



Analisis Sentimen Masyarakat Pada Twitter Terhadap Debt Collector Menggunakan Metode Naïve Bayes Classifiers

Rizqy Arya Pratama¹, Iqbal Maulana², Oman Komarudin³

^{1,2,3}Universitas Singaperbangsa Karawang

Received: 8 Januari 2024
Revised: 16 Januari 2024
Accepted: 21 Januari 2024

Abstract

Saat ini banyak masyarakat yang membutuhkan pinjaman uang tunai. Bersumber pada hasil laporan Otoritas Jasa Keuangan (OJK) jumlah penyaluran pinjaman online telah mencapai Rp18,73 triliun sepanjang tahun 2022. Debt collector hadir untuk melakukan penagihan bagi masyarakat yang tidak mampu melunasi pinjaman yang telah diajukan. Meski begitu kinerja debt collector banyak mengalami kecaman dari masyarakat, kinerja debt collector sering menjadi perbincangan di media sosial twitter. Oleh karena itu perlu dilakukan analisis terhadap isu tentang debt collector pada twitter untuk melihat opini masyarakat tentang praktik debt collector itu sendiri. Data yang digunakan dalam penelitian ini sejumlah 600 data. Pelabelan data menggunakan Ahli Bahasa Indonesia sebagai validator untuk mengetahui kelas sentimennya. Pada tahap preprocessing data dibersihkan untuk mengurangi atribut yang kurang berpengaruh terhadap proses klasifikasi. Hasil accuracy tertinggi didapatkan dengan menggunakan Naive Bayes Classifier dan TF-IDF yaitu sebesar 78.3% dengan presentase data 90:10 (90% data latih dan 10% data uji). Untuk nilai precision tertinggi diperoleh dari pengujian dengan penggunaan data 90:10 dengan menggunakan Naive Bayes Classifier dengan metode BoW yaitu sebesar 80%. Sedangkan nilai recall tertinggi diperoleh dari pengujian dengan penggunaan data 90:10 menggunakan metode BoW yaitu sebesar 80%. Penggunaan BoW berhasil meningkatkan nilai akurasi pada sebagian besar skenario pembagian data dalam pengujian. Proses klasifikasi menghasilkan kata yang paling sering muncul pada masing-masing kelas sentimen yang divisualisasikan dengan word cloud dan diagram fishbone. Kata "kejar" merupakan kata paling dominan pada data tweet negatif, sedangkan kata "bantu" yang paling dominan pada data tweet positif terhadap debt collector pada twitter. Penggambaran menggunakan diagram fishbone memberikan penyelesaian atas pendapat dan pengalaman negatif masyarakat terhadap debt collector yaitu salah satunya penggunaan kata kasar yang bisa diatasi dengan melakukan pelatihan untuk semua anggota tim tentang etika berkomunikasi.

Keywords: Analisis Sentimen, Debt collector, Algoritma Naïve Bayes

(*) Corresponding Author: 1910631170130@student.unsika.ac.id

How to Cite: Pratama, R., Maulana, I., & Komarudin, O. (2024). Analisis Sentimen Masyarakat Pada Twitter Terhadap Debt Collector Menggunakan Metode Naive Bayes Classifier. *Jurnal Ilmiah Wahana Pendidikan*, 10(7), 929-941.

PENDAHULUAN

Uang tunai dan uang elektronik adalah alat pertukaran yang diakui di seluruh dunia. Saat ini banyak masyarakat yang membutuhkan pinjaman uang tunai untuk bertahan hidup, membeli sesuatu, ataupun untuk membayar hutang. Bahkan pada era digital sekarang akses untuk melakukan pinjaman tidak hanya berupa uang tunai (fisik) tetapi juga melakukan pinjaman uang berupa elektronik atau e-wallet yang telah banyak disediakan oleh pihak pinjaman online. Pinjaman online merupakan model bisnis berbasis internet yang memenuhi kebutuhan pinjaman

antar perantara keuangan. Platform ini ditujukan untuk perusahaan menengah dan kecil dimana menurut mereka persyaratan pinjaman bank mungkin terlalu tinggi (Hsueh., 2017).

Bersumber pada hasil laporan Otoritas Jasa Keuangan (OJK) jumlah penyaluran pinjaman *online* telah mencapai Rp18,73 triliun sepanjang bulan Januari 2023 meningkat 35,72% dari bulan Januari 2022. Dengan adanya peningkatan yang terjadi, sementara keadaan ekonomi masyarakat Indonesia yang tidak selalu stabil menjadikan pinjaman online tidak luput dari penunggakan pembayaran dikarenakan masyarakat tidak mampu untuk melunasi pinjaman yang telah diajukan. Karena penunggakan pembayaran tersebut pihak jasa keuangan atau bank melakukan penagihan secara langsung menggunakan jasa *debt collector* apabila sudah dilakukan penagihan secara daring namun tidak mendapatkan jawaban dari masyarakat yang melakukan pinjaman. *Debt collector* sendiri dapat dikatakan sebagai individu atau sekumpulan orang yang memberikan jasa dalam menagih utang seseorang atau lembaga yang menyewa jasa mereka (Kasmir, 2020).

