

Analisis Sentimen Pinjaman Online Di Media Sosial Twitter Menggunakan Metode Naive Bayes

Muhammad Imam Ghozali*, Wibowo Harry Sugiharto, Ary Fajar Iskandar

Program Studi Teknik Informatika, Fakultas Teknik, Universitas Muria Kudus, Kudus, Indonesia

Email: ^{1,*}imam.ghozali@umk.ac.id, ²wibowo.harrys@umk.ac.id, ³aryfajar011@gmail.com

Email Penulis Korespondensi: imam.ghozali@umk.ac.id

Abstrak—Analisis sentimen digunakan untuk mengukur kecenderungan opini masyarakat terhadap suatu kejadian yang sedang berlangsung atau telah terjadi. Salah satu kasus yang dianalisis dalam penelitian ini adalah Pinjaman Online atau yang umum dikenal sebagai Pinjol. Data penelitian mengenai Pinjaman Online diperoleh dari media sosial Twitter dengan menggunakan kata kunci Pinjaman Online. Metode analisis yang digunakan adalah Naive Bayes. Sebelum proses analisis sentimen dilakukan, data teks diambil melalui crawling dari API Twitter dengan menggunakan aplikasi Rapidminer. Data tersebut kemudian melalui tahap text pre-processing. Selanjutnya, data diberikan pembobotan menggunakan metode TF-IDF. Hasil penelitian ini menunjukkan kecenderungan konflik sentimen positif dan negatif pada tiap tweet yang dibahas oleh pengguna Twitter mengenai Pinjaman Online. Kesimpulan yang dapat diambil dari analisis sentimen menggunakan klasifikasi algoritma Naive Bayes dengan menggunakan data dari Twitter tentang Pinjaman Online adalah sebagai berikut: Dari 2931 data yang digunakan, setelah melalui proses text preprocessing, terdapat 2912 data yang tersedia. Dari jumlah tersebut, sentimen negatif tercatat sebanyak 68,61% dengan 1998 data, sementara sentimen positif tercatat sebanyak 31,39% dengan 914 data. Analisis sentimen dari pengguna Twitter mengenai Pinjaman Online memiliki tingkat akurasi sebesar 80%.

Kata Kunci: Pinjaman Online; Naive Bayes; Analisis Sentimen; Twitter

Abstract—Sentiment analysis is conducted to measure public opinion tendencies towards ongoing or past events. One of the cases analyzed in this study is Online Loans, commonly known as "Pinjol" in Indonesian. The research data regarding Online Loans was collected from the Twitter social media platform using the keyword "Pinjaman Online." The analysis method employed in this study is Naive Bayes. Prior to the sentiment analysis process, text data was obtained through crawling from the Twitter API using the Rapidminer application. The data was then subjected to text pre-processing. Subsequently, the data underwent TF-IDF weighting. The results of this research demonstrate the tendencies of positive and negative sentiment conflicts within each tweet discussed by Twitter users regarding Online Loans. The conclusion drawn from the sentiment analysis using the Naive Bayes classification algorithm with data obtained from Twitter concerning Online Loans is as follows: Out of the 2931 data used, after undergoing text pre-processing, a total of 2912 data were available. Among them, negative sentiment accounted for 68.61% with 1998 data, while positive sentiment accounted for 31.39% with 914 data. The sentiment analysis of Twitter users regarding Online Loans achieved an accuracy rate of 80%.

Keywords: Online Loans; Naive Bayes; Sentiment Analysis; Twitter

1. PENDAHULUAN

Munculnya berbagai macam perkembangan teknologi, terutama di bidang keuangan. Salah satu yang cukup populer akhir-akhir ini yaitu pinjaman online. Dimana pinjaman online merupakan perkembangan teknologi dari konsep peminjaman uang yang kini bisa melalui online atau menggunakan teknologi untuk memudahkan transaksi.

Pinjaman Online juga dapat membantu pembiayaan usaha kecil menengah untuk mengembangkan usahanya, dengan menjadi tempat transaksi peminjaman yang mudah dan aman disatu sisi hal ini menimbulkan masalah terhadap masyarakat Indonesia yang belum siap menerima perubahan dalam aktivitas ekonomi. FinTech sekarang mampu menggantikan peran lembaga keuangan formal seperti bank, bahkan sebagai sistem pembayaran baru[1]. Masalah lain yang muncul ialah banyak pemberi pinjaman online yang belum memiliki regulasi yang memadai, sehingga potensi risiko bagi para peminjam menjadi lebih tinggi. Hal ini diperparah dengan maraknya kasus penipuan dan tindakan yang tidak etis oleh beberapa pemberi pinjaman online. Oleh karena itu timbullah berbagai macam persepsi masyarakat terhadap penggunaan pinjaman online.

Salah satu tempat yang digunakan untuk menyampaikan tanggapan dan keluhan adalah dengan menggunakan media sosial, salah satunya yaitu twitter[2]. Oleh karena itu, analisis sentimen terhadap pinjaman online di Twitter menjadi penting untuk mendapatkan pemahaman tentang pandangan dan opini masyarakat mengenai layanan ini. Sentiment analysis atau opinion mining mengacu pada bidang yang luas dari pengolahan bahasa alami, komputasi linguistik dan text mining yang bertujuan menganalisa pendapat, sentimen, evaluasi, sikap, penilaian dan emosi seseorang apakah pembicara atau penulis berkenan dengan suatu topik, produk, layanan, organisasi, individu, ataupun kegiatan tertentu[3].

