Training Data*. Pada *Training Data* ini disiapkan sebanyak 13 record data dan 3 atribut yaitu *Kata*, *Dokumen*, *TF* (*Pembobotan Kata*) serta 1 label yaitu *Sentimen*.

b) **Perhitungan Data Uji/Testing Data**

Dalam data uji pada *Decision Tree* ditambahkan perhitungan *Entropy* dan *Gain*. Rumus hitung *Entropy* dalam pengujian ini adalah $Entropy(S) = \sum_{i=1}^n -p_i * \log_2 p_i$ dan

rumus *Gain* adalah $Gain(S, A) = Entropy(S) - \sum_{Sv} \frac{Sv}{S} Entropy(Sv)$.

Berikut ini tabel data uji/Testing Data yang ditunjukkan pada Tabel 6.

Tabel 6 Data Uji/Testing Data Decision Tree

	Jumlah	POSITIF	NEGATIF	NEUTRAL	Entropy	Gain
Kata	Total	13	4	2	7	1,419556
	pikada	1	1	0	0	0
	jabat	3	1	1	1,584962507	1
	kawal	2	1	0	1	1
	ke pisa	1	0	1	0	0
	daberah	2	0	0	2	0
	kader	1	0	0	1	0
	demokrasi	1	0	0	1	0
	proses	1	1	0	0	0
	politik	1	0	0	1	0
	Total					0,8999
Dokumen	D1	3	2	0	1	0,918295834
	D2	5	2	0	3	0,970950594
	D3	3	0	0	3	0
	D4	2	0	2	0	0
	Total					0,9342
TF	1	8	3	1	4	1,405639062
	2	5	1	1	3	1,370950594
	Total					0,0273

Pada hitung data uji ini, dicari nilai *Gain* yang terbesar, kemudian dicari nilai *Entropy* yang selain nol. Apabila ada nilai *Entropy* > 0 maka akan dibuatkan tabel data latih baru sesuai dengan jumlah atribut yang memiliki nilai *entropy* > 0.

Dalam simulasi ini didapatkan atribut *Kata* memiliki *Gain* yang paling besar, yaitu 0,8999 dengan *Kata jabat* dan *kawal* yang memiliki nilai *entropy* > 0 yaitu 1,5849625007 dan 1. Untuk itu dibuatkan data training baru yang berisi tentang kata *jabat*. Berikut ini hasil data training kata *jabat* yang ditunjukkan pada Tabel 7.

Tabel 7 Data Latih/Training Data Kata Jabat

	Kata	Dokumen	TF	Sentimen
jabat	jabat	D4	1	NEGATIF
	jabat	D1	2	POSITIF
	jabat	D1	1	NEUTRAL

Berikut ini hasil simulasi tabel data uji dari kata jabat yang ditunjukkan pada Tabel 8.

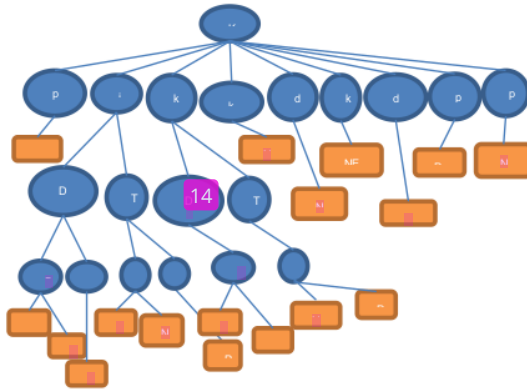
Tabel 8 Data Uji/Testing Data Kata Jabat

	Jml	POS	NEG	NEU	Entropy	Gain
Total	13	4	2	7	1,419556	
Dokumen	D1	2	1	0	1	1
	D2	0	0	0	0	0
	D3	0	0	0	0	0
	D4	1	0	1	0	0
	Total					1,265
TF	1	2	0	1	1	1
	2	1	1	0	0	0
	Total					1,265

Proses pembuatan data uji dan data latih ini akan berakhir apabila tidak ditemukan *gain*

yang lebih besar dari yang lain. Jikalau ternyata ditemukan gain yang lebih besar maka harus diperiksa nilai dari entrophynya, apabila nilai semua entrophy = 0 maka proses berakhir.

c) Desain Pohon Keputusan/Decision Tree
 Pada simulasi ini dihasilkan gambar desain *Decision Tree* yang ditunjukkan pada Gambar 5.



Gambar 5 Desain Decision Tree

Susunan desain ini diawali dari atribut yang memiliki gain terbesar menuju atribut dengan gain terkecil. Hal ini ditandai dengan simbol bulat biru, sedangkan untuk hasil sentimennya ditandai oleh persegi panjang warna jingga.

d) Hasil RapidMiner

Tahap selanjutnya adalah implementasi pada aplikasi RapidMiner. Hasil akurasi dari *Decision Tree* senilai 33,61 % ditunjukkan pada Gambar 6.

accuracy: 33.61% +/- 0.64% (micro average: 33.61%)				
	true pos	true neu	true neg	class precision
pred pos	0	0	0	0.00%
pred neu	1	0	1	0.00%
pred neg	4455	4432	4501	33.62%
class recall	0.00%	0.00%	99.98%	

Gambar 6 Hasil akurasi dari *Decision Tree*

25

3. Klasifikasi pada K-NN

Pada proses pengklasifikasian text mining menggunakan *Algoritma K-NN* ini dibutuhkan langkah-langkah sebagai berikut :

- a) Persiapan dokumen yang telah disimbolkan dengan angka
- b) Persiapkan data uji/Testing Data
- c) Persiapkan Data Latih/Training Data

d) Hitung Nilai Ecludian Distance / K-NN Model

Dengan menggunakan rumus $d(p, q) = \sqrt{\sum_{i=1}^n (q_i - p_i)^2}$ pada tahap ini dihitung jarak terdekat dari data uji/Testing Data ke data latih/Training Data. Kemudian hitung nilai terkecil dari data jarak yang sudah dibuat apakah sesuai dengan nilai sentimen Neg(1), Neu(2), Pos(3). Pengujian ini dilakukan hingga K=15. Hasil dari tahap K-NN ditunjukkan pada Tabel 9.