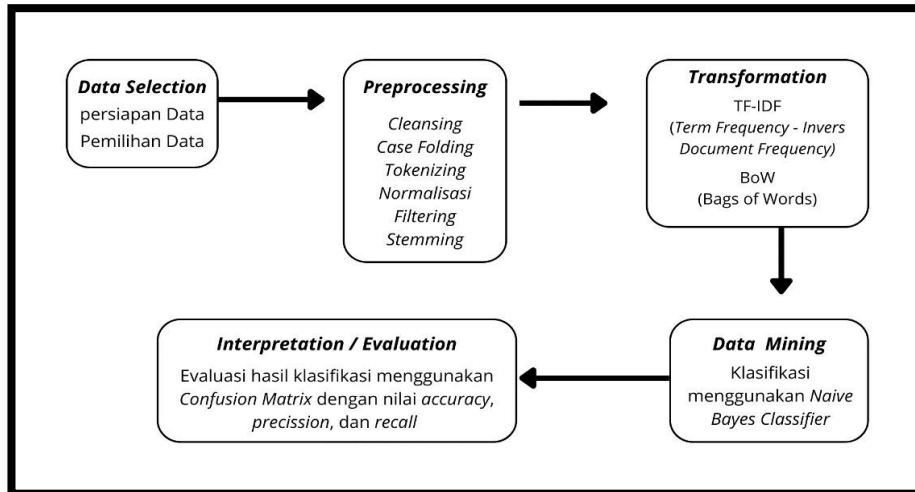
Twitter merupakan salah satu media sosial yang banyak digunakan oleh masyarakat di Indonesia, yang mana jumlah pengguna twitter di Indonesia pada Januari 2023 telah mencapai 24 juta pengguna. Hal ini menunjukkan bahwa media sosial twitter masih menjadi salah satu media sosial yang sering digunakan untuk menyampaikan pendapat terhadap suatu isu. *Debt collector* merupakan salah satu isu yang seringkali menjadi perbincangan juga di twitter, *debt collector* memang bertugas menagih utang atas nama perusahaan atau lembaga keuangan tertentu. Akan tetapi terdapat beberapa kasus dimana *debt collector* melakukan tindakan yang tidak sesuai dengan aturan dan etika dalam menjalankan tugasnya, seperti melakukan pemerasan, ancaman, atau kekerasan yang membuat keresahan di lingkungan masyarakat. Akibatnya, respon masyarakat tentang *debt collector* ini cukup beragam termasuk di media sosial twitter, yang mana terdapat respon positif dan juga respon negatif.

Berdasarkan penelitian sebelumnya yang dilakukan oleh Dedy, Ema & Ainul (2022) mengenai perbandingan kinerja model TF-IDF dan BoW untuk klasifikasi opini publik tentang kebijakan BLT minyak goreng. Dari penelitian tersebut, dapat diketahui hasil dari pengujian dua metode pembobotan yaitu TFIDF dan BoW dalam melakukan pembobotan dengan objek penelitian opini public tentang kebijakan BLT minyak goreng dengan menggunakan 90% data latih dan 10% data uji yang dimana terdiri dari 25,62% tweet positif, 14,38% tweet negatif dan 60% tweet netral dengan akurasi yang di dapatkan menggunakan metode pembobotan TF-IDF sebesar 69% dan menggunakan metode pembobotan BoW sebesar 72%.

Penelitian ini akan membahas klasifikasi sentimen masyarakat terhadap *debt collector* pada twitter menjadi dua sentimen yaitu sentimen positif dan sentiment negatif dengan metode *Naïve Bayes Classifier* dan membandingkan ekstraksi fitur TF-IDF dan BoW. Dengan harapan hasil dari penelitian ini dapat mengklasifikasikan sentimen masyarakat terhadap *debt collector* dengan baik sehingga informasi yang telah dihasilkan tersebut dapat memberikan pengetahuan tentang *debt collector* yang dapat dimanfaatkan dan digunakan oleh pihak yang membutuhkan.

METODE

Metode atau langkah-langkah untuk mencapai tujuan dari penelitian ini adalah metodologi *Knowledge Discovery in Database (KDD)*, yang terdiri dari lima langkah yaitu eleksi data, *pre-processing*, transformasi, data mining dan interpretasi/evaluasi.



Gambar 1 Alur Tahapan Penelitian

Tahap pertama yaitu menyeleksi dan mengambil data di Twitter tentang opini masyarakat terhadap agen penagihan (*debt collector*) mulai dari tanggal 1 Januari 2023 hingga 30 April 2023 dengan kata kunci yang digunakan yakni (“*debt collector*”). Ditahap *data selection* ini dilakukan seleksi data dimana teknik data *crawling* akan digunakan untuk memperoleh data yang relevan dengan isu penelitian. Tahap kedua adalah data *preprocessing* yang dilakukan untuk mengurangi atribut yang kurang berpengaruh terhadap proses klasifikasi. Tahap selanjutnya adalah data *transformation* yaitu proses pengkodean data untuk mengubah skala data asli ke bentuk lain. Pada fase ini data diubah menjadi numerik kemudian data tersebut dibagi menjadi data training dan data testing.

Langkah selanjutnya adalah *data mining*, dimana klasifikasi tanggapan masyarakat terhadap agen penagihan (*debt collector*) pada twitter yang akan dilakukan dengan menggunakan bantuan tool RStudio. Metode klasifikasi yang digunakan dalam penelitian ini adalah *Naive Bayes classifier*.

Tahap terakhir pada metodologi KDD adalah evaluasi. Pada tahap ini dilakukan pengujian metode *Naive Bayes Classifiers* terhadap hasil klasifikasi dengan menggunakan *Confusion Matrix* yang hasilnya akan diperoleh berupa nilai *accuracy*, *precision*, dan *recall*. Selanjutnya akan ditampilkan juga hasil klasifikasi ulasan sentimen positif dan sentimen negatif dengan cara memvisualisasikan katakata yang sering muncul dalam ulasan kedua sentimen tersebut dalam bentuk *word cloud* dan *diagram fishbone*.