Metode Naive Bayes merupakan salah satu metode analisis sentimen yang cukup populer dan sering digunakan dalam penelitian-penelitian terkait. Metode ini memiliki tingkat akurasi yang cukup tinggi dan dapat menghasilkan hasil analisis dengan cepat dan efektif. Beberapa kasus sebelumnya adalah penerapan pada kebijakan Pembatasan Kegiatan Masyarakat (PPKM) saat pandemic COVID19. Penelitian ini bertujuan untuk mengevaluasi sentimen masyarakat terhadap PPKM dengan mengklasifikasikan tweet menggunakan teknik text mining. Algoritma K-Nearest Neighbor (KNN) dan Naive Bayes Classifier (NBC) digunakan dengan data dari 3.516 tweet yang dikumpulkan selama PPKM tahun lalu. Hasilnya, algoritma NBC lebih baik dibandingkan KNN dengan akurasi sebesar 79,67% dibandingkan

78,86%. Sentimen masyarakat terhadap PPKM didapatkan sebagai berikut: 36,83% sentimen positif (1.295 tweet), 54,15% sentimen netral (1.902 tweet), dan 9,02% sentimen negatif (317 tweet)[4]. Metode serupa juga digunakan untuk menganalisis kebijakan-kebijakan pemerintah lain seperti kenaikan tarif dasar listrik [5], isu penundaan pemilu [6], kebijakan penanganan covid [7], penyelenggaraan BPJS [8], dll. Hal ini dapat menjadi acuan dalam pengambilan keputusan dari apa luaran analisis.

Berbagai pro dan kontra dalam pinjaman online tak terhindarkan. Mulai dari pro kontra pinjaman online dalam pandangan agama maupun pro kontra pinjaman online yang legal dan ilegal. Hal tersebut membuat pandangan terhadap pinjaman online sangat majemuk atau bisa dikatakan bisa dilihat dari berbagai sisi. Oleh karena itu, penulis berusaha mengangkat isu pinjaman online ini dengan menggunakan analisis sentimen di media sosial twitter. Diharapkan hasil penelitian ini dapat memberikan wawasan yang bermanfaat bagi masyarakat dan pihak yang berkepentingan dalam industri pinjaman online untuk meningkatkan regulasi dan memberikan pelayanan yang lebih baik bagi para peminjam.

2. METODOLOGI PENELITIAN

Dalam penelitian ini, kami mengumpulkan data tweet dari Twitter menggunakan metode scraping. Scraping adalah proses mengambil data dari situs web. Data yang diperoleh kemudian akan diolah dan dianalisis dengan fokus pada sentimen yang terkandung dalam data tersebut. Kami menggunakan metode analisis sentimen dengan menggunakan algoritma klasifikasi Naive Bayes. Tahapan penelitian yang dilakukan terdiri dari:

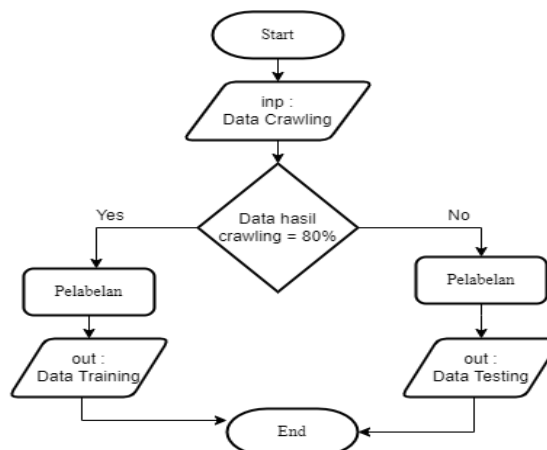
- Pengumpulan Data: Data tweet diambil dari Twitter menggunakan teknik scraping. Proses scraping dilakukan untuk mendapatkan sejumlah tweet yang relevan dengan topik penelitian.
- Preprocessing Data: Data tweet yang terkumpul kemudian melalui tahap preprocessing, termasuk tahap pembersihan data seperti penghilangan karakter khusus, penghapusan tautan, dan penghapusan tanda baca yang tidak diperlukan. Selain itu, dilakukan juga proses tokenisasi untuk memecah kalimat menjadi kata-kata.
- Analisis Sentimen: Setelah data tweet dipreprocessing, langkah selanjutnya adalah menerapkan algoritma klasifikasi Naive Bayes untuk melakukan analisis sentimen. Algoritma ini digunakan untuk mengklasifikasikan tweet menjadi sentimen positif, negatif, atau netral berdasarkan kata-kata yang terkandung di dalamnya.
- Evaluasi dan Interpretasi Hasil: Hasil dari analisis sentimen dievaluasi dan diinterpretasikan untuk mendapatkan pemahaman tentang sentimen masyarakat terkait topik penelitian.

2.1 Penambangan Data

Tahap berikutnya adalah mengambil atau mengumpulkan data yang akan digunakan dalam analisis sentimen ini. Proses pengambilan data menggunakan software RapidMiner. Data yang diperoleh untuk melakukan penelitian ini adalah dengan cara scraping yaitu data yang diambil langsung dari tweet menggunakan Twitter API[9].

2.2 Data Training dan Data Testing

Pembagian data training dan data testing dilakukan dengan perbandingan 80:20 dari total 500 data yang telah dilabeli secara manual. Data training digunakan sebagai data latih dalam proses analisis sentimen. Sementara itu, data testing digunakan untuk menguji seberapa akurat proses analisis sentimen ini, seperti yang terlihat dalam Gambar 1. Pembagian ini memungkinkan evaluasi kinerja model dalam mengklasifikasikan sentimen dengan menggunakan data yang tidak digunakan dalam proses pelatihan.



