

IMPLEMENTASI NAÏVE BAYES CLASSIFIER UNTUK ANALISIS SENTIMEN APLIKASI PINJAMAN ONLINE PADA MEDIA SOSIAL TWITTER

Shinta May Khoiria Azzahra¹✉^{#1}, Nirma Ceisa Santi^{*2}, Sunu Wahyudi³

*Program studi Teknik Informatika, Program Studi Sistem Informasi, Program Studi Sistem Komputer
Universitas Nahdlatul Ulama Sunan Giri*

¹Shintamay25@gmail.com

²nirmaCeisa@unugiri.ac.id

³Sunuwahyudi@unugiri.ac.id

Abstract — Sistem analisis sentimen adalah sistem yang digunakan untuk melakukan proses analisis otomatis pada tweet aplikasi pinjaman online berbahasa Indonesia untuk mendapatkan informasi termasuk informasi sentimen yang menjadi bagian dari tweet tersebut. Data tersebut diklasifikasikan dengan menggunakan Naive Bayes. Sistem analisis sentimen dibagi menjadi 5 (lima) tahap, yaitu scraping, pre-processing, pembobotan kata, pembangunan model dan klasifikasi sentimen. Metode TF-IDF (Term Frequency - Inverse Document Frequency) digunakan untuk pembobotan kata. Data yang ada akan diklasifikasikan menjadi 2 (dua) kelas positif dan negatif. Hasil pengujian dengan 500 data dari tweet aplikasi Kredivo menunjukkan bahwa pada pengujian 2 kelas (negatif dan positif) didapatkan hasil terbaik pada 90% data latih dan 10% data uji dengan nilai akurasi sebesar 91,12%.

Keywords— naïve bayes classifier; online loan application; sentiment analysis

I. PENDAHULUAN

Perkembangan teknologi semakin hari semakin meningkat. Hal ini dapat dilihat dari berbagai aspek kehidupan yang serba digital dan semakin canggih, mulai dari layanan belanja yang semakin mudah dengan adanya sistem e-commerce, layanan teknologi finansial, media sosial, dan masih banyak lagi sistem lainnya yang memudahkan pekerjaan manusia. Perkembangan teknologi telah membawa banyak perubahan dalam gaya hidup umat manusia, mulai dari yang muda hingga yang tua. Kecanggihan internet sangat memudahkan para pelaku bisnis untuk menawarkan barang atau jasa kepada konsumen, begitu juga sebaliknya, konsumen dapat dengan mudah mencari barang yang mereka butuhkan.

Dampak dari perkembangan teknologi tidak hanya memberikan dampak positif tetapi juga memberikan dampak negatif bagi setiap individu yang mengakibatkan pembiayaan masyarakat semakin tinggi karena kebutuhan masyarakat yang semakin meningkat. Terutama pada perilaku konsumtif di kalangan remaja yang mengikuti pola gaya hidup masyarakat dari media sosial, berlomba-lomba mengikuti perkembangan yang terjadi di media sosial tanpa mempertimbangkan pro dan kontra. Perilaku konsumtif sebagai gaya hidup baru dalam tatanan sosial tidak lepas dari kemudahan yang didapatkan akibat pertumbuhan dunia

online kita saat ini, hal ini berdampak pada perekonomian Indonesia [2].

Hal ini menciptakan suatu perubahan sosial yang merubah perilaku masyarakat menjadi konsumerisme yang ditandai dengan kehidupan yang serba mewah dan berlebihan. Barang-barang yang sedang menjadi tren memberikan kepuasan dan kesenangan bagi orang yang memilikinya, meskipun barang tersebut bukan merupakan barang primer [3]. ketika gaya hidup meningkat, maka perilaku konsumtif juga akan meningkat, artinya semakin mewah gaya hidup seseorang, maka perilaku konsumtif juga akan meningkat, begitu juga sebaliknya, jika gaya hidup seseorang menurun, maka pola perilaku konsumtif orang tersebut juga akan menurun [4].