HASIL & PEMBAHASAN

Hasil

Data Selection

Pada tahap ini akan memilih data yang sesuai untuk digunakan dalam penelitian dari data hasil *crawling* sebanyak 3.000 tweet terseleksi sebanyak 1.000 data tweet. Setelah data terkumpul selanjutnya dilakukan proses *labelling*. Proses *labelling* dengan bantuan ahli Bahasa Indonesia dilakukan terhadap 1.000 data tweet yang didapatkan dalam tahap *data selection*. Sebagian data penelitian dapat dilihat pada Gambar 2 dan jumlah data *labelling* pada Gambar 3..

1	Date	Tweet	Username
2	30/04/2023 02:22	@dontbry_@r0bbeni baiknya gausah diangkat sma lgsg di blokir aja kak, bbrp hari yg lalu aku sering di te	oliivyou
3	30/04/2023 00:57	@indarani Yang itu kayaknya tipe yang lagi ngumpet dari debt collector atau stalker deh ðŸ™ˆ	mutonk
4	29/04/2023 23:22	dia tuh model orang yg kalo temennya mau pinjem duit, dia yg pinjem pake paylater sampe didatengin	rexgazing
5	29/04/2023 22:31	@yookarinza Deg degan, apalagi kalo sambil didatengin debt collector	karrivna
6	29/04/2023 21:56	Username sih investor, tp cara pikirnya kayak debt collector ðŸ™ˆ,	LelakiFavoritmu
7	29/04/2023 17:58	mukanya kusem banget akibat dikejar debt collector ya?	NJTTWclsd
8	29/04/2023 15:54	@cursedkidd yg nagih mah debt collector bkn sales gmn sih	dijudesin
9	29/04/2023 15:18	@TM_One @unifi @helpmeunifi you guys upah Along ke nak kutip due. Debt collector cakap ikut suka je	tamizhacchi
10	29/04/2023 14:05	@mas_aro @HooHooTan @neohistoria_id Maksudnya? Orang Maluku merantau ke Jakarta jadi debt colle	dantepamungkas
11	29/04/2023 13:02	@jek__@shitlicious Duit hasil jd debt collector pinjol bkn ini staffff	zweetheart
12	29/04/2023 12:12	gua ngehargain bgt sih profesionalisme para debt collector yang nagih nya ga pake ribut2.	pajakkalianda
13	29/04/2023 11:49	Sekuriti depan gedung. Ternyata ada orang yg beneran dateng, bapak2 pake jaket hitam gayanya bukan k.	lulalalands
14	29/04/2023 10:12	Terima kasih kepada debt collector yang ngebantuin saya nagih hutang ke temen saya, kalo di tagih sendi	heridians
15	29/04/2023 09:59	Bukak kedai untung ribu2 lepastu semua benda nak hutang itu ini, bukak bisnes dah maju jadi semua	akmaltaufk
16	29/04/2023 09:11	@blank0429 Khas kadrun ngurusin utang pemerintah sampai lupa utangnya sendiri. Mending lo lunasi an	triliuner007
17	29/04/2023 06:24	@yulisr05 @aimluckiest @recehtapisayng Penekanan pada "kalo lo pinter". Kalo lo gak pinter, dlm bbrp t	Don_Nat
18	29/04/2023 05:14	@convomf Tagih ke dia nder, trus jgn mau namamu dipake lagi. Wajib dilunasi, jgn main2 sama hutang nc	catatanumbi
19	29/04/2023 00:13	Makin kesini, debt collector makin enggak kaku, mereka bisa ngasih solusi penyelesaian hutang dengan t	TheaEvanglista

Gambar 2 Dataset Penelitian dari Hasil Teknik *Crawling*



Gambar 3 Jumlah Ulasan Data Labelling

Berdasarkan gambar 3 menunjukkan adanya ketimpangan data antara ulasan negatif dan positif yang berada di dalam *dataset*, maka akan dilakukan *undersampling* yaitu menghilangkan ulasan negatif yang berupa kelas mayoritas, diambil data *sampling* yaitu sebanyak 300 data negatif dan 300 data positif. Sebanyak 600 data tweet tanggapan masyarakat terhadap *debt collector* telah diberi label oleh ahli Bahasa Indonesia. Proses pelabelan dilakukan dengan menggunakan dua label yaitu label positif dan label negatif. Hasil dari pelabelan data yang berdasarkan kolom tweet terhadap kolom keterangan ditunjukkan pada Gambar 4.

1	Date	Tweet	Keterangan (positif / Negatif)
2	30/04/2023 02:22	@dontbry_@r0bbeni baiknya gausah diangkat sma lgsg di blokir aja kak, bbrp hari yg lalu ak	Negatif
3	30/04/2023 00:57	@Indarani Yang itu kayaknya tipe yang lagi ngumpet dari debt collector atau stalker deh δY™-	Negatif
4	29/04/2023 23:22	dia tuh model orang yg kalo temennya mau pinjem duit, dia yg pinjem pake paylater samp	Negatif
5	29/04/2023 22:31	@yookarinza Deg degan, apalagi kalo sambil didatengin debt collector	Negatif
6	29/04/2023 21:56	Username sih investor, tp cara pikirnya kayak debt collector δY™,	Negatif
7	29/04/2023 17:58	mukanya kusem banget akibat dikejar debt collector ya?	Negatif
8	29/04/2023 15:54	@cursedkidd yg nagih mah debt collector bkn sales gmn sih	Negatif
9	29/04/2023 15:18	@TM_One @unifi @helpmeunifi you guys upah Along ke nak kutip due. Debt collector caka	Negatif
10	29/04/2023 14:05	@mas_aro @HooHooTan @neohistoria_id Maksudnya? Orang Maluku merantau ke Jakarta j	Negatif
11	29/04/2023 13:02	@jek_@shitlicious Duit hasil jd debt collector pinjol bkn ini staffff	Negatif
12	29/04/2023 12:12	gua ngehargain bgt sih profesionalisme para debt collector yang nagih nya ga pake ribut2.	Positif
13	29/04/2023 11:49	Sekuriti depan gedung. Ternyata ada orang yg beneran datang, bapak2 pake jaket hitam gay	Negatif
14	29/04/2023 10:12	Terima kasih kepada debt collector yang ngebantuin saya nagih hutang ke temen saya, kalo	Positif
15	29/04/2023 09:59	Bukak kedai untung ribu2 lepastu semua benda nak hutang itu ini, bukak bisnis dah maju ja	Negatif
16	29/04/2023 09:11	@blank0429 Khas kadrun ngurusin utang pemerintah sampai lupa utangnya sendiri. Mendin	Negatif
17	29/04/2023 06:24	@yulistr05 @aimluckiest @recehtapisayng Penekanan pada "kalo lo pinter". Kalo lo gak pinte	Negatif
18	29/04/2023 05:14	@convomf Tagih ke dia nder, trus jgn mau namamu dipake lagi. Wajib dilunasi, jgn main2 sai	Negatif
19	29/04/2023 00:13	Makin kesini, debt collector makin enggak kaku, mereka bisa ngasih solusi penyelesaian hut	Positif