Gambar 1. Diagram Alir Pembagian Data Training dan testing

2.3 Text Pre-Processing

Langkah selanjutnya setelah tahap pemberian label adalah tahap *Text Preprocessing*. *Text Preprocessing* dilakukan untuk merubah beberapa kata asing ke Bahasa Indonesia dan membuang beberapa karakter yang tidak diperlukan[10]. Tahap *text preprocessing* yang dilakukan dalam tahap ini adalah untuk mengolah hasil *crawling* data yang berbentuk kalimat

tidak terstruktur menjadi data yang tersusun dari kata dasar. *Text Pre-processing* merupakan proses ekstraksi pola (informasi dan pengetahuan yang berguna) dari sejumlah data tak terstruktur yang nantinya akan diperoleh pola-pola data, tren dan ekstraksi pengetahuan yang potensial dari data teks[11]. Data yang diambil dari twitter adalah data mentah yang mengandung bagian-bagian yang tidak diperlukan pada saat pengolahan data. Sehingga perlu upaya untuk menyiapkan data agar data layak untuk digunakan pada pengolah selanjutnya[12].



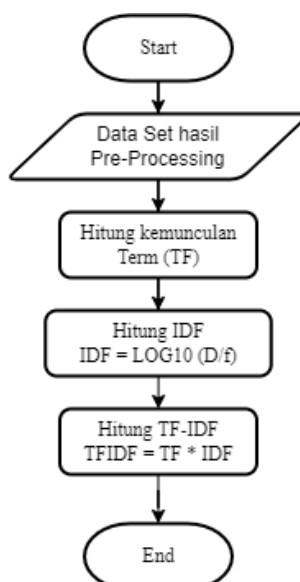
Gambar 2. Diagram Alir Proses Text Pre-processing

Berikut penjelasan gambar 2 tahapan Text Pre- Processing, antara lain:

- Stop Words Removal merupakan penghapusan kata sambung, karena kata-kata tersebut dianggap tidak memiliki arti dan membuat teks menjadi sulit untuk di analisa[13].
- Casefolding adalah sebuah proses menjadikan seluruh huruf menjadi huruf kecil (lowercase)[14].
- Tokenizing untuk mengubah semua karakter huruf menjadi huruf[15].
- Cleansing adalah tahapan dimana karakter dan tanda baca yang tidak diperlukan dihilangkan dari teks. Berfungsi untuk mengurangi noise pada dataset[16].
- Stemming adalah perubahan kata berimbuhan menjadi kata dasar. [17].

2.4 Pembobotan TF-IDF

Diagram alir proses pembobotan TF-IDF dapat dilihat pada Gambar 3 berikut ini.



Gambar 3. Diagram Alir Pembobotan TF-IDF

2.5 Klasifikasi Naive Bayes

Dalam pengklasifikasian data selanjutnya akan diklasifikasikan menggunakan metode *Naive Bayes* melalui alur sebagai berikut. Salah satu metoda dalam klasifikasi data adalah *Naive Bayes* yaitu salah satu metode machine learning yang memanfaatkan perhitungan probabilitas dan statistik yang dikemukakan oleh ilmuwan Inggris Thomas Bayes, cara kerja *Naive Bayes* yaitu memprediksi probabilitas di masa depan berdasarkan pengalaman di masa sebelumnya[18].

Klasifikasi *Naive Bayes* juga memperlihatkan tingginya akurasi dan cepat ketika digunakan untuk dataset dengan jumlah besar[19]. Performa *naive bayes* yang kompetitif dalam proses klasifikasi walaupun menggunakan asumsi keidependenan atribut (tidak ada kaitan antar atribut). Asumsi keidependenan atribut ini pada data sebenarnya jarang terjadi, namun walaupun asumsi keidependenan atribut tersebut dilanggar performa pengklasifikasian *naive bayes* cukup tinggi, hal ini dibuktikan pada berbagai penelitian empiris[20].

$$\text{Prior positif} = (\text{jumlah documen positif})/(\text{Jumlah seluruh documen}) \quad (1)$$

$$\text{Prior negatif} = (\text{jumlah dokumen negatif})/(\text{Jumlah seluruh document}) \quad (2)$$

Kemudian menghitung nilai conditional probability positif dan negatif. Berdasarkan pada perhitungan conditional pada persamaan 1 dan 2.

$$\text{Conditional positif} = (\text{bobot tfidf[kata]+1})/(\text{Jumlah tfidf+jumlah sumidf}) \quad (3)$$