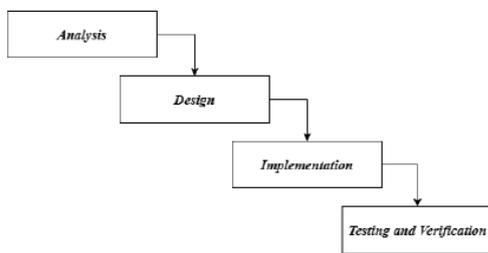
Lembaga pembiayaan memanfaatkan situasi ini untuk memenuhi kebutuhan masyarakat, terutama masyarakat yang berperilaku konsumtif, baik pembiayaan dana maupun barang modal. Di era digital ini banyak sekali lembaga yang menawarkan pinjaman online dan kredit online, hal ini sangat membantu untuk membiayai kebutuhan masyarakat yang membutuhkan dana dengan cepat.

Fitur kredit online menawarkan layanan kredit tanpa menggunakan kartu kredit, salah satu aplikasi yang menawarkan hal tersebut adalah Aplikasi Kredivo, Akulaku, Indodana, Julo, Kredit Pintar, Tunaiku, Rupiah Cepat, dan

masih banyak lagi. Dengan banyaknya aplikasi pinjaman online yang tersedia, calon pengguna tentu akan merasa bingung untuk memilih aplikasi mana yang terbaik dari segi fitur, pelayanan dan feedback pengguna. Untuk mengetahuinya, perlu dilakukan analisis sentimen untuk mengetahui persepsi pengguna aplikasi pinjaman online sehingga calon pengguna dapat membandingkan aplikasi mana yang cocok untuk digunakan.

Pada penelitian ini, peneliti merancang dan membangun sebuah aplikasi analisis sentimen dengan menggunakan bahasa pemrograman PHP untuk memudahkan masyarakat dalam memilih aplikasi pinjaman online. Untuk melakukan analisis sentimen terhadap aplikasi pinjaman online, peneliti menggunakan data tweet aplikasi Kredivo pada media sosial Twitter sebagai objek. Pada proses klasifikasi, algoritma Naïve Bayes digunakan untuk mengklasifikasikan sentimen positif dan negatif.

II. METODE PENELITIAN



Gambar 1 Metode Waterfall

Rancang bangun aplikasi analisis sentimen menggunakan metode waterfall. Pengembangan dan pembangunan perangkat lunak dapat dilihat pada Gambar 1 yang terdiri dari analisis, desain, implementasi, pengujian dan verifikasi, Pada proses analisis penulis melakukan pengumpulan data mengenai kebutuhan yang diperlukan untuk perhitungan dan aplikasi. Setelah dilakukan analisis, penulis menentukan alur proses pembangunan perhitungan dan aplikasi, meliputi perancangan basis data, flowchart, activity diagram, use case, mock up aplikasi. Implementasi aplikasi dibangun dengan menggunakan bahasa pemrograman PHP dan metode Naïve Bayes. Setelah aplikasi selesai dibangun, akan dilakukan pengujian dengan menggunakan sistem black box untuk mengevaluasi aplikasi yang telah dibangun.

A. Twiter

Sumber data penelitian ini adalah Twitter, yang merupakan layanan jejaring sosial yang membantu pengguna berbagi dan membaca pesan berbasis teks hingga 140 karakter. [5]. Selain itu, di Twitter tidak ada batasan dalam berbagi konten, pengguna bebas berekspresi sesuai dengan apa yang mereka inginkan tanpa adanya takedown dari sistem. Sehingga sumber data Twitter sangat cocok digunakan sebagai sumber data dalam penelitian yang membahas tentang aplikasi pinjaman online. Aplikasi pinjaman online adalah platform penyedia fasilitas layanan jasa keuangan yang beroperasi secara online, aplikasi atau layanan peminjaman uang berbasis teknologi informasi merupakan salah satu jenis dari Financial Technology

Implementation. Artikel harus terdiri dari minimal delapan halaman. Jangan mengubah ukuran huruf atau spasi untuk memasukkan lebih banyak teks pada jumlah halaman yang terbatas. [6].