Gambar 4 Hasil Data Setelah Dilakukan *Labelling*

Data Preprocessing

Data yang didapatkan dari hasil proses *data selection* dan telah melalui tahap *labelling* tidak dapat langsung digunakan karena data mengandung banyak *noise* (gangguan). Oleh karena itu, diperlukan tahap *preprocessing* atau *cleansing* untuk membersihkan data dan mengurangi atribut yang kurang berpengaruh terhadap proses klasifikasi.

Berikut adalah urutan tahapan yang diperlukan:

1. *Cleansing*

Pada tahap *cleansing* data dibersihkan dengan menghilangkan karakter yang tidak berpengaruh seperti tanda baca (*punctuation*), simbol-simbol yang biasa digunakan seperti simbol karakter “@#\$\$%^&*()_”.

2. *Case Folding*

Tahap *case folding* dilakukan dengan mengubah semua huruf yang ada di dokumen teks ulasan menjadi huruf kecil (*lowercase*), sehingga huruf kapital dan huruf kecil tidak terdeteksi memiliki perbedaan arti. *Case folding* menjadi salah satu tahap yang penting karena dalam proses pembacaan *corpus*, mesin akan membaca secara *case sensitive*.

3. *Tokenizing*

Proses *tokenizing* ini dilakukan untuk memisahkan setiap kata yang dihubungkan dengan karakter spasi menjadi setiap kata yang dihimpun pada array. Secara sederhana, *tokenizing* merupakan proses pemenggalan kalimat menjadi kata-kata yang terpisah.

4. Normalisasi

Setelah proses tokenisasi maka selanjutnya dilakukan proses normalisasi dimana mengubah kata tidak baku yang ada pada tweet menjadi kata baku.

Pada proses normalisasi akan digunakan dokumen khusus yang berisi daftar kata baku dan tidak baku, dokumen tersebut digunakan untuk mengubah kata tidak baku pada data *tweet* menjadi kata baku.

5. *Filtering*

Tahap *filtering* atau *stopword removal*, yakni menyaring kata-kata yang relevan guna proses klasifikasi selanjutnya. Contoh *stopword* dalam bahasa Indonesia yaitu kata-kata konjungsi seperti “dan”, “namun”, “atau”, dll. Pada proses *filtering*, kata-kata yang ada dalam dokumen teks ulasan dibandingkan dengan kata-kata yang terdapat pada *stopword*, jika dalam dokumen terdapat kata yang dimuat pada *stopword* maka kata tersebut akan dihilangkan (dihapus).

6. *Stemming*

Proses *stemming* dilakukan dengan mengubah kata yang berimbuhan menjadi kata dasar. *Package* yang digunakan untuk proses *stemming* adalah *package* *katadasaR*.

Data Transformation

Pada tahap ini dilakukan pembobotan menggunakan TF-IDF (*Term Frequency Inverse Document Frequency*) dan menggunakan BoW (*Bag of Words*). Perhitungan pembobotan kata dilakukan dengan menentukan nilai *Term Frequency* (TF) terlebih dahulu. Hasil dari pembobotan kata menggunakan TF-IDF (*Term Frequency Inverse Document Frequency*) dan BoW (*Bag of Words*) dapat dilihat pada gambar 5 dan 6.

docs	ambil	beli	collector	debt	debtcollector	diberentiin	gara	gitu	miung	sembunyi
50	0.000000	0	0.1343011	0.1197136	0	0	0	0.000000	0	0
51	0.000000	0	0.1343011	0.1197136	0	0	0	0.000000	0	0
52	0.000000	0	0.1343011	0.1197136	0	0	0	4.169925	0	0
53	0.000000	0	0.1343011	0.1197136	0	0	0	0.000000	0	0
54	0.000000	0	0.2686022	0.2394271	0	0	0	0.000000	0	0
55	0.000000	0	0.1343011	0.1197136	0	0	0	0.000000	0	0
56	0.000000	0	0.1343011	0.1197136	0	0	0	0.000000	0	0
57	0.000000	0	0.1343011	0.1197136	0	0	0	0.000000	0	0
58	6.754888	0	0.1343011	0.1197136	0	0	0	0.000000	0	0
59	0.000000	0	0.1343011	0.1197136	0	0	0	0.000000	0	0