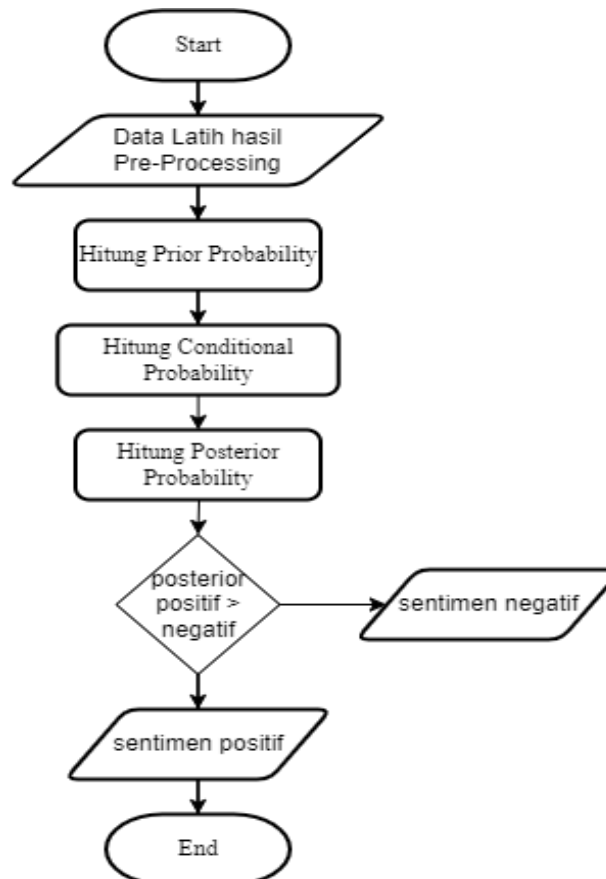
$$\text{Conditional negatif} = (\text{bobot tfidf[kata]+1})/(\text{Jumlah tfidf+jumlah sumidf}) \quad (4)$$

Perhitungan selanjutnya yaitu, posterior probability positif dan negatif

$$\text{Posterior positif} = \text{Prior positif} * \text{Conditional positif} \quad (5)$$

$$\text{Posterior negatif} = \text{Prior negatif} * \text{Conditional negatif} \quad (6)$$

Jika nilai posterior positif lebih besar daripada nilai posterior negatif maka kalimat atau dokumen tersebut akan diklasifikasikan menjadi sentimen positif. Begitu pun sebaliknya, Jika nilai posterior negatif lebih besar daripada nilai posterior positif maka kalimat atau dokumen tersebut akan diklasifikasikan menjadi sentimen negatif dapat dilihat pada gambar 4.



Gambar 4. Diagram Alir Proses Klasifikasi

3. HASIL DAN PEMBAHASAN

Pada bab keempat, akan dibahas hasil dan pembahasan dari penelitian yang telah dilakukan. Hasil dari analisis yang telah dilakukan akan dibahas secara detail, berdasarkan tahap-tahap yang telah dilakukan sesuai dengan metode yang digunakan. Implementasi dari penelitian ini juga akan dibahas, yang didasarkan pada pengolahan data *tweet* dari Twitter yang telah difilter menggunakan *keyword* pinjaman *online*.

3.1 Penambangan Data

Tahap penambangan data sendiri, dalam penelitian ini dilakukan dengan cara *crawling*. Dalam tahap *crawling* menggunakan *tools* RapidMiner, yang kemudian data hasil *crawling* akan disimpan dalam bentuk ekstensi excel (.xls).

Tabel 1. Data Hasil *Crawling*

No	Komentar
1	nah iyak, paman aja bisa ngasih pinjaman online sama cicilan hp buat mempermudah mitranya. ini membantu banget buat mitra min https://t.co/L8m1u2Fbol
2	@DediiAjii @Partono_ADjem @pinjollaknat ATM/no rek saya dpinjam sama temen saya bro, sebelumnya saya gak tau sama sekali kalau ATM/no rek saya disalahgunakan untuk bisnis pinjaman online itu, pertama saya tau hal ini dari seseorang yang menghubungi saya m
3	Bunga rendah Tidak ada biaya layanan
4	Gilaaa gini cara nagih pinjaman online, buat yg punya masalah dgn pinjol kalo kalian sdh di permalukan melalui kontak yg ada di HP mending gak usah di bayar, krn bayar nama kita sdh jelek. 3 hari blm jatuh tempo kita di suruh bayar, kalo gak kita dibuat m
5	@AREAJULID Betul banget tuh, temen gw beli kartu baru eh udah di telponin debt collector suruh bayar pinjaman online katanya. Trus di gangguin mulu. Padahal itu kateu bener2 baru beli bnget
6	Ada solusi ngak ya buat yang beginian. Mungkin ada yang melakukan pinjaman online, tapi pakai nomor hp saya sebagai saudara ya. Lagian ini bapak siapa. Kesela aja nih lapor polisi online kayaknya gk pernah berfungsi? #polisi #pinjol https://t.co/wqq3bZRQ2
...	...
500	@goldenclosset @karirfess Temanku di tempat kerjanya digininiin 2 senior, mana buat judi online itu. Slot apa ya sebutannya, kurang ngerti juga. Iya sih tiap gajian diganti, tapi apa ga capek gaji abis buat judi & ganti pinjaman. Iya kalo badan sehat teru

Crawling data dilakukan tiap seminggu sekali dengan menggunakan *keyword* 'Pinjaman Online'. Dan data *limit* tiap *crawling* berjumlah kurang lebih 250 data. Sekitar 400 data akan diambil untuk menjadi data latih dan 100 data menjadi data uji. Data selama bulan oktober sampai desember tahun 2023 akan dijadikan satu dan akan dilakukan proses klasifikasi dengan metode *naive bayes*. Data tersebut akan menjadi hasil dari penelitian ini.

3.2 Data Training dan Data Testing

Data pelabelan akan dipecah atau dibagi menjadi data training atau disebut data latih dan juga data testing atau data uji dengan perbandingan 80:20. Data pelabelan sejumlah 500 data yang akan dibagi menjadi data latih dan uji. Sehingga dengan pembagian tersebut dapat diketahui data latih sejumlah 400 data tweet yang dapat dilihat pada Gambar 4.8, dan data uji sejumlah 100 data tweet dapat dilihat pada Tabel 2. Masing masing komentar selanjutnya diberikan label 1 untuk positif dan 0 untuk negatif.