B. Sentimen Analisis

Analisis Sentimen adalah cabang dari penelitian text mining, yang berhubungan dengan pengelolaan data dalam bidang yang lebih luas untuk tujuan tertentu[7]. Analisis sentimen dapat dilakukan dengan mudah. Dengan data mining, yang merupakan proses semi-otomatis yang menggunakan teknik matematika, kecerdasan buatan, statistik, dan pembelajaran mesin untuk mengekstrak dan mengidentifikasi informasi tertentu yang berpotensi berguna yang tersimpan dalam database [8]. Dengan data yang besar, diperlukan preprocessing agar proses pembobotan dan klasifikasi dapat dilakukan.

C. Naïve bayes

Metode yang digunakan untuk melakukan proses klasifikasi adalah Naïve Bayes. Naïve Bayes merupakan metode klasifikasi berbasis probabilitas sederhana yang dimaksudkan untuk digunakan dengan asumsi tidak ada saling ketergantungan antara satu kelas dengan kelas yang lain[9]. Teorema Bayes adalah teorema yang mengacu pada konsep probabilitas bersyarat. Probabilitas Bayesian digunakan untuk menyelesaikan masalah ketidakpastian data berdasarkan rumus Bayes yang telah dinyatakan[10]. Naive Bayes menggunakan konsep probabilitas dalam menentukan kelas dokumen. Metode ini menggunakan asumsi bahwa dalam sebuah dokumen kemunculan sebuah kata tidak mempengaruhi kemunculan kata lainnya dan tidak munculnya sebuah kata tidak mempengaruhi tidak munculnya kata lainnya. Naïve Bayes classifier mengasumsikan bahwa ada atau tidak adanya suatu fitur tertentu pada suatu kelas tidak mempengaruhi keberadaan fitur yang lain.

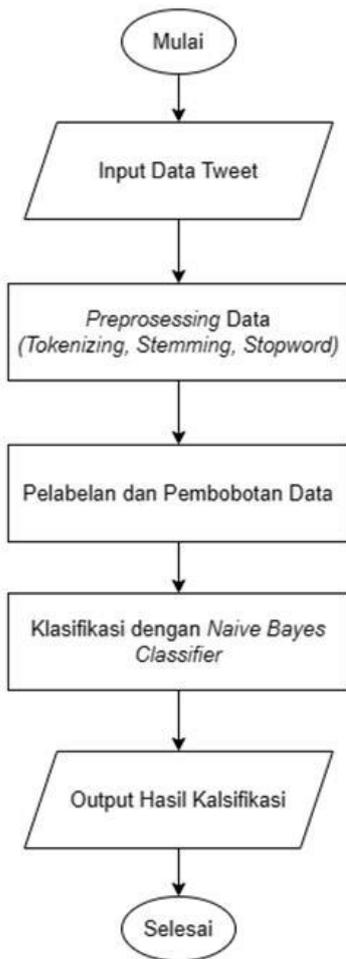
$$P(A|B) = \frac{P(A|B)P(A)}{P(B)} \quad (1)$$

Pada rumus di atas merupakan rumus dari naïve bayer, berikut penjelesan rumusnya.

- A : Data yang kelasnya belum diketahui
- B : Hipotesis data B dari suatu kelas
- $P(A/B)$: Probabilitas hipotesis A dalam kondisi B (probabilitas a posteriori)
- $P(A)$: Probabilitas hipotesis A (probabilitas prior)
- $P(B/A)$: Probabilitas B di bawah kondisi hipotesis A
- $P(B)$: Probabilitas B

D. Alur penelitaan

Dalam penelitian ini, alur proses kegiatan analisis sentimen akan dijelaskan dalam alur berikut:



1. Pengumpulan data
 Data yang digunakan dalam penelitian ini diambil dari kumpulan tweet berbahasa Indonesia yang diambil oleh Twitter. Data tweet ini diperoleh dengan membuat program scraping menggunakan bahasa pemrograman Python. Scraping adalah proses pengambilan data pada sebuah website secara otomatis dan menyimpannya langsung pada perangkat komputer[11]. Pada proses scraping, secara otomatis akan mengambil data tweet yang mengandung kata berupa nama aplikasi Kredivo. Yang kemudian siap untuk dilakukan proses preprocessing
2. Preprocessing tokenizing
 Pada tahap ini, tweet akan diperiksa dari karakter pertama hingga karakter terakhir. Jika sebuah karakter tidak memiliki pemisah kata seperti titik (.), koma (,), spasi, dan tanda pemisah lainnya, maka karakter tersebut akan digabungkan dengan karakter berikutnya[12].
3. Preprocessing stemming
 Pada tahap ini, kumpulan tweet yang telah melewati tahap tokenizing akan melalui tahap stemming. Setiap kata dalam tweet akan diperiksa. Apakah ada kata berimbuhan? Jika terdapat kata berimbuhan, seperti me-, di-, -kan dan lain-lain. Maka imbuhan tersebut akan dihilangkan dari kata tersebut dan hanya kata

dasarnya saja yang akan diambil untuk proses selanjutnya[13].

4. Preprocessing stopwords
 Pada tahap ini, kumpulan tweet yang telah melewati tahap tokenizing akan melalui tahap stopword removal. Setiap kata yang ada di dalam tweet akan diperiksa. Jika terdapat kata-kata yang tidak ada hubungannya dengan analisis sentimen, seperti kata penghubung, kata depan, dan kata ganti, maka kata-kata tersebut akan dihilangkan[14].

III. HASIL DAN PEMBAHASAN

Implementasi merupakan tindakan untuk menjalankan rencana yang telah disusun secara matang dan terperinci. Implementasi dilakukan setelah desain dianggap baik. Implementasi dalam pengembangan perangkat lunak biasanya disebut dengan konstruksi, yaitu suatu proses perubahan spesifikasi sistem menjadi yang dapat dijalankan atau dapat dikaitkan sebagai tahap penerjemahan hasil desain/logika dan fisik ke dalam kode program.

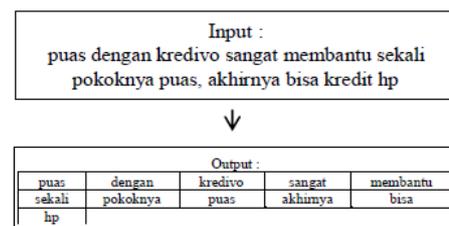
Berdasarkan pemaparan dan metode penelitian yang telah dijelaskan. Selanjutnya akan dijelaskan hasil dan pembahasan dari penelitian dan aplikasi yang telah dikembangkan

A. Scraping data

Setelah dilakukan scraping data menggunakan pemrograman Python terhadap data tweet, aplikasi Kredivo menghasilkan 500 data tweet yang akan dibagi menjadi dua yaitu, 100 tweet sebagai data latih dan 400 tweet sebagai data uji

B. Tokenizing

Pada tahap ini, data tweet yang telah didapatkan pada proses scraping akan dilakukan preprocessing dengan melakukan tokenizing. Gambaran dari tahap tokenizing dapat dilihat pada bagan berikut:



Gambar 2 Proses Tokenizing

C. Steeming

Pada tahap ini, data tweet yang telah melalui preprocessing tokenizing akan dilakukan preprocessing stemming. Gambaran tahapan stemming dapat dilihat pada gambar berikut:

Input :				
puas	dengan	kredivo	sangat	membantu
sekali	pokoknya	puas	akhirnya	bisa
hp				

↓

Output :				
puas	dengan	kredivo	sangat	bantu
sekali	pokok	puas	akhir	bisa
hp				

Gambar 3 Processing Steeming

D. Stopword

Pada tahap ini, data tweet yang telah melalui preprocessing stemming akan mengalami preprocessing stopwords. Gambaran dari tahap stopwords dapat dilihat pada gambar berikut:

Input :				
puas	dengan	kredivo	sangat	bantu
sekali	pokok	puas	akhir	bisa
hp				

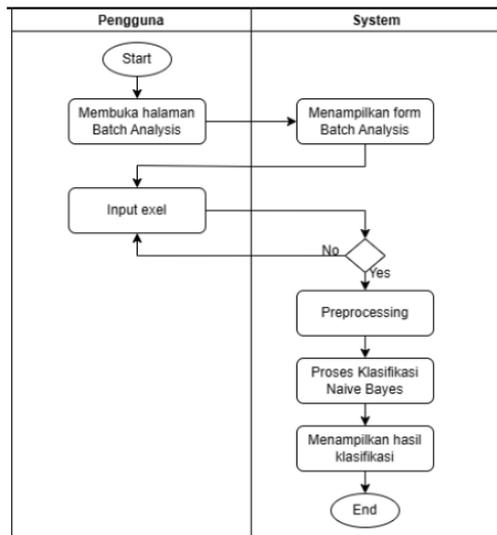
↓

Output :				
puas	kredivo	bantu	puas	kredit
hp				

Gambar 4 Preprocessing Stopword

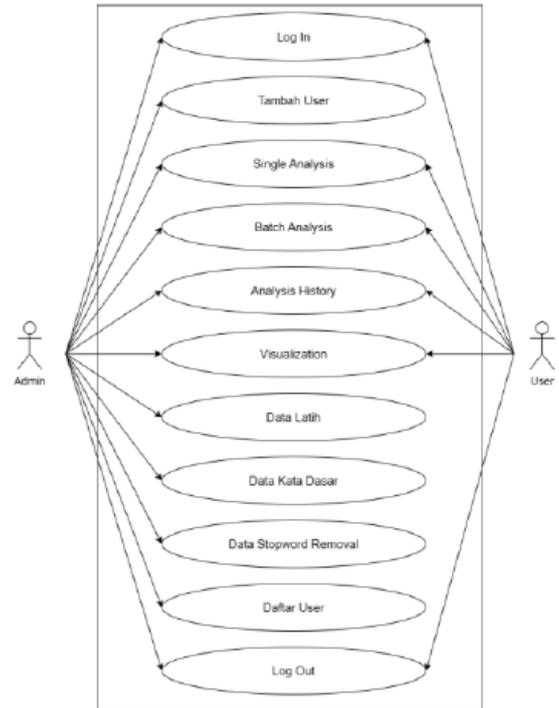
E. Naïve bayes analisis

Setelah dilakukan preprocessing, data tweet siap untuk dilakukan proses pembobotan dan pelabelan dan selanjutnya dilakukan klasifikasi menggunakan Naïve Bayes. Pada gambar 5 merupakan activity diagram untuk analisis sentimen pengajuan pinjaman online



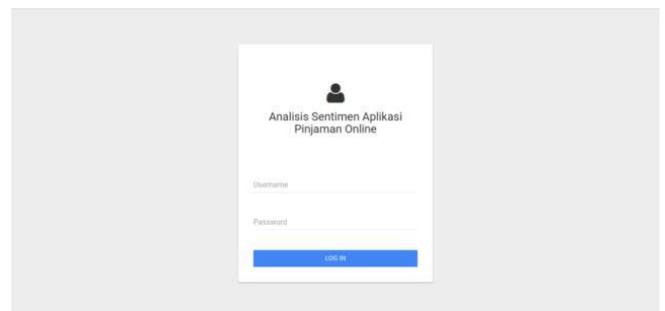
Gambar 5 Activity Diagram

Gambar tersebut menjelaskan alur kegiatan sistem dalam analisis sentimen, mulai dari awal hingga akhir. Pada aplikasi ini terdapat dua role yang berbeda yaitu admin dan user, masing-masing role memiliki kapasitas dan hak akses masing-masing yang akan dijelaskan pada gambar 6.