Gambar 5 Nilai Bobot Menggunakan TF-IDF

docs	collector	debt	miung	sembunyi	ambil	beli	debtcollector	diberentiin	gara	gitu
50	1	1	0	0	0	0	0	0	0	0
51	1	1	0	0	0	0	0	0	0	0
52	1	1	0	0	0	0	0	0	0	1
53	1	1	0	0	0	0	0	0	0	0
54	1	1	0	0	0	0	0	0	0	0
55	1	1	0	0	0	0	0	0	0	0
56	1	1	0	0	0	0	0	0	0	0
57	1	1	0	0	0	0	0	0	0	0
58	1	1	0	0	1	0	0	0	0	0
59	1	1	0	0	0	0	0	0	0	0
60	1	1	0	0	0	0	0	0	0	0

Gambar 6 Nilai Bobot Menggunakan BoW

Proses pembobotan dilakukan dengan membandingkan bobot setiap kata pada *data testing* dengan kata pada *data training*. Jika bobot probabilitas positif lebih besar maka hasilnya tweet tersebut masuk ke dalam kelas positif. Jika bobot probabilitas negative lebih besar maka hasilnya tweet tersebut masuk ke dalam kelas negatif.

Data Mining

Pada tahap *data mining* merupakan tahap klasifikasi dari teks ulasan menggunakan *Naïve Bayes Classifier*. Sebelum dilakukan klasifikasi, perlu adanya pembagian data latih dan data uji. Data latih merupakan data yang digunakan dalam melatih atau membentuk model. Sedangkan data uji adalah data yang digunakan untuk menguji apakah *classifier* yang telah dilatih berhasil melakukan klasifikasi dengan benar. Pada penelitian ini dilakukan 4 kali pengujian yaitu pertama menggunakan *presentase* 90% data latih dan 10% data uji, pengujian yang kedua menggunakan *presentase* 80% data latih dan 20% data uji, serta pengujian ketiga menggunakan *presentase* 70% data latih dan 30% data uji, serta pengujian keempat menggunakan *presentase* 60% data latih dan 40% data uji. Berikut merupakan jumlah data latih dan data uji pada masing-masing pengujian.

Presentase Latih : Uji (%)	Data Latih	Data Uji
90:10	540	60
80:20	480	120
70:30	420	180
60:40	360	240

Tabel 1 Pembagian Data Latih dan Data Uji

Analisis sentimen masyarakat terhadap *debt collector* pada twitter menggunakan *Naïve Bayes Classifier* pada RStudio didapatkan hasil klasifikasi yang disajikan dalam bentuk tabel *confusion matrix* yang berisi informasi mengenai jumlah data klasifikasi yang sebenarnya (*actual*) dengan jumlah data klasifikasi hasil prediksi.

Evaluation

Pada tahap evaluasi ini, dilakukan untuk mengukur kinerja klasifikasi dilakukan menggunakan *confusion matrix* dengan nilai *accuracy*, *precision*, dan *recall* yang kemudian hasil klasifikasi setiap pengujian akan dibandingkan hasilnya dengan menggunakan TF-IDF (*Term Frequency Inverse Document Frequency*) dan menggunakan BoW (Bag of Words) serta hasil perhitungan kinerja klasifikasi dari RStudio untuk data dengan setiap pengujian. Untuk mendapatkan hasil klasifikasi menggunakan perhitungan sebagai berikut.

$$Accuracy = \frac{TP+TN}{(TP+TN+FP+FN)}$$

$$Precision = \frac{TN}{(TN+FP)}$$

$$Recall = \frac{TN}{(TN+FN)}$$

Pembahasan

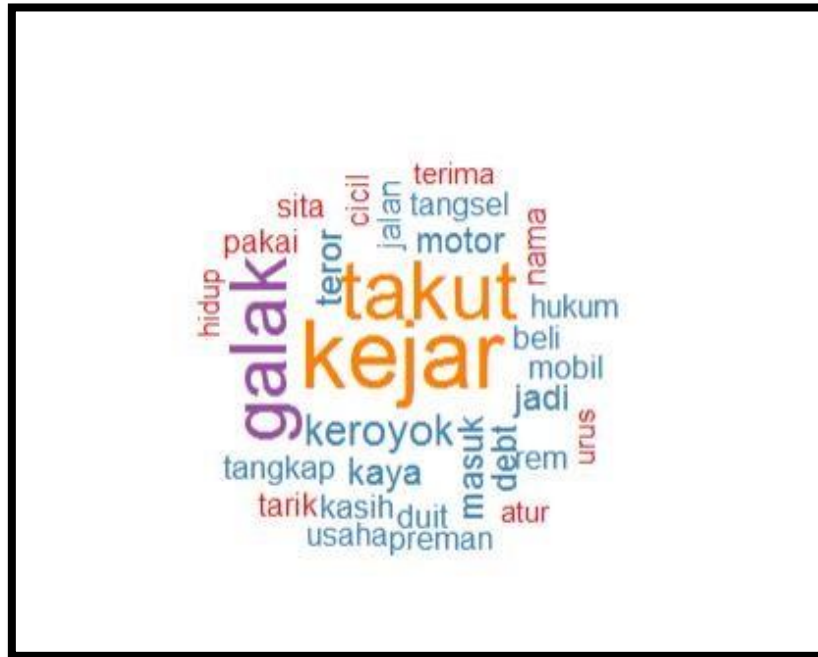
Hasil analisis sentimen menggunakan *Naïve Bayes Classifier* dalam mengklasifikasi data tweet masyarakat terhadap *debt collector* pada twitter menjadi dua kelas yaitu kelas positif dan kelas negatif dengan menggunakan beberapa perbandingan data latih dan data uji, yaitu 90:10 (90% data latih dan 10% data uji), 80:20 (80% data latih dan 20% data uji), 70:30 (70% data latih dan 30% data uji), 60:40 (60% data latih dan 40% data uji). Hasil akurasi menggunakan *Naïve Bayes Classifier* untuk penggunaan data 90:10 dengan menggunakan TF-IDF (*Term Frequency Inverse Document Frequency*) mencapai nilai akurasi tertinggi yakni sebesar 78.3%. Sedangkan untuk pengujian dengan menggunakan BoW (Bag of Words) mencapai nilai akurasi tertinggi diperoleh pada penggunaan data 90:10 dengan nilai yaitu sebesar 76.6%. Untuk nilai *precision* tertinggi diperoleh dari pengujian dengan menggunakan data 90:10 dengan menggunakan BoW yaitu sebesar 80%, sedangkan untuk nilai *recall* tertinggi diperoleh dari pengujian