Tabel 2. Data Training

No	Komentar	Sentimen
1	nah iyak, paman aja bisa ngasih pinjaman online sama cicilan hp buat mempermudah mitranya. ini membantu banget buat mitra min https://t.co/L8m1u2Fbol	1
2	@DediiAjii @Partono_ADjem @pinjollaknat ATM/no rek saya dpinjam sama temen saya bro, sebelumnya saya gak tau sama sekali kalau ATM/no rek saya disalahgunakan untuk bisnis pinjaman online itu, pertama saya tau hal ini dari seseorang yang menghubungi saya m	0
3	Bunga rendah Tidak ada biaya layanan	1
4	Gilaaa gini cara nagih pinjaman online, buat yg punya masalah dgn pinjol kalo kalian sdh di permalukan melalui kontak yg ada di HP mending gak usah di bayar, krn bayar nama kita sdh jelek. 3 hari blm jatuh tempo kita di suruh bayar, kalo gak kita dibuat m	0
5	@AREAJULID Betul banget tuh, temen gw beli kartu baru eh udah di telponin debt collector suruh bayar pinjaman online katanya. Trus di gangguin mulu. Padahal itu kateu bener2 baru beli bnget	0
6	Ada solusi ngak ya buat yang beginian. Mungkin ada yang melakukan pinjaman online, tapi pakai nomor hp saya sebagai saudara ya. Lagian ini bapak siapa. Kesela aja nih lapor polisi online kayaknya gk pernah berfungsi? #polisi #pinjol https://t.co/wqq3bZRQ2	0
...	...	
400	Berterimakasih kepada diri sendiri karena tidak pernah terjerumus ke pinjol demi gaya hidup ??	0

3.3 Text Pre-Processing

Data *crawling* yang diperoleh belum bisa digunakan untuk klasifikasi dikarenakan data masih tidak terstruktur dan banyak *noise*. Tahapan yang perlu dilalui dan pre-processing diantaranya yaitu *cleansing*, *case folding*, *tokenizing* dan *Stopword Removal*.

Tabel 3. Data Text Pre-Processing

Data Set Awal	Hasil Text Pre-processing
Waspada dengan segala Pinjaman Online Ilegal yang berupa dari Aplikasi di media sosial.	['waspada', 'pinjaman', 'online', 'ilegal', 'aplikasi', 'media', 'sosial']

3.4 Pembobotan TF-IDF

Data latih yang telah di tokenizing kemudian dihitung dengan menggunakan rumus tf idf, sehingga tiap term memiliki bobot positif maupun negatif.

No	Kata	Bobot
1	bunga	1.25964
2	rendah	1.82391
3	biaya	1.39794
4	layan	1.75696
5	atm	2.12494
6	no	1.60206
7	rek	2.60206
8	dpinjam	2.60206
9	sama	1.02228
10	temen	1.60206

Gambar 5. Data nilai IDF tiap kata

Gambar 5 merupakan function *hitunfIdf* dimana menggunakan rumus $idf = \text{round}(\log_{10}(\frac{SD}{\$f}), 5)$. Yang akan menghasilkan nilai *\$idf*. Jumlah kata yang telah melalui proses pre-processing sejumlah 2017 data, dimana tiap kata memiliki nilai idf.

No	Kata	Kemunculan	Bobot TF-IDF
1	nah	1	2
2	iyak	1	2.60206
3	paman	1	2.60206
4	aja	8	8.99952
5	ngasih	1	2.30103
6	pinjam	136	26.82328
7	online	105	25.03725
8	sama	7	7.15596
9	cicil	3	5.47173
10	hp	2	3.39794

No	Kata	Kemunculan	Bobot TF-IDF
1	atm	2	4.24988
2	no	9	14.41854
3	rek	2	5.20412
4	dpinjam	1	2.60206
5	sama	36	36.80208
6	temen	13	20.82678
7	bro	1	2.60206
8	gak	21	26.04693
9	tau	10	15.2288
10	sekali	8	12.81648

Gambar 6 Data TF dan TFIDF Kata Positif dan Negatif

Pada Gambar 6 menunjukkan bahwa tiap kata akan dihitung kemunculannya atau biasa disebut TF dan akan memiliki suatu bobot TF-IDF

3.4 Klasifikasi Dengan Algoritma Naive Bayes

Tahap klasifikasi dengan Naive Bayes classifier diawali dengan membentuk bobot setiap term dari data validasi dengan menggunakan teknik tf-idf. Hasil komputasi tf-idf ini kemudian diolah menggunakan algoritma Naive Bayes untuk membangun model probabilistik.

Kalimat sentimen akan melalui proses *preprocessing* dan menghasilkan kalimat yang memiliki suatu nilai positif ataupun negatif dengan melalui perhitungan TfIdf tiap kata.