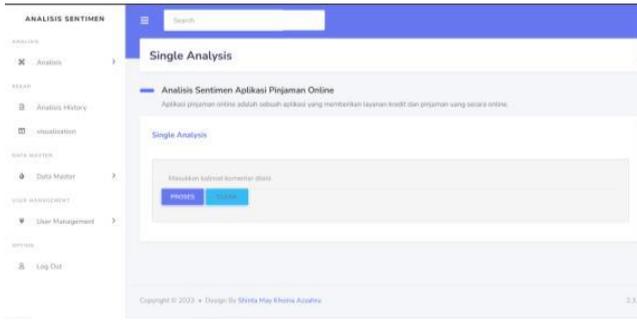
Gambar 6 Use Case

Perancangan aplikasi telah berhasil diimplementasikan menghasilkan sebuah aplikasi berbasis website yang dibangun menggunakan bahasa pemrograman PHP dengan framework CodeIgniter. Aplikasi ini dapat diakses oleh admin dan pengguna, pengguna yang dimaksud adalah masyarakat atau calon pengguna aplikasi pinjaman online yang sedang mencari aplikasi yang dirasa cocok. Aplikasi yang dibangun memiliki fitur login, analisis tunggal, analisis batch, visualisasi, history, master data dan logout. Yang akan dijelaskan pada setiap halaman pada gambar berikut.



Gambar 7 Halaman Login

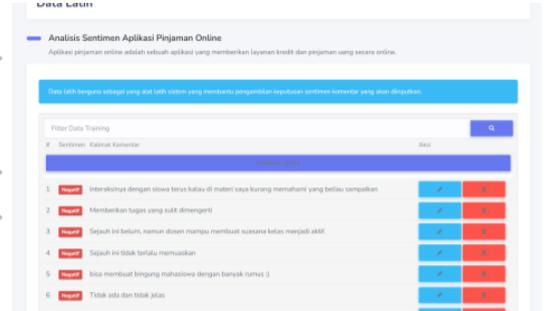
Berikut ini adalah tampilan form login untuk user sebelum dapat masuk ke dalam aplikasi. Pengguna dikelompokkan menjadi 2, yaitu: Admin dan user. User dapat memasukkan username dan password yang telah disediakan. Halaman login dapat dilihat seperti pada gambar 7 berikut.



Gambar 8 Halaman Analisis

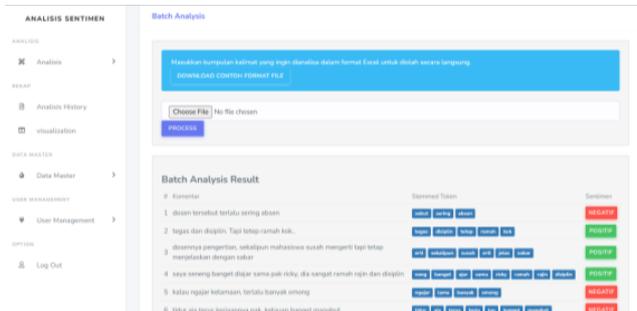
Halaman Single Analysis memiliki navbar dan juga sidebar, sidebar terdiri dari beberapa menu, pada bagian body sendiri terdapat control form untuk menginput kalimat atau tweet yang akan dianalisa sentimennya, terdapat dua tombol untuk submit dan untuk membersihkan atau menghapus kalimat. Single Analysis dapat digunakan oleh admin dan pengguna.

Pada halaman visualisasi terdapat diagram untuk menunjukkan visualisasi yang telah diinput. Terdapat dua diagram, yaitu diagram lingkaran untuk menunjukkan persentase dan diagram batang untuk menunjukkan jumlah. Halaman ini dapat diakses oleh admin dan user. Halaman visualisasi dapat dilihat seperti pada gambar 11 berikut



