

Analisis Sentimen Ulasan Pengguna Aplikasi Pinjaman Online Di Google Playstore Menggunakan Naive Bayes

Faisal Akbar Pangestu¹, Irwansyah^{2*}

akbarfaisall197@gmail.com¹, irwansyah@uhamka.ac.id²

^{1,2} Universitas Muhammadiyah Prof. DR. HAMKA Jakarta

Informasi Artikel

Diterima : 21 Mei 2024

Direvisi : 19 Jun 2024

Disetujui : 8 Agu 2024

Kata Kunci

AdaKami, google playstore, Naive Bayes

Abstrak

Aplikasi AdaKami merupakan platform pinjaman online yang tidak memerlukan jaminan dan bertujuan untuk memberikan akses keuangan yang berkualitas kepada jutaan individu di Indonesia. Pada google playstore aplikasi AdaKami di unduh sebanyak 10 juta kali. Permasalahan pada penelitian ini adalah bagaimana performa algoritma Naive Bayes untuk mengklasifikasi data pada ulasan aplikasi AdaKami seberapa banyak ulasan yang mengandung nilai positif maupun negatif. Tujuan dari penelitian ini adalah bagaimana menerapkan algoritma Naive Bayes pada ulasan aplikasi AdaKami. Manfaat dari penelitian ini mengetahui keakuratan metode Naive Bayes dalam menganalisis ulasan pada google playstore. Penelitian ini menghasilkan evaluasi dari klasifikasi algoritma Naive Bayes pada ulasan aplikasi AdaKami pada google playstore dan mendapatkan hasil nilai accuracy sebesar 78,92%, precision 78,79%, dan recall 74,82%. Pada penelitian ini lebih banyak ulasan negatif dibandingkan ulasan positif.

Keywords

AdaKami, google playstore, Naive Bayes

Abstract

The AdaKami application is an online loan platform that attempts to give millions of people in Indonesia high-quality financial access without the need for collateral. Ten million downloads of the AdaKami app have been made through the Google Playstore. The issue in this study is how well the AdaKami application reviews are classified using the Naive Bayes method, specifically how many of the reviews have positive or negative values. How to use the Naive Bayes method for the AdaKami application review is the goal of this study. The knowledge of the accuracy of the Naive Bayes approach in examining Google Playstore