penggunaan data 90:10 dengan menggunakan BoW yaitu sebesar 80%. Penggunaan data dengan perbandingan jumlah data latih dan data uji yang berbeda mempengaruhi tingkat akurasi pengklasifikasian data ulasan menggunakan *Naïve Bayes Classifier*. Selain itu, penggunaan metode pembobotan BoW juga berhasil meningkatkan nilai akurasi pada sebagian besar rasio pengujian. Hasil perhitungan terhadap 4 skenario menggunakan TF-IDF dan menggunakan BoW dapat dilihat pada tabel 2.

Tabel 2 Hasil Perhitungan Pemodelan *Naïve Bayes Classifier* Menggunakan TFIDF dan Menggunakan BoW

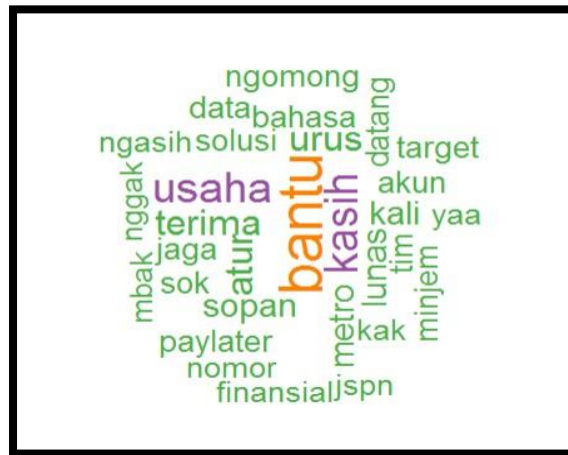
Skenario	Dataset	Accuracy	Precision	Recall
Skenario 1 (90:10)	TF-IDF	78.3%	77.7%	75%
	BoW	76.6%	80%	80%
Skenario 2 (80:20)	TF-IDF	73.3%	70.1%	72.7%
	BoW	76.6%	77.4%	77.4%
Skenario 3 (70:30)	TF-IDF	71.6%	68.2%	70.7%
	BoW	73.8%	75.5%	76.2%
Skenario 4 (60:40)	TF-IDF	73.3%	72.5%	73.7%
	BoW	73.7%	79.3%	64.1%

Hasil klasifikasi analisis sentiment tweet masyarakat terhadap *debt collector* dapat divisualisasikan dengan menggunakan *word cloud* dan *diagram fishbone* untuk mengetahui gambaran atau informasi umum mengenai data tweet terhadap *debt collector* pada twitter. Berikut merupakan pembahasan visualisasi kata dari masing-masing kelas sentimen.



Gambar 7 *Word Cloud* Ulasan Negatif

Berdasarkan gambar 7 diperoleh informasi bahwa masyarakat mengalami rasa takut dengan perlakuan *debt collector* di masyarakat dan banyak mengeluhkan bahwa *debt collector* sebagian besar menggunakan terror dalam melakukan tugasnya.

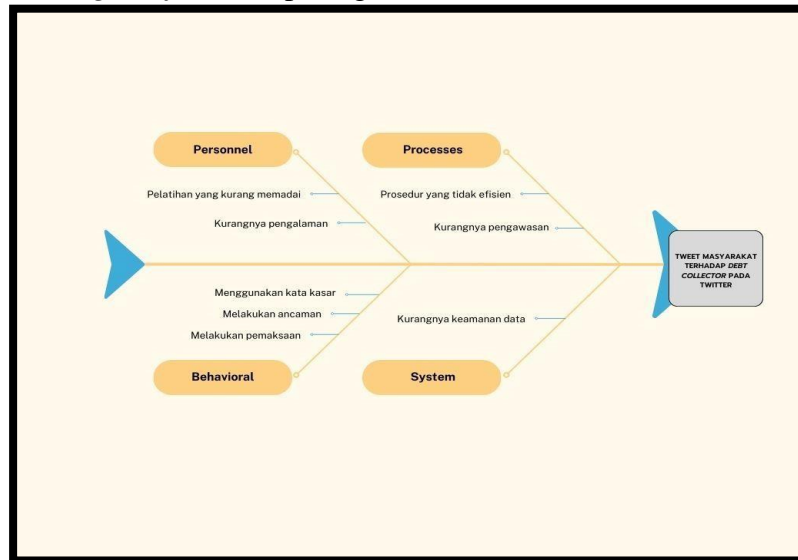


Gambar 8 *word cloud* Ulasan Positif

Berdasarkan gambar 8 Dapat diperoleh informasi bahwa masyarakat terbantu dengan hadirnya *debt collector* dalam penyelesaian hutang dan beberapa diantaranya juga memberikan solusi agar persoalan terhadap hutang dapat diselesaikan.

Semakin besar ukuran kata dalam *word cloud* maka semakin tinggi pula frekuensi kata tersebut, artinya kata tersebut sering digunakan sebagai tweet terhadap *debt collector* di twitter. Berdasarkan hasil ulasan negatif yang didapatkan dari data tweet, maka dapat diperoleh informasi mengenai masalah-masalah yang

terjadi terkait dengan tweet masyarakat terhadap *debt collector* pada twitter berdasarkan *diagram fishbone* pada gambar 4.36 berikut ini.