Gambar 11 Halaman Training

Dengan menggunakan aplikasi ini, calon pengguna aplikasi pinjaman online dapat membandingkan dan mengkomparasi antar aplikasi untuk menemukan aplikasi yang terbaik, baik dari segi fitur maupun persepsi komunitas Twitter. Aplikasi ini mendukung semua jenis data, baik itu CSV formal maupun XLX, sehingga memudahkan pengguna aplikasi. Pada pengujian metode yang dilakukan dengan menggunakan aplikasi rapid miner, berdasarkan hasil analisis terhadap aplikasi yang telah dibangun didapatkan akurasi sebesar 91,12% yang berarti aplikasi ini valid dan dapat digunakan karena memiliki tingkat akurasi yang tinggi. Metode pengujian dapat dilihat pada gambar 12 berikut



Gambar 9 Analisis Batch analysis

Pada analisis batch di bagian body terdapat keterangan yang berisi format file Excel yang sesuai dan pengguna dapat mengunduh file tersebut. Kemudian terdapat choose file untuk mengimpor file Excel dan tombol process untuk melakukan proses analisis. Kemudian hasil dari analisis sentimen akan muncul dalam bentuk tabel, berisi angka, komentar, data preprocessing dan sentimen yang dideskripsikan dalam bentuk rentang, jika positif maka sukses, negatif bahaya dan netral warning. Analisis batch dapat diakses oleh admin dan user. Halaman analisis batch dapat dilihat seperti pada Gambar 9 berikut.

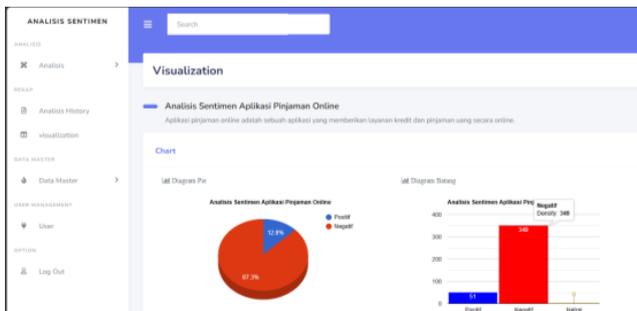
accuracy: 91.12%

	true positif	true negatif	class precision
pred. positif	280	52	84.34%
pred. negatif	1	264	99.62%
class recall	99.64%	83.54%	

Gambar 12 Akurasi

IV. SIMPULAN

Kesimpulan yang dapat diambil dari implementasi Naïve Bayes untuk analisis sentimen terhadap aplikasi pinjaman online di Twitter adalah sistem klasifikasi analisis sentimen dibangun dengan menggunakan bahasa pemrograman PHP, sistem dibangun dengan menggunakan algoritma Naïve Bayes classifier untuk mengetahui persepsi dan sentimen pengguna atau masyarakat terhadap aplikasi pinjaman online, dan pada penelitian ini menggunakan data tweet aplikasi Kredivo sebagai data uji. Hasil yang diperoleh dari implementasi Naïve Bayes dalam mengklasifikasikan data tweet aplikasi Kredivo ke dalam kelas positif dan negatif dengan 100 data latih dan 400 data uji diperoleh hasil klasifikasi kelas negatif sebesar 87,3%, kelas positif 12,8%. Sehingga dapat diketahui bahwa persepsi pengguna aplikasi Kredivo adalah negatif. Sedangkan berdasarkan hasil uji akurasi yang dilakukan pada Rapid Miner, metode Naïve



Gambar 10 Halaman Visualisasi Data

Bayes dalam mengklasifikasikan data tweet aplikasi Kredivo menghasilkan akurasi sebesar 91,12%.