No	Komentar	Positif	Negatif	Hasil
1	rt butuh uang desak aju dana darurat adakami sekarang limit pinjam hingga rp cicil tan	$P(rt Positif) = 10.40824 + 1 / 3702.269 + 4779.824 = 0.001345$ $P(butuh Positif) = 19.1976 + 1 / 3702.269 + 4779.824 = 0.002381$ $P(uang Positif) = 18.23456 + 1 / 3702.269 + 4779.824 = 0.002268$ $P(desak Positif) = 6.37482 + 1 / 3702.269 + 4779.824 = 0.000869$ $P(aju Positif) = 29.43632 + 1 / 3702.269 + 4779.824 = 0.003588$ $P(dana Positif) = 14.81469 + 1 / 3702.269 + 4779.824 = 0.001864$ $P(darurat Positif) = 2 + 1 / 3702.269 + 4779.824 = 0.000354$ $P(adakami Positif) = 6.37482 + 1 / 3702.269 + 4779.824 = 0.000869$ $P(sekarang Positif) = 5.70927 + 1 / 3702.269 + 4779.824 = 0.000791$ $P(limit Positif) = 13.18256 + 1 / 3702.269 + 4779.824 = 0.001672$ $P(pinjam Positif) = 26.82328 + 1 / 3702.269 + 4779.824 = 0.00328$ $P(hingga Positif) = 2.30103 + 1 / 3702.269 + 4779.824 = 0.000389$ $P(rp Positif) = 3.39794 + 1 / 3702.269 + 4779.824 = 0.000518$ $P(cicil Positif) = 5.47173 + 1 / 3702.269 + 4779.824 = 0.000763$ $P(tan Positif) = 2.60206 + 1 / 3702.269 + 4779.824 = 0.000425$ $P(rt butuh uang desak aju dana darurat adakami sekarang limit pinjam hingga rp cicil tan Positif) = 0.001345 * 0.002381 * 0.002268 * 0.000869 * 0.003588 * 0.001864 * 0.000354 * 0.000869 * 0.000791 * 0.001672 * 0.00328 * 0.000389 * 0.000518 * 0.000763 * 0.000425 * 0.3325 = 1.223874725884E-45$	$P(rt Negatif) = 15.61236 + 1 / 7818.679 + 4779.824 = 0.001319$ $P(butuh Negatif) = 8.95888 + 1 / 7818.679 + 4779.824 = 0.00079$ $P(uang Negatif) = 18.23456 + 1 / 7818.679 + 4779.824 = 0.001527$ $P(desak Negatif) = 0 + 1 / 7818.679 + 4779.824 = 7.9E-5$ $P(aju Negatif) = 2.55968 + 1 / 7818.679 + 4779.824 = 0.000283$ $P(dana Negatif) = 10.77432 + 1 / 7818.679 + 4779.824 = 0.000935$ $P(darurat Negatif) = 6 + 1 / 7818.679 + 4779.824 = 0.000556$ $P(adakami Negatif) = 0 + 1 / 7818.679 + 4779.824 = 7.9E-5$ $P(sekarang Negatif) = 5.70927 + 1 / 7818.679 + 4779.824 = 0.000533$ $P(limit Negatif) = 1.64782 + 1 / 7818.679 + 4779.824 = 0.00021$ $P(pinjam Negatif) = 33.13464 + 1 / 7818.679 + 4779.824 = 0.002709$ $P(hingga Negatif) = 2.30103 + 1 / 7818.679 + 4779.824 = 0.000262$ $P(rp Negatif) = 27.18352 + 1 / 7818.679 + 4779.824 = 0.002237$ $P(cicil Negatif) = 7.29564 + 1 / 7818.679 + 4779.824 = 0.000658$ $P(tan Negatif) = 0 + 1 / 7818.679 + 4779.824 = 7.9E-5$ $P(rt butuh uang desak aju dana darurat adakami sekarang limit pinjam hingga rp cicil tan Negatif) = 0.001319 * 0.00079 * 0.001527 * 7.9E-5 * 0.000283 * 0.000935 * 0.000556 * 7.9E-5 * 0.000533 * 0.00021 * 0.002709 * 0.000262 * 0.002237 * 0.000658 * 7.9E-5 * 0.6675 = 9.0087593041833E-51$	Positif

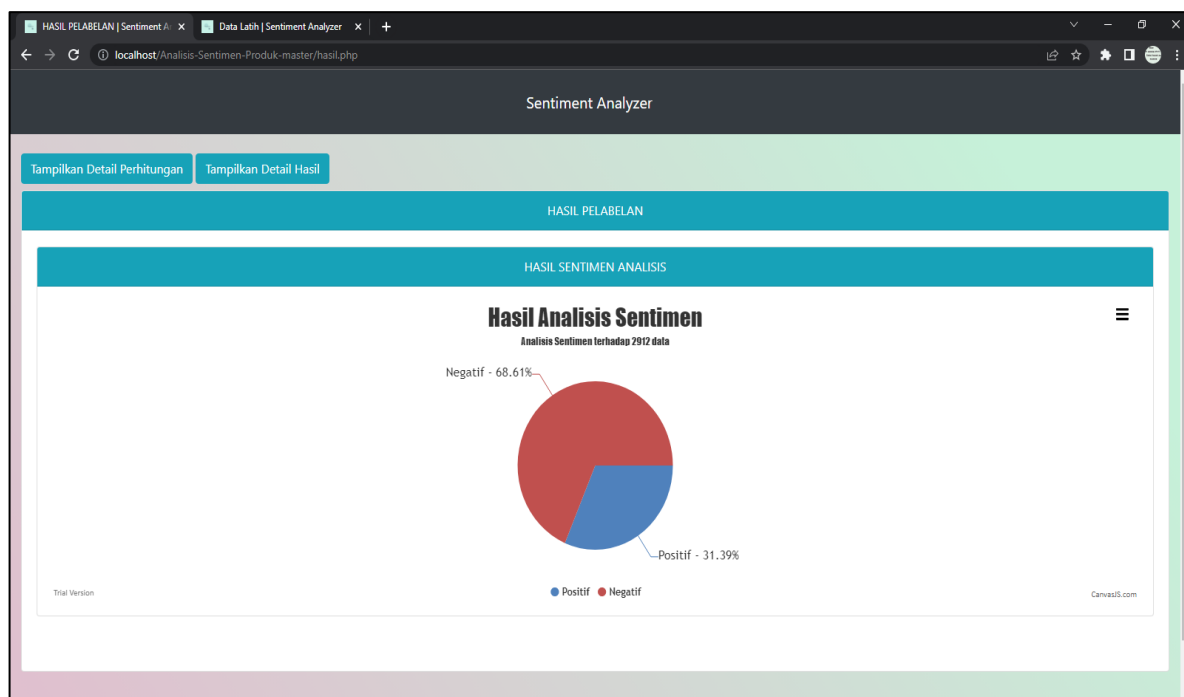
Gambar 7. Gambar Proses Perhitungan Bobot tiap Data Sentimen