Gambar 9 *Diagram Fishbone* Tweet Masyarakat Terhadap *Debt collector* Pada Twitter

Pada gambar 9 terdapat berbagai informasi mengidentifikasi faktor-faktor yang menjadi penyebab tweet negatif terhadap tweet masyarakat terhadap *debt collector* pada twitter. Faktor-faktor ini meliputi masalah dalam aspek *personnel*, *processes*, *behavioral*, dan *system* yang terkait dengan tweet masyarakat terhadap *debt collector* pada twitter. Permasalahan yang digunakan dalam gambar 9 diperoleh melalui visualisasi yang telah dilakukan sebelumnya untuk memberikan gambaran yang jelas tentang permasalahan yang ada. Rencana solusi untuk mengatasi masalah teridentifikasi dalam tweet masyarakat terhadap *debt collector* pada twitter dapat ditemukan dalam tabel 3.

Tabel 3 Rencana Pemecahan Masalah Tanggapan Masyarakat Terhadap *Debt collector* pada Twitter

No	Faktor	Permasalahan	Pemecahan Masalah
1	Personnel (Personil)	Pelatihan yang kurang memadai	Tingkatkan program pelatihan untuk mengedukasi kolektor hutang tentang praktik terbaik dalam penagihan hutang, hukum terkait, dan keterampilan komunikasi yang efektif.
		Kurangnya pengalaman	Berikan dukungan dan mentorship kepada kolektor yang kurang berpengalaman, dan pertimbangkan pengembangan jalur karir yang jelas.

2	Behavioral (Perilaku)	Menggunakan kata kasar	Selenggarakan pelatihan tambahan untuk semua anggota tim tentang etika komunikasi yang sesuai dalam penagihan hutang. Berikan panduan tentang bahasa yang harus digunakan dan bahasa yang harus dihindari.
		Melakukan ancaman	Berikan pelatihan kepada anggota tim tentang cara berkomunikasi dengan peminjam secara positif dan menghindari ancaman. Tetapkan kebijakan yang melarang penggunaan ancaman dalam komunikasi penagihan. Jelaskan konsekuensi bagi pelanggaran kebijakan ini.
		Melakukan pemaksaan	Berikan pelatihan kepada kolektor tentang hukum-hukum yang mengatur penagihan hutang dan etika dalam praktik penagihan. Tetapkan kebijakan melarang pemaksaan dalam proses penagihan hutang. Jelaskan konsekuensi yang akan diterapkan jika ada pelanggaran kebijakan ini.
3	Processes (Proses)	Prosedur yang tidak efisien	Evaluasi dan perbarui prosedur penagihan hutang untuk memastikan efisiensi. Terapkan alat dan sistem yang dapat membantu tugas-tugas rutin.
No	Faktor	Permasalahan	Pemecahan Masalah
4	System (Sistem)	Kurangnya pengawasan	lebih sering memeriksa pekerjaan mereka melalui audit reguler, memberikan saran yang membantu, dan memastikan manajemen selalu mendukung mereka.
		Kurangnya keamanan data	Tetapkan kebijakan yang jelas tentang bagaimana data peminjam harus dikelola dan disimpan dengan aman. Berikan pelatihan kepada tim <i>debt collector</i> tentang praktik keamanan data yang baik.

KESIMPULAN

Berdasarkan penelitian yang telah dilakukan dapat disimpulkan bahwa Hasil *accuracy* tertinggi adalah klasifikasi menggunakan *Naïve Bayes Classifier* dengan menggunakan metode TF-IDF yaitu sebesar 78.3% dengan presentasi penggunaan data 90:10 (90% data latih dan 10% data uji). Untuk nilai *precision* tertinggi diperoleh dari pengujian dengan data 90:10 dengan menggunakan metode BoW yaitu sebesar 80%, sedangkan nilai *recall* tertinggi diperoleh dari pengujian dengan penggunaan data 90:10 menggunakan metode BoW yaitu sebesar 80%. Informasi yang didapatkan dari proses pengklasifikasian adalah penggunaan data dengan perbandingan jumlah data latih dan data uji yang berbeda mempengaruhi tingkat akurasi pengklasifikasian data tweet masyarakat terhadap *debt collector* pada twitter menggunakan *Naïve Bayes Classifier*. Selain itu, penggunaan metode BoW (*Bag of Words*) juga berhasil meningkatkan nilai akurasi pada sebagian besar skenario pengujian.

Gambaran atau informasi umum mengenai data tweet masyarakat terhadap *debt collector* pada twitter diperoleh dari visualisasi dengan menggunakan *word cloud* dan *diagram fishbone*. Dalam proses ini, terdapat beberapa faktor yang mempengaruhi gambaran atau informasi umum yang diungkapkan oleh masyarakat terhadap *debt collector*. Faktor-faktor ini meliputi 4 bagian yaitu *personnel*, *behavioral*, *processes*, dan *system* terkait dengan interaksi masyarakat dengan *debt collector*. Pendapat dan pengalaman masyarakat terhadap *debt collector* tercermin dari dalam data tweet yang telah dikumpulkan dan faktor-faktor seperti pelatihan yang kurang memadai, kurangnya pengalaman, penggunaan kata kasar, ancaman, pemaksaan, prosedur yang tidak efisien, kurangnya pengawasan, sistem yang ketinggalan zaman, dan kurangnya keamanan data adalah kesimpulan yang didapatkan dan diharapkan pihak-pihak yang terkait bisa melakukan perbaikan untuk meningkatkan *performa* dan kepercayaan masyarakat terhadap *debt collector*.