Tabel 9 K-NN Model

Distance	K-NN Model				
	K = 1	K = 5	K = 9	K = 13	K = 15
7,000	0	0	0	3	3
2,449	0	2	2	2	2
7,348	0	0	0	1	1
1,414	2	2	2	2	2
8,775	0	0	0	0	1
4,123	0	0	3	3	3
3,742	0	1	1	1	1
2,449	0	2	2	2	2
3,742	0	2	2	2	2
4,583	0	0	2	2	2
5,000	0	0	3	3	3
6,164	0	0	0	2	2
7,348	0	0	0	1	1
8,367	0	0	0	0	2
5,568	0	0	2	2	2

e) Hitung Nilai Persentase Sentimen Data Uji/Testing Data

Adapun tabel prediksi Sentimen data uji ditunjukkan pada Tabel 10.

Tabel 10 Persentase Sentimen Data Uji/Testing Data

	NEGATIF NEUTRAL POSITIF		
	1	2	3
P(NEGATIF/NEUTRAL/POSITIF)	27%	53%	20%
TOTAL	100%		

f) Hitung Nilai K-NN Model Prediction

Hasil dari K-NN Model Prediction ditunjukkan pada Tabel 11.

Tabel 11 K-NN Model Prediction

Distance	PERSENTASE NEGATIF NEUTRAL POSITIF			CLASS PREDICTION
	JARAK	1	2	
7,000	47%	20%	0%	27% NEUTRAL
2,449	16%	0%	0%	0% POSITIF
7,348	49%	22%	0%	29% NEUTRAL
1,414	9%	0%	0%	0% POSITIF
8,775	58%	32%	5%	38% NEUTRAL
4,123	27%	1%	0%	7% NEUTRAL
3,742	25%	0%	0%	5% POSITIF
2,449	16%	0%	0%	0% POSITIF
3,742	25%	0%	0%	5% POSITIF
4,583	31%	4%	0%	11% NEUTRAL
5,000	33%	7%	0%	13% NEUTRAL
6,164	41%	14%	0%	21% NEUTRAL
7,348	49%	22%	0%	29% NEUTRAL
8,367	56%	29%	2%	36% NEUTRAL
5,568	37%	10%	0%	17% NEUTRAL

g) Hitung Nilai Confusion Matrix dan Akurasi
 Hasil perhitungan nilai Confusion Matrix dan Akurasi ini dapat dilihat pada Tabel 12.

Tabel 12 Confusion Matrix dan Akurasi

Confusion Table	Predicted ↓	CLASS		
		NEGATIF	NEUTRAL	POSITIF
			1	2
	NEGATIF	0	0	0
	NEUTRAL	3	4	3
	POSITIF	1	4	0
Accuracy =		27%		

h) Hasil Akurasi RapidMiner

Pada penelitian ini hasil akurasi dari K-NN didapatkan nilai 32,82 % ditunjukkan pada Gambar 7.

accuracy: 32.82% +/- 1.79% (micro average: 32.82%)

	true pos	true neu	true neg	class precision
pred pos	1482	1438	1531	32.99%
pred neu	1432	1473	1511	33.36%
pred neg	1562	1521	1480	32.14%
class recall	32.81%	33.24%	32.43%	

Gambar 7 Hasil akurasi dari K-NN

4.4.2. Skenario Kedua

Mengacu pada Gambar 1 tentang alur penelitian maka optimasi pada Algoritma Bat ini dilakukan pada hasil klasifikasi pembobotan kata pada RapidMiner atau jika dikerjakan secara manual maka optimasi diawali pada hasil Term Weighting TF-IDF / Pembobotan Kata dimana hasil dari pembobotan kata tersebut menjadi bahan olahan data pada Algoritma Bat. Tahapan proses pada skenario kedua ini dijelaskan pada tahapan berikut:

1. Pengidentifikasian kekelawar

Berikut ini tabel sebelum dilakukan pengidentifikasian kekelawar seperti yang ditunjukkan pada Tabel 13.

Tabel 13 Identifikasi kekelawar pada tabel klasifikasi kata

Bat_	1	2	3	4	5	6	7	
Sentiment	Id	Texta1	Texta2	Texta3	Texta4	Texta5	Texta6	Texta7
neg	1299792569195310000,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
neg	1300236866134790000,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
neg	1300285780498160000,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
neg	1300381180550810000,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
neg	13003806508948000,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
neg	1300380310589380000,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
neg	1300380133115620000,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
neg	1300380128711600000,0	0,0	0,0	0,0	0,1	0,0	0,0	0,0
neg	1300380077587270000,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
neg	1300380044246670000,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
neg	1300379894170280000,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
neg	1300379782207560000,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
neg	1300379694840250000,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
neg	1300379657506760000,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
neg	1300379619703440000,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
neg	1300379483694760000,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0

2. Pengolahan pada tabel optimasi

- a) Penentuan fungsi obyektif/fitness
- b) Penentuan nilai β , α , γ , ϵ , f_{min} , f_{max} , A_0
- c) Hitung Frekuensi

Dengan menggunakan rumus $f_i = f_{min} + (f_{max} - f_{min})\beta$ maka dihasilkan data frekuensi seperti pada Gambar 8.