Tiap probabilitas kemunculan kata dalam kondisi positif ataupun negatif ditampilkan, yang kemudian akan dihitung probabilitas mana yang mempunyai nilai probabilitas lebih besar. Bisa dilihat pada gambar 7, nilai probabilitas positif lebih besar daripada negatif sehingga hasilnya menunjukkan bahwa kalimat tersebut memiliki nilai sentimen positif.

No	Komentar	Sentimen
1	RT @adakamiofficial: ??Butuh uang mendesak? Ajukan dana darurat di Adakami sekarang! ??Limit Pinjaman Hingga Rp 20.000.000 🟢Bisa cicil, tan...	Positif
2	@SmgMenfess2 Padahal kui bisa jadi Akbar mung nawarke jasa pinjaman online	Positif
3	@BakaTito Hai, Kak Tito. Maaf ya untuk mendapat pesan atau telepon spam. Kami sarankan, Kakak silakan konfirmasi ke pihak yang menghubungi, bahwa Kakak tidak melakukan pinjaman online, dan Kami sarankan juga, Kakak melakukan blokir mandiri diperangkat yang digunakan. Terima kasih??-Dofi	Positif
4	@Telkomsel min ini gimana saya di teror karena nomor saya disalah gunakan oleh orang yang tidak bertanggung jawab untuk melakukan pinjaman online	Negatif
5	Ojek online : OJOL Pinjaman online : PINJOL Kontrakan online : ?????	Negatif
6	7 PINJAMAN Online Terbaik dan Pasti Cair, Jangan TERJEBAK Pinjol Ilegal! Ngeri Kalau Sampai Data Pribadi Terse https://t.co/r84bceflwv	Negatif
7	@ojkindonesia Sore, untuk lapor penyalahgunaan data pribadi bagaimana ya? teman saya semenjak KTPnya hilang jadi terdaftar ke beberpa aplikasi pinjaman online, dia tidak pernah melakukan pinjaman online ke akun tersebut, Tolong di bantu karna jumlahnya tidak sedikit ????	Negatif
8	RT @hyang_wisnu: POLDA METRO: 58 APLIKASI PINJOL ILEGAL SUDAH DIBLOKIR KOMINFO Polda Metro Jaya membongkar 58 aplikasi pinjaman online (pi...	Negatif
9	Info pinjaman bukan online dong! urgent hiks	Positif

Gambar 8. Hasil Perhitungan Kalimat Sentimen

Data komentar yang belum memiliki label akan dilakukan klasifikasi secara otomatis. Pembobotan dengan Tf-Idf dan model klasifikasi dari data latih atau data training sebelumnya yang akan menentukan klasifikasi positif atau negatif terhadap data komentar yang belum memiliki label. Dari total data komentar sebanyak 2931 data, setelah dilabeli secara otomatis dengan model klasifikasi menghasilkan hasil sebagai berikut: sentimen negatif sebanyak 68.61% dengan 1998 data, sentimen positif sebanyak 31.39% dengan 914 data



Gambar 9. Hasil Analisis Sentimen Pinjaman Online

4. KESIMPULAN

Kesimpulan yang dapat diambil dari sentimen analisis menggunakan klasifikasi algoritma Naive Bayes dengan data yang diambil dari Twitter dan pinjaman online sebagai objek penelitian adalah Dari 2931 data yang digunakan, dan telah melalui proses text preprocessing menjadi 2912 data. Sentimen negatif sebanyak 68.61% dengan 1998 data, dan sentimen positif sebanyak 31.39% dengan 914 data. Analisis sentimen dari pengguna Twitter tentang pinjaman online memiliki tingkat akurasi sebesar 80%.

REFERENCES

- [1] I. A. K. S. B. Rahardjo, Budi; Khairul, "Pengaruh Financial Technology (Fintech) Terhadap Perkembangan UMKM di Kota Magelang," *Pros. Semin. Nas. Dan Call Pap. Fak. Ekon. Univ. Tidar*, pp. 347–356, 2019.
- [2] R. Wati, S. Ernawati, P. Studi, S. Informasi, U. Bina, S. Informatika, P. Studi, S. Informasi, U. Nusa, and J. Timur, "Analisis Sentimen Persepsi Publik Mengenai PPKM Pada Twitter Berbasis SVM Menggunakan Python," vol. 06, pp. 240–247, 2021.
- [3] L. Ardiani, H. Sujaini, and T. Tursina, "Implementasi Sentiment Analysis Tanggapan Masyarakat Terhadap Pembangunan di Kota Pontianak," *J. Sist. dan Teknol. Inf.*, vol. 8, no. 2, p. 183, 2020.
- [4] A. Mustofa and R. Novita, "Klasifikasi Sentimen Masyarakat Terhadap Pemberlakuan Pembatasan Kegiatan Masyarakat Menggunakan Text Mining Pada Twitter," *Build. Informatics, Technol. Sci.*, 2022.
- [5] A. Kusuma and A. Nugroho, "Analisa Sentimen Pada Twitter Terhadap Kenaikan Tarif Dasar Listrik Dengan Metode Naive Bayes," *J. Ilm. Teknol. Inf. Asia*, vol. 15, no. 2, p. 137, Dec. 2021.
- [6] A. Perdana, A. Hermawan, and D. Avianto, "Analisis Sentimen Terhadap Isu Penundaan Pemilu di Twitter Menggunakan Naive Bayes Classifier," *J. Sisfokom (Sistem Inf. dan Komputer)*, 2022.
- [7] N. P. G. Naraswati, R. Nooraeni, D. C. Rosmilda, D. Desinta, F. Khairi, and R. Damaiyanti, "Analisis Sentimen Publik dari Twitter Tentang Kebijakan Penanganan Covid-19 di Indonesia dengan Naive Bayes Classification," *SISTEMASI*, 2021.
- [8] M. Yunus, M. Husni, and M. M. Mufadhhal, "Klasifikasi Sentimen Terhadap Badan Penyelenggara Jaminan Sosial (BPJS) Pada Media Sosial Twitter Menggunakan Naive Bayes," *SMATIKA J.*, 2021.
- [9] A. Pramitha, E. Supriyati, and T. Listyorini, "Sentimen Analisis Terhadap Kaum Gay Pada Twitter Menggunakan Metode Naive Bayes," *J. Inf. Eng. Educ. Technol.*, vol. 5, no. 1, pp. 1–4, 2021.
- [10] M. Syarifuddin, "Analisis Sentimen Opini Publik Mengenai Covid-19 Pada Twitter Menggunakan Metode Naive Bayes Dan Knn," *Inti Nusa Mandiri*, vol. 15, no. 1, pp. 23–28, 2020.
- [11] F. Ratnawati, "Implementasi Algoritma Naive Bayes Terhadap Analisis Sentimen Opini Film Pada Twitter," *INOVTEK Polbeng - Seri Inform.*, vol. 3, no. 1, p. 50, 2018.
- [12] C. Prianto, N. H. Harani, and I. Firmansyah, "Analisis Sentimen Terhadap Kandidat Presiden Republik Indonesia Pada Pemilu 2019 di Media Sosial Twitter," *J. Media Inform. Budidarma*, vol. 3, no. 4, p. 405, 2019.

- [13] Maiti and Bidinger, "Implementasi Teknik Web Scraping Dan Klasifikasi Sentimen Menggunakan Metode Naive Bayes Classifier Dan Asosiasi Teks," *J. Chem. Inf. Model.*, vol. 53, no. 9, pp. 1689–1699, 2017.
- [14] P. Nurmawati, E. Supriyati, and T. Listyorini, "Analisis Sentimen Terhadap Penggemar Kpop Di Media Sosial Twitter Menggunakan Naive Bayes (Studi Kasus Penggemar Grup Bts)," *JIEET (Journal Inf. Eng. Educ. Technol.)*, vol. 4, no. 2, pp. 86–89, 2020.
- [15] A. Z. Amrullah, A. Sofyan Anas, and M. A. J. Hidayat, "Analisis Sentimen Movie Review Menggunakan Naive Bayes Classifier Dengan Seleksi Fitur Chi Square," *Jurnal*, vol. 2, no. 1, pp. 40–44, 2020.
- [16] W. Yulita, E. D. Nugroho, and M. H. Algifari, "Analisis Sentimen Terhadap Opini Masyarakat Tentang Vaksin Covid-19 Menggunakan Algoritma Naive Bayes Classifier," *Ejurnal.Tekokrat.Ac.Id*, vol. 2, no. 2, pp. 1–9, 2021.
- [17] A. C. Sitepu, W. Wanayumini, and Z. Situmorang, "Analisis Kinerja Support Vector Machine dalam Mengidentifikasi Komentar Perundungan pada Jejaring Sosial," *J. Media Inform. Budidarma*, vol. 5, no. 2, p. 475, 2021.
- [18] D. Darwis, N. Siskawati, and Z. Abidin, "Penerapan Algoritma Naive Bayes Untuk Analisis Sentimen Review Data Twitter Bmkg Nasional," *J. Tekno Kompak*, vol. 15, no. 1, p. 131, 2021.
- [19] E. S. Basryah, A. Erfina, C. Warman, D. Digital, and P. Store, "Analisis Sentimen Aplikasi Dompot Digital Di Era 4 . 0 Pada Masa Pandemi Covid-19 Di Play Store," pp. 189–196, 2021.
- [20] S. Syarli and A. Muin, "Metode Naive Bayes Untuk Prediksi Kelulusan (Studi Kasus: Data Mahasiswa Baru Perguruan Tinggi)," *J. Ilm. Ilmu Komput.*, vol. 2, no. 1, pp. 22–26, 2016.