REFERENCES

- Aprilia, R. A., & Anisa, H. (2019). Analisis Sentimen Masyarakat pada Media Sosial Twitter Terhadap Partai Politik Peserta Pemilu 2019 Menggunakan *Naïve Bayes Classifier*. *e-Proceeding of Engineering*, 6(2), 8617.
- Arjuni, S. (2016). Analisis Sentimen Masyarakat Terhadap Produk Televisi Berlangganan Di Media Sosial Twitter. Tugas Akhir. *Universitas Widyatama Bandung*.
- Arsi, P. (2021). Analisis Sentimen Wacana Pemindahan Ibu Kota Indonesia Menggunakan Algoritma Support Vector Machine (SVM). *Jurnal Teknologi Informasi Dan Ilmu Komputer*, 8(1), 147.
- Dedy, S, Ema, U, & Ainul, Y. (2022). Perbandingan Kinerja Model TF-IDF dan BoW untuk Klasifikasi Opini Publik Tentang Kebijakan BLT Minyak Goreng. *Jurnal Teknik Industri*, 12(3).
- Dianati, D. P. (2022). Analisis Sentimen Kinerja Dewan Perwakilan Rakyat (DPR) Pada Twitter Menggunakan Metode Naive Bayes Classifier. *Jurnal Informatika dan Teknik Elektro Terapan (JITET)*, 10(1).

- Durroh, B. D. (2023). Analysis of Quality Control of Tea Products Using the Fishbone Diagram Approach at Pt Candi Loka, Indonesia. *Asian Journal of Research in Crop Science*, 8(January), 16-24.
<https://doi.org/10.9734/ajrcs/2023/v8i1154>
- Fahmi, A. R. (2020). Analisis Sentiment Masyarakat Selama Bulan Ramadhan Dalam Menghadapi Pandemi Covide-19. *Jurnal Informatika Dan Sistem Informasi (JIFoSI)*, 1(1), 608-617.
<https://doi.org/10.9734/ajrcs/2023/v8i1154>
- Fajrin, A. A. (2018). Penerapan Data Mining Untuk Analisis Pola Pembelian Konsumen Dengan Algoritma Fp-Growth Pada Data Transaksi Penjualan Spare Part Motor. *Kumpulan Jurnal Ilmu Komputer (KLIK)*, 5(1), 27-36.
- Gata, W. (2017). Akurasi Text Mining Menggunakan Algoritma K- Nearest Neighbour pada Data Content Berita SMS. *Jurnal Format*, 6(1), 1-13.
- Kaplan, A. (2015). Users of the world, unite! The challenges and opportunities of social media. *ELSEVIER*, 53(1), 59-68.
- Kolekar, N. V., Rao, G., Dey, S., Mane, M., Jadhav, V., & Patil, S. (2016). Sentiment Analysis and Classification Using Lexicon Based Approach and Addressing Polarity Shift Problem. *Journal Of Theoretical and Applied Information Technology*, 90(1), 118-125.
- Lingga, A. A. (2019). Analisis Sentimen Masyarakat terhadap Hasil Quick Count Pemilihan Presiden Indonesia 2019 pada Media Sosial Twitter Menggunakan Metode Naive Bayes Classifier. *Indonesian Journal of Applied Statistics*, 2(1).
- Liu, B. (2011). *Sentimen analysis and opinion mining*. Morgan & Claypool Chicago: Springer.
- Maarif, M. R. (2018). Content Analysis on Twitter Users Interaction within First 100 Days of Jakarta's New Government by Using Text Mining. *Journal Pekommas*, 3(2), 137.
<https://doi.org/10.30818/jpkm.2018.2030203>
- Manik, G. E. (2021). Analisis Sentimen Pada Review Pengguna E-Commerce Bidang Pangan Menggunakan Metode Support Vector Machine (Studi Kasus: Review Sayurbox dan Tanihub pada Google Play). *Senamika*, 2(2), 64-74.
- Nugroho, R. A. (2021). Implementasi Naïve Bayes Classifier untuk Klasifikasi Emosi Tweet Berbahasa Indonesia pada Spark. *JPTIHK*, 5(1), 301-310.
- Paulina, W. B. (2020). Analisis Sentimen Berbasis Aspek Ulasan Pelanggan Terhadap Kertanegara Premium Guest House Menggunakan Support Vector Machine. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, 4(4), 1141-1149.
- Praptiwi, D. Y. (2018). Analisis Sentimen Online Review Pengguna E-Commerce Menggunakan Metode Support Vector Machine Dan Maximum Entropy (Studi Kasus: Review Bukalapak Pada Google Play). Tugas Akhir. *Universitas Islam Indonesia*.

- Priza, P. K. (2022). Analisis Sentimen Opini Publik Terhadap Program Vaksinasi Covid-19 Di Indonesia Pada Twitter Menggunakan Metode Naive Bayes Classifier. *Informatics Journal*, 7(3).
- Rodriguez, N. (2021). *SOUL: Scala Oversampling and Undersampling Library for Imbalancing Classification*. *ELLSEVIER*.
<https://doi.org/10.1016/j.softx.2021.100767>
- Syarifuddin, M. (2020). Analisis Sentimen Opini Publik Mengenai Covid-19 Pada Twitter Menggunakan Metode Naive Bayes dan KNN. *Inti Nusa Mandiri*, 15(1).
- Yanuar, N. F. (2021). Analisis Sentimen Opini Publik Terhadap Undang-Undang Cipta Kerja pada Twitter Menggunakan Metode Naive Bayes Classifier. *Prosiding Seminar Nasional Sains Teknologi dan Inovasi Indonesia*, 3, 201212.
- Yonathan, S. M. (2018). Analisis Sentimen Terhadap Pemerintahan Joko Widodo Pada Media Sosial Twitter Menggunakan Algoritma Naive Bayes Classifier. *Prosiding SINTAK*