AVN13 FREKUENSI (f) $f_i = f_{min} + (f_{max} - f_{min})\beta$

A	AVN	AVO	AVP	AVQ	AVR	AVS	AVT	AVU	AVV	AVW
Index	1	2	3	4	5	6	7	8	9	10
X	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10
Bat_	Bat_1	Bat_2	Bat_3	Bat_4	Bat_5	Bat_6	Bat_7	Bat_8	Bat_9	Bat_10
Sentiment	ada	agar	agusyudf	aja	ajang	akan	anak	anda	andrerod	sapa
total										
neg/Awal										
Iterasi										
1	0,82818	0,02707	0,97228	0,91931	0,91305	0,32636	0,10479	0,64083	0,99867	0,88945
2	0,82818	0,02707	0,97228	0,91931	0,91305	0,32636	0,10479	0,64083	0,99867	0,88945
3	0,82818	0,02707	0,97228	0,91931	0,91305	0,32636	0,10479	0,64083	0,99867	0,88945
4	0,82818	0,02707	0,97228	0,91931	0,91305	0,32636	0,10479	0,64083	0,99867	0,88945
5	0,82818	0,02707	0,97228	0,91931	0,91305	0,32636	0,10479	0,64083	0,99867	0,88945
6	0,82818	0,02707	0,97228	0,91931	0,91305	0,32636	0,10479	0,64083	0,99867	0,88945
7	0,82818	0,02707	0,97228	0,91931	0,91305	0,32636	0,10479	0,64083	0,99867	0,88945
8	0,82818	0,02707	0,97228	0,91931	0,91305	0,32636	0,10479	0,64083	0,99867	0,88945
9	0,82818	0,02707	0,97228	0,91931	0,91305	0,32636	0,10479	0,64083	0,99867	0,88945
10	0,82818	0,02707	0,97228	0,91931	0,91305	0,32636	0,10479	0,64083	0,99867	0,88945

Gambar 8 Hasil perhitungan frekuensi

d) Hitung velocity/Kecepatan (V)

Dengan menggunakan rumus $v_i^t = v_i^{t-1} + (x_i^t - x_{*}) f_i$ maka dihasilkan tabel kecepatan kekelawar yang ditunjukkan pada Gambar 9.

TV13 KECEPATAN (V)

A	TV	TW	TX	TY	TZ	UA	UB	UC	UD	UE
Index	1	2	3	4	5	6	7	8	9	10
X	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10
Bat_	Bat_1	Bat_2	Bat_3	Bat_4	Bat_5	Bat_6	Bat_7	Bat_8	Bat_9	Bat_10
Sentiment	ada	agar	agusyudf	aja	ajang	akan	anak	anda	andrerod	sapa
total										
neg/Awal										
Iterasi										
1	2,64219	0,24132	1,1485	2,79901	0,09923	0,9832	1,59269	1,38142	0,68394	0,12992
2	6,90899	1,39083	2,28248	5,53612	0,01099	1,37426	2,39509	3,33592	3,14476	1,93924
3	11,0294	1,69547	5,13273	8,20698	0,77889	1,18141	4,15734	5,8492	6,47648	3,63565
4	14,9952	2,02396	7,94986	11,6525	3,17118	0,41026	5,15429	8,04961	11,5049	6,97653
5	19,6679	3,2139	11,5971	15,0432	7,18335	-0,08225	5,40034	9,95677	18,2262	11,8971
6	25,4975	5,24356	16,4583	17,9709	12,386	-0,44519	4,86003	12,1184	26,1721	16,3429
7	31,327	7,27321	22,1359	20,0822	18,405	0,00824	5,13608	15,0963	34,1179	19,9723
8	36,6614	10,1189	28,9841	23,3443	25,5725	0,5841	5,45145	18,3347	41,6325	24,7413
9	43,4584	12,1901	37,4122	26,5509	32,6794	2,2193	5,85127	22,0495	50,7482	30,227
10	51,0451	15,0488	45,8403	30,5451	39,7864	3,85449	6,25109	26,5719	59,0764	34,9253

Gambar 9 Hasil perhitungan kecepatan kekelawar

e) Hasil Perhitungan Loudness/Kebisingan (A)

Dengan rumus $A_i^{t+1} = \alpha A_i^t$, $0 \leq A \leq 1$ maka dihasilkan nilai kebisingan seperti yang ditunjukkan pada Gambar 10.

AHR13 LOUDNESS/KEBISINGAN (A)

A	AHR	AHS	AHT	AHU	AHV	AHW	AHX	AHY	AHZ	AIA
Index	1	2	3	4	5	6	7	8	9	10
X	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10
Bat_	Bat_1	Bat_2	Bat_3	Bat_4	Bat_5	Bat_6	Bat_7	Bat_8	Bat_9	Bat_10
Sentiment	ada	agar	agusyudf	aja	ajang	akan	anak	anda	andrerod	sapa
total										
neg/Awal										
Iterasi										
1	0,882	0,88	0,9	0,9	0,9	0,882	0,882	0,882	0,9	0,882
2	0,86436	0,86	0,882	0,882	0,882	0,86436	0,86436	0,86436	0,9	0,86436
3	0,84707	0,85	0,86436	0,86436	0,86436	0,84707	0,84707	0,84707	0,9	0,84707
4	0,83013	0,85	0,84707	0,84707	0,84707	0,83013	0,83013	0,83013	0,9	0,84707
5	0,81353	0,85	0,83013	0,83013	0,83013	0,81353	0,81353	0,81353	0,9	0,84707
6	0,79726	0,85	0,81353	0,81353	0,81353	0,79726	0,79726	0,79726	0,9	0,84707
7	0,78131	0,85	0,79726	0,79726	0,79726	0,78131	0,78131	0,78131	0,9	0,84707
8	0,78131	0,85	0,78131	0,78131	0,78131	0,79726	0,76569	0,76569	0,9	0,84707
9	0,78131	0,85	0,76569	0,76569	0,76569	0,79726	0,75037	0,75037	0,9	0,84707
10	0,78131	0,85	0,75037	0,75037	0,75037	0,79726	0,73537	0,73537	0,9	0,84707

Gambar 10 Hasil loudness/kebisingan kekelawar

f) Hitung *Pulse emission rate*
 Dengan rumus $r_i^{t+1} = r_i^0 [1 - \exp(-\gamma t)]$ maka dihasilkan data seperti pada Gambar 11.

AAT13 Pulse Emission Rate/Detak Sonar (R)										
A	AA1	AAU	AAV	AAW	AAZ	AA1	AA2	AA3	AA4	AA5
Index	1	2	3	4	5	6	7	8	9	10
X	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10
Bat	Bat_1	Bat_2	Bat_3	Bat_4	Bat_5	Bat_6	Bat_7	Bat_8	Bat_9	Bat_10
Sentiment total	ada	agar	agusyudhaja	ajang	akan	anak	anda	androsopa		
neg/Awai iterasi	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01
1	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01
2	0,00859	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01
3	0,00859	0,01	0,00947	0,01	0,01	0,01	0,01	0,01	0,00947	0,01
4	0,00842	0,01	0,00928	0,01	0,0098	0,01	0,01	0,00928	0,01	0,01
5	0,00836	0,01	0,00921	0,01	0,00973	0,00993	0,00993	0,00921	0,01	0,01
6	0,00836	0,01	0,00919	0,01	0,00973	0,0099	0,0099	0,00919	0,01	0,01
7	0,00836	0,01	0,00918	0,01	0,00972	0,0099	0,0099	0,00919	0,01	0,01
8	0,00836	0,01	0,00918	0,01	0,00971	0,0099	0,00989	0,00919	0,01	0,01
9	0,00836	0,01	0,00917	0,00999	0,00971	0,0099	0,00989	0,00919	0,01	0,01
10	0,00836	0,01	0,00917	0,00999	0,00971	0,0099	0,00989	0,00919	0,01	0,01

Gambar 11 Hasil *pulse emission rate* kelelawar

g) Hitung update posisi kelelawar
 Dengan rumus $X_{new} = X_{old} + \epsilon A^t$ maka ditentukan posisi X yang terbaru dimana nilai Epsilon (ϵ) bernilai $-1 \leq \epsilon \leq 1$. Dengan ini hasil dari update posisi kelelawar ditampilkan pada Gambar 12.

MX13 POSISI (X) $\leftarrow X(t-1) + (\epsilon A(t-1))$												
A	MX	MY	MZ	NA	NB	NC	ND	NE	NF	NG	NH	NI
Index	1	2	3	4	5	6	7	8	9	10	11	12
X	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12
Bat	Bat_1	Bat_2	Bat_3	Bat_4	Bat_5	Bat_6	Bat_7	Bat_8	Bat_9	Bat_10	Bat_11	Bat_12
Sentiment total	ada	agar	agusyudhaja	ajang	akan	anak	anda	androsopa	artis	asin		
neg/Awai iterasi	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01
1	2,27521	0,24403	1,24597	2,89116	0,13077	0,10582	1,6032	1,44566	0,78395	0,21908	1,6446	0,22766
2	3,60921	1,12803	0,36197	2,00716	-0,75323	0,13192	0,7192	1,44566	1,66795	1,10308	2,5286	1,11166
3	2,74073	0,25955	1,23045	1,13868	-0,73656	1,58768	1,44566	1,66795	0,2346	2,5286	0,24318	0,24318
4	1,88698	0,25955	0,3787	1,13868	0,10052	-1,59031	0,73393	0,59191	2,5217	1,08835	1,67485	-0,61058
5	1,88698	1,0999	0,3787	0,29833	0,94087	-1,59031	-0,10643	-0,24844	3,36206	1,92871	0,83449	-1,45093
6	2,71495	1,92787	1,20486	-0,52964	1,76884	-1,59031	-0,9344	-0,24844	4,19002	1,10074	0,00633	-2,2789
7	2,71495	1,92787	2,02104	-1,34601	2,58521	-0,77394	-0,11802	0,56793	4,19002	0,28437	0,8229	-1,46253
8	1,90905	2,73796	2,82493	-0,54012	3,3911	-0,77394	-0,11802	0,56793	3,38413	1,09028	0,8229	-0,65666
9	2,70527	1,92744	3,82325	-1,33843	2,58478	0,02238	-0,11802	0,56793	4,18043	1,09028	0,02658	-1,45295
10	3,48091	1,72488	3,82325	-0,54891	2,58478	0,02238	-0,11802	1,35546	3,39393	0,30273	0,84111	-0,66542

Gambar 12 Hasil update posisi kelelawar

h) Hitung nilai fungsi obyektif / *fitness* kelelawar
 Dengan rumus $f(x) = x^2$ didapatkan hasil seperti pada Gambar 13.

B13 FUNGSI OBYEKTIF/FITNESS $f(x) = x^2$										
A	B	C	D	E	F	G	H	I	J	K
Index	1	2	3	4	5	6	7	8		
X	X1	X2	X3	X4	X5	X6	X7	X8		
Bat	Bat_1	Bat_2	Bat_3	Bat_4	Bat_5	Bat_6	Bat_7	Bat_8		
Sentiment total	ada	agar	agusyudhono	aja	ajang	akan	anak	anda		
neg/Awai iterasi	13,14214161	0,059550059	0,119693909	14,3729	0,0171	1,03209	2,57024	0,29775		
1	13,14214161	0,059550059	0,119693909	14,3729	0,0171	1,03209	2,57024	0,29775		
2	7,426765085	0,059550059	0,119693909	8,35883	0,0171	1,03209	2,57024	0,29775		
3	7,426765085	0,059550059	0,119693909	4,02871	0,0171	0,0174	0,51724	0,29775		
4	7,426765085	0,059550059	0,119693909	1,2966	0,0171	0,0174	0,51724	0,29775		
5	3,560688168	0,059550059	0,119693909	1,2966	0,0101	0,0174	0,51724	0,29775		
6	3,560688168	0,059550059	0,119693909	0,089	0,0101	0,0174	0,01133	0,06172		
7	3,560688168	0,059550059	0,119693909	0,089	0,0101	0,0174	0,01133	0,06172		

Gambar 13 Hasil perhitungan fungsi obyektif / *fitness* posisi kelelawar

i) Rekap Data Fitness
 Hasil rekap iterasi ini nantinya akan dimasukkan kedalam rapidminer untuk dilakukan pengujian kembali hasil akurasi datanya. Hasil rekap iterasi ini ditunjukkan pada Gambar 14.

A1 sentiment										
A	B	C	D	E	F	G	H	I	J	K
Index	1	2	3	4	5	6	7	8	9	10
Sentimen	ada	agar	agusyudhaja	ajang	akan	anak	anda			
1	13,14214	0,05955	0,119694	14,37293	0,0171	1,032094	2,570242	0,29775		
2	7,426765	0,05955	0,119694	8,35883	0,0171	1,032094	2,570242	0,29775		
3	7,426765	0,05955	0,119694	4,289721	0,0171	0,038385	0,613398	0,29775		
4	7,426765	0,05955	0,119694	1,736502	0,0171	0,038385	0,613398	0,29775		
5	4,37027	0,05955	0,119694	1,736502	0,000145	0,038385	0,613398	0,29775		
6	4,37027	0,05955	0,119694	0,429436	0,000145	0,038385	0,029879	0,006		
7	4,37027	0,05955	0,119694	0,000491	0,000145	0,038385	0,029879	0,006		
8	4,37027	0,05955	0,119694	0,000491	0,000145	0,038385	0,022804	0,006		
9	4,37027	0,05955	0,119694	1,05E-05	0,000145	0,004246	0,022804	0,006		
10	4,37027	0,05955	0,119694	2,73E-05	0,000145	0,004246	0,022804	0,006		
11	4,37027	0,05955	0,119694	2,73E-05	0,000145	0,004246	0,022804	0,006		
12	4,37027	0,05955	0,119694	1,05E-05	0,000145	0,004246	0,022804	0,006		
13	4,37027	0,05955	0,119694	1,05E-05	0,000145	0,004246	0,022804	0,006		
14	4,37027	0,05955	0,119694	1,05E-05	0,000145	0,004246	0,022804	0,006		
15	4,37027	0,05955	0,119694	1,05E-05	0,000145	0,004246	0,021731	0,006		
16	4,37027	0,05955	0,119694	1,05E-05	0,000145	0,004246	0,021731	0,006		
17	4,37027	0,05955	0,119694	1,05E-05	0,000145	0,004246	0,021731	0,006		
18	4,37027	0,05955	0,119694	1,05E-05	0,000145	0,004246	0,021731	0,006		
19	4,37027	0,05955	0,119694	1,05E-05	0,000145	0,004246	0,021731	0,006		
20	4,37027	0,05955	0,119694	1,05E-05	0,000145	0,004246	0,021731	0,006		

Gambar 14 Rekap fungsi obyektif/fitness posisi kelelawar pada seluruh sentiment

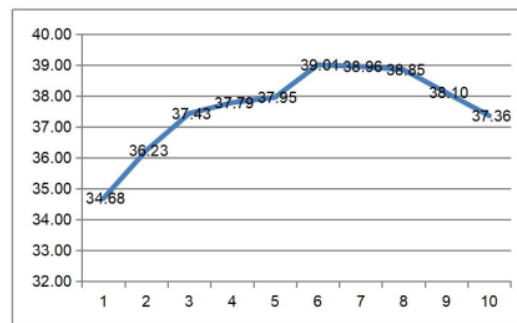
3. Hitung nilai akurasi

a) Optimasi pada Naive Bayes Classifier
 Hasil pengujian optimasi Naive Bayes ini ditunjukkan pada Tabel 13.

Tabel 13 Akurasi Naive Bayes dengan 100 Bat

Jumlah Bat	Percobaan akurasi pada fitness ke-i									
	1	2	3	4	5	6	7	8	9	10
100	34,68	36,23	37,43	37,79	37,95	39,01	38,96	38,85	38,10	37,36

Hasil penelitian menunjukkan bahwa nilai pada fitness ke-6 merupakan nilai fitness yang paling optimal dengan tingkat akurasi sebesar 39,01 %. Hal ini dapat dilihat pada Gambar 15.



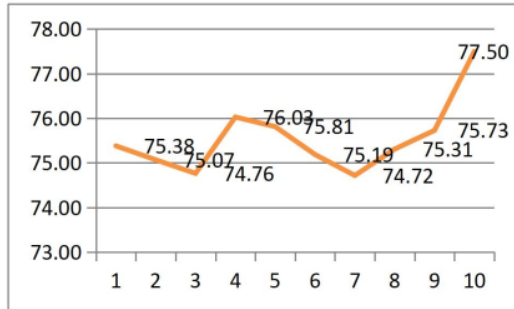
Gambar 15 Nilai Pengujian Akurasi Naive Bayes Untuk 100 Bat

b) Optimasi pada Decision Tree
 Hasil dari 10 kali percobaan optimasi pada Algoritma Bat pada Decision Tree seperti yang ditunjukkan pada Tabel 14 di bawah ini.

Tabel 14 Akurasi Decision Tree dengan 100 Bat

Jumlah Bat	Percobaan akurasi pada fitness ke-i									
	1	2	3	4	5	6	7	8	9	10
100	75,38	75,07	74,76	76,03	75,81	75,19	74,72	75,31	75,73	77,50

Hasil penelitian menunjukkan bahwa nilai pada fitness ke-10 merupakan nilai fitness yang paling optimal dengan tingkat akurasi sebesar 77,50 %. Hal ini dapat dilihat pada Gambar 16.



Gambar 16 Nilai Pengujian Akurasi Decision Tree Untuk 100 Bat

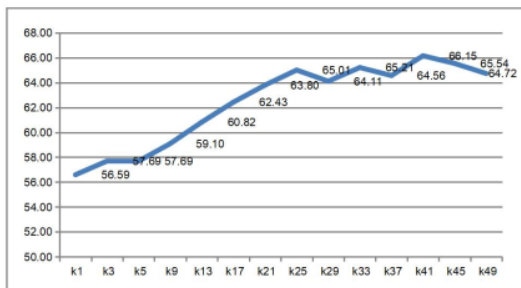
c) Optimasi pada K-NN

Dari percobaan klasifikasi data yang telah dioptimasi kemudian menggunakan RapidMiner, akurasi K-NN dengan nilai k yang digunakan dapat dilihat pada Tabel 15.

Tabel 15 Akurasi KNN dari berbagai nilai K

k	1	3	5	9	13	17	21	25	29	33	37
akurasi	56,59	57,69	57,69	59,10	60,82	62,43	63,80	65,01	64,11	65,21	64,56
k	41	45	49								
akurasi	66,15	65,54	64,72								

Hasil penelitian menunjukkan bahwa nilai K=41 merupakan nilai k yang paling optimal dengan tingkat akurasi sebesar 66,15 %. Hal ini juga dapat dilihat dalam bentuk grafik yang ditunjukkan pada Gambar 17.



Gambar 17 Nilai Pengujian Akurasi KNN

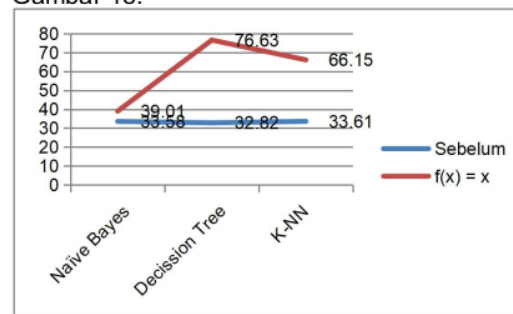
d) Rekap Hasil Optimasi

Hasil perhitungan nilai akurasi dari ketiga Algoritma klasifikasi pada penelitian ini ditunjukkan pada Tabel 16.

Tabel 16 Perbandingan hasil optimasi

Klasifikasi	Optimasi (%)	
	Sebelum	f(x)=x^2
Naive Bayes	33,58	39,01
Decision Tree	32,82	76,63
K-NN	33,61	66,15

Berikut ini hasil perbandingan optimasi dalam bentuk grafik seperti yang ditunjukkan pada Gambar 18.



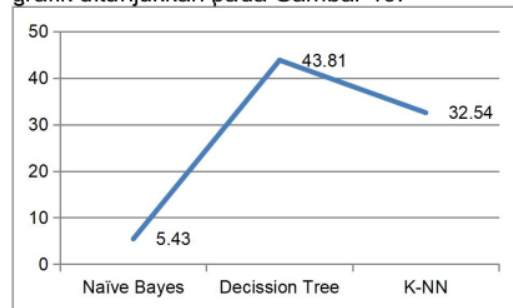
Gambar 18 Grafik perbandingan akurasi hasil optimasi dan sebelum optimasi

Dalam tabel dan gambar perbandingan hasil optimasi diatas menunjukkan bahwa selisih perhitungannya ditunjukkan pada tabel 17.

Tabel 17 Perbandingan selisih hasil optimasi dibandingkan sebelum optimasi

Klasifikasi	f(x)=x^2 - Sebelum Optimasi
Naive Bayes	5,43
Decision Tree	43,81
K-NN	32,54

Berikut ini hasil tampilan grafik dari selisih peningkatan hasil optimasi dibandingkan dengan sebelum dilakukan optimasi. Gambar grafik ditunjukkan pada Gambar 19.



Gambar 19 Selisih akurasi hasil optimasi dibandingkan sebelum optimasi

V. PENUTUP

5.1 Kesimpulan

Pada penelitian ini metode optimasi menggunakan algoritma Bat menunjukkan suatu peningkatan nilai akurasi yang cukup signifikan dibandingkan sebelum optimasi pada kasus algoritma Naïve Bayes, *Decision Tree* dan K-NN. Dari hasil pengujian diatas menunjukkan kesimpulan sebagai berikut :

1. Pada hasil perbandingan peningkatan optimasi, didapatkan bahwa tiap algoritma klasifikasi mengalami peningkatan akurasi dengan menggunakan fitness $f(x) = x^2$ yaitu mulai dari 5,43% hingga 43,81%
2. Pada hasil simulasi akurasi didapatkan Algoritma *Decision Tree* menempati selisih yang paling tinggi yaitu 43,81% dari sebelum optimasi. Hal ini bisa disimpulkan bahwa dalam penelitian ini algoritma keelawar cocok digunakan untuk mengoptimasi algoritma klasifikasi *Decision Tree* selanjutnya diikuti KNN (32,54 %) kemudian Naïve Bayes (5,43 %).

5.2 Saran

Pada penelitian selanjutnya, saran yang dapat diberikan adalah :

1. Para peneliti juga bisa mencoba persamaan fungsi yang lain pada fungsi obyektif/fitness pada algoritma Bat.
2. Para peneliti juga dapat melakukan iterasi > 150 kali pada tiap sentimennya serta jumlah Bat > 100 sehingga bisa menghasilkan kesimpulan uji yang lebih kompleks.

DAFTAR PUSTAKA

- [1] P. Yudha, "APJII: Jumlah Pengguna Internet di Indonesia Tembus 171 Juta Jiwa," *Kompas.com*, 2019. <https://tekno.kompas.com/read/2019/05/16/03260037/apjii-jumlah-pengguna-internet-di-indonesia-tembus-171-juta-jiwa>
- [2] We Are Social and Hootsuite Survei, "Riset Ungkap Pola Pemakaian Medsos Orang Indonesia," *Kompas.Com*, 2018. <https://tekno.kompas.com/read/2018/03/01/10340027/riset-ungkap-pola-pemakaian-medsos-orang-indonesia>
- [3] H. David, *Teori Komunikasi, Media, Teknologi, Dan Masyarakat*, vol. III, no. 3. Bogor: Ghalia Indonesia, 2014.
- [4] Iskandar, "Pilkada Serentak di Indonesia Hasilkan 370.000 Tweets, 13 Desember 2015," 2015. <https://www.liputan6.com/tekno/read/2388665/pilkada-serentak-di-indonesia-hasilkan-370000-tweets>
- [5] Charlie M. Sianipar, "Jumlah Pengguna Twitter Indonesia Naik Pesat," *Tagar.Id*, 2019. <https://www.tagar.id/jumlah-pengguna-twitter-indonesia-naik-pesat>
- [6] P. P. O. Mahawardana, I. A. P. F. Imawati, and I. W. Dika, "Analisis Sentimen Berdasarkan Opini dari Media Sosial Twitter terhadap 'Figure Pemimpin' Menggunakan Python," *J. Manaj. dan Teknol. Inf.*, vol. 12, no. 2, pp. 50–56, 2022, [Online]. Available: <https://ojs.mahadewa.ac.id/index.php/jmti/article/view/2111>
- [7] 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," in *Procedia Engineering*, 2013, vol. 53, pp. 453–462. doi: 10.1016/j.proeng.2013.02.059.
- [8] R. Mohamed, M. M. Yusof, and N. Wahid, "The effectiveness of Bat algorithm for data handling in various applications," in *Proceedings - 6th IEEE International Conference on Control System, Computing and Engineering, ICCSCE 2016*, 2017, pp. 151–156. doi: 10.1109/ICCSCE.2016.7893562.
- [9] H. Muhamad, C. A. Prasajo, N. A. Sugianto, L. Surtiningsih, and I. Cholissodin, "Optimasi Naïve Bayes Classifier Dengan Menggunakan Particle Swarm Optimization Pada Data Iris," *J. Teknol. Inf. dan Ilmu Komput.*, vol. 4, no. 3, p. 180, 2017, doi: 10.25126/jtiik.201743251.
- [10] A. Kumar and A. Jaiswal, "Empirical study of twitter and tumblr for sentiment analysis using soft computing techniques," in *Lecture Notes in Engineering and Computer Science*, 2017, vol. 1, pp. 472–476.
- [11] H. Sumarmo, "Komparasi Algoritma Klasifikasi Machine Learning," *BINA Insa. ICT J.*, vol. VOL 4, no. ISSN : 2527-9777, pp. 189–196, 2017.
- [12] S. V. and M. M., "Review: Sentiment Analysis using SVM Classification Approach," *Int. J. Comput. Appl.*, vol. 181, no. 37, pp. 1–8, 2019, doi: 10.5120/ijca2019917993.
- [13] J. C. Zavala-Díaz, M. A. Cruz-Chávez, J. López-Calderón, J. A. Hernández-Aguilar, and M. E. Luna-Ortiz, "A multi-branch-and-bound binary parallel algorithm to solve

- the knapsack problem 0-1 in a multicore cluster," 2019. doi: 10.3390/app9245368.
- [14] A. Chandra, I. M. Widiartha, and A. Muliantara, "Analisis & implementasi algoritma kekeluar sebagai fitur selektor dalam klasifikasi dermatology," *J. Ilmu Komput. Univ. Udayana*, vol. IX, no. 2, pp. 15–26, 2016.
- [15] H. Khurana and S. K. Sahu, "Bat inspired sentiment analysis of Twitter data," 2018. doi: 10.1007/978-981-10-6875-1_63.
- [16] M. Bilal, H. Israr, M. Shahid, and A. Khan, "Sentiment classification of Roman-Urdu opinions using Naïve Bayesian, Decision Tree and KNN classification techniques," *J. King Saud Univ. - Comput. Inf. Sci.*, vol. 28, no. 3, pp. 330–344, 2016, doi: 10.1016/j.jksuci.2015.11.003.
- [17] M. A. Hearst, "Text data mining: Issues, techniques, and the relationship to information access," in *Presentation notes for UW/MS workshop on data mining*, 1997.
- [18] R. Feldman and I. Dagan, "Knowledge Discovery in Textual Databases (KDD)," in *International Conference on Knowledge Discovery and Data Mining (KDD)*, 1995, pp. 112–117. [Online]. Available: <http://www.aaai.org/Papers/KDD/1995/KDD95-012.pdf>
- [19] J. Feldman, R & Sanger, *The text mining handbook: advanced approaches in analyzing unstructured data*, vol. 44, no. 10. New York: Cambridge Univ. Press, 2007. doi: 10.5860/choice.44-5684.
- [20] C. Triawati, M. A. Bijaksana, N. Indrawati, and W. A. Saputro, "Pemodelan Berbasis Konsep Untuk Kategorisasi Artikel Berita," in *Seminar Nasional Aplikasi Teknologi Informasi (SNATI)*, 2009, vol. 2009, no. Snati, pp. 48–53.
- [21] S. Roaf, *Introduction to second edition*. MIT Press, 2014. doi: 10.5325/j.ctv1c9hns8
- [22] KBBI Kemendikbud, "KBBI Daring," *Badan Pengembangan dan Pembinaan Bahasa, Kementerian Pendidikan dan Kebudayaan Republik Indonesia*, 2021. <https://kbbi.kemdikbud.go.id/entri/optimas> (accessed Jun. 12, 2020).
- [23] A. Hasad, "Algoritma optimasi dan aplikasinya," Bogor, 2011.
- [24] K. Dave, S. Lawrence, and D. M. Pennock, "Mining the peanut gallery: Opinion extraction and semantic classification of product reviews," Princeton, NJ; Pasadena, CA, 2003. doi: 10.1145/775152.775226.
- [25] D. T. Larose, "Naïve Bayes Estimation and Bayesian Networks," in *Data Mining Methods and Models*, Hoboken, NJ, USA: John Wiley & Sons, Inc., 2006, pp. 204–239. doi: 10.1002/0471756482.ch5.
- [26] M. Kubek, "Natural Language Processing and Text Mining," in *Studies in Big Data*, vol. 62, London, England: Springer Verlag, 2020, pp. 35–52. doi: 10.1007/978-3-030-23136-1_4.
- [27] J. Arguello, "Naïve Bayes Text Classification," *The University of North Carolina*, 2013. https://ils.unc.edu/courses/2013_fall/inls613_001/lectures/04NaiveBayesClassification.pdf (accessed Apr. 12, 2020).
- [28] K. Jain and S. Kaushal, "A Comparative Study of Machine Learning and Deep Learning Techniques for Sentiment Analysis," in *2018 7th International Conference on Reliability, Infocom Technologies and Optimization: Trends and Future Directions, ICRITO 2018*, 2018, pp. 483–487. doi: 10.1109/ICRITO.2018.8748793.
- [29] Jyoti and S. Rao, "A survey on sentiment analysis and opinion mining," in *ACM International Conference Proceeding Series*, 2016, vol. 12-13-NaN-2016, pp. 7–11. doi: 10.1145/2979779.2979832.
- [30] A. P. Jain and V. D. Katkar, "Sentiments analysis of Twitter data using data mining," in *Proceedings - IEEE International Conference on Information Processing, ICIP 2015*, 2016, pp. 807–810. doi: 10.1109/INFOP.2015.7489492.
- [31] M. Kumaresan, "Sentiment Analysis using Vader," DELHI TECHNOLOGICAL UNIVERSITY, 2016.
- [32] X. S. Yang, *A new metaheuristic Bat-inspired Algorithm*, vol. 284. Cambridge, UK: Springer Verlag, 2010. doi: 10.1007/978-3-642-12538-6_6.

14%

SIMILARITY INDEX

PRIMARY SOURCES

1	repository.unair.ac.id Internet	91 words — 2%
2	www.coursehero.com Internet	58 words — 1%
3	media.neliti.com Internet	56 words — 1%
4	digilib.unhas.ac.id Internet	54 words — 1%
5	informatikalogi.com Internet	50 words — 1%
6	journal.untar.ac.id Internet	49 words — 1%
7	etheses.uin-malang.ac.id Internet	47 words — 1%
8	5aed14f6-7642-4701-ac9b-d522954b0ab5.filesusr.com Internet	21 words — < 1%
9	Jefferson A. Costales, Christian M. De Los Santos, Jeff Jojer Jones E. Catulay, Michael G. Albino. "Sentiment Analysis for Twitter Tweets: A Framework to Detect	20 words — < 1%

Sentiment Using Naïve Bayes Algorithm", 2022 4th
International Conference on Computer Communication and the
Internet (ICCCI), 2022

Crossref

10	digilib.uinsby.ac.id Internet	20 words — < 1%
11	ijcseonline.org Internet	20 words — < 1%
12	ocs.unud.ac.id Internet	15 words — < 1%
13	text-id.123dok.com Internet	13 words — < 1%
14	www.astro.ucla.edu Internet	13 words — < 1%
15	openlibrarypublications.telkomuniversity.ac.id Internet	12 words — < 1%
16	api.intechopen.com Internet	11 words — < 1%
17	publishing-widyagama.ac.id Internet	11 words — < 1%
18	www.uni-ulm.de Internet	11 words — < 1%
19	www.scielo.br Internet	10 words — < 1%
20	jurnal.kharisma.ac.id Internet	9 words — < 1%

21	scopedatabase.com Internet	9 words — < 1%
22	123dok.com Internet	8 words — < 1%
23	docobook.com Internet	8 words — < 1%
24	publikasi.dinus.ac.id Internet	8 words — < 1%
25	www.ejurnal.stmik-budidarma.ac.id Internet	8 words — < 1%
26	doaj.org Internet	6 words — < 1%
27	ejournal.uksw.edu Internet	6 words — < 1%

EXCLUDE QUOTES OFF

EXCLUDE BIBLIOGRAPHY OFF

EXCLUDE SOURCES OFF

EXCLUDE MATCHES OFF