DAFTAR PUSTAKA

- [1] A. Muhammadin and I. A. Sobari, "Analisis Sentimen Pada Ulasan Aplikasi Kredivo Dengan Algoritma Svm Dan Nbc," *Reputasi J. Rekayasa Perangkat Lunak*, vol. 2, no. 2, pp. 85–91, 2021, doi: 10.31294/reputasi.v2i2.785.
- [2] E. Dewi, "Potret Pendidikan di Era Globalisasi Teknosentrisme dan Proses Dehumanisasi," *Sukma J. Pendidik.*, vol. 3, no. 1, pp. 93–116, 2019, doi: 10.32533/03105.2019.
- [3] I. Maulana, J. M. br. Manulang, and O. Salsabila, "Pengaruh Social Media Influencer Terhadap Perilaku Konsumtif di Era Ekonomi Digital," *Maj. Ilm. Bijak*, vol. 17, no. 1, pp. 28–34, 2020, doi: 10.31334/bijak.v17i1.823.
- [4] R. Wahyuni, H. Irfani, and I. A. Syahrina, "Pengaruh Gaya Hidup Dan Literasi Keuangan Terhadap Perilaku Konsumtif Berbelanja Online Pada Ibu Rumah Tangga Di Kecamatan Lubuk Begalung Kota Padang," *J. Benefita*, vol. 4, no. 3, pp. 548–559, 2019.
- [5] D. Duei Putri, G. F. Nama, and W. E. Sulistiono, "Analisis Sentimen Kinerja Dewan Perwakilan Rakyat (DPR) Pada Twitter Menggunakan Metode Naive Bayes Classifier," *J. Inform. dan Tek. Elektro Terap.*, vol. 10, no. 1, pp. 34–40, 2022, doi: 10.23960/jitet.v10i1.2262.
- [6] D. S. Utami and A. Erfina, "Analisis Sentimen Pinjaman Online di Twitter Menggunakan Algoritma Support Vector Machine (SVM)," *SISMATIK (Seminar Nas. Sist. Inf. dan Manaj. Inform.*, vol. 1, no. 1, pp. 299–305, 2021.
- [7] 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, doi: 10.33365/jtk.v15i1.744.
- [8] D. P. Utomo and M. Mesran, "Analisis Komparasi Metode Klasifikasi Data Mining dan Reduksi Atribut Pada Data Set Penyakit Jantung," *J. Media Inform. Budidarma*, vol. 4, no. 2, p. 437, 2020, doi: 10.30865/mib.v4i2.2080.
- [9] A. Ridwan, "Penerapan Algoritma Naive Bayes Untuk Klasifikasi Penyakit Diabetes Mellitus," *J. SISKOM-KB (Sistem Komput. dan Kecerdasan Buatan)*, vol. 4, no. 1, pp. 15–21, 2020, doi: 10.47970/siskom-kb.v4i1.169.
- [10] R. D. Fitriani¹, Hasbi Yasin², and Tarno³, "Penanganan Klasifikasi Kelas Data Tidak Seimbang Dengan Random Oversampling Pada Naive Bayes," *GAUSSIAN*, vol. 10, pp. 11–20, 2021.
- [11] O. S. Udang, M. Tabaru, E. A. M. Sampetoding, and E. S. Manapa, "Pengolahan Data Siswa SMA Negeri 1 Sambuara Kabupaten Kepulauan Talaud Pada Aplikasi DAPODIK," *J. Dyn. Saint*, vol. 6, no. 1, pp. 7–11, 2021, doi: 10.47178/dynamicsaint.v6i1.1193.
- [12] H. Syah and A. Witanti, "Analisis Sentimen Masyarakat Terhadap Vaksinasi Covid-19 Pada Media Sosial Twitter Menggunakan Algoritma Support Vector Machine (Svm)," *J. Sist. Inf. dan Inform.*, vol. 5, no. 1, pp. 59–67, 2022, doi: 10.47080/simika.v5i1.1411.
- [13] [M. Resa, A. Yudianto, A. Rahim, P. Sukmasetya, and R. A. Hasani, "PERBANDINGAN METODE SUPPORT VECTOR MACHINE DENGAN METODE LEXICON DALAM ANALISIS SENTIMEN BAHASA," vol. 6, no. 1, 2022.
- [14] W. Yulita, "Analisis Sentimen Terhadap Opini Masyarakat Tentang Vaksin Covid-19 Menggunakan Algoritma Naive Bayes Classifier," *J. Data Min. dan Sist. Inf.*, vol. 2, no. 2, p. 1, 2021, doi: 10.33365/jdmsi.v2i2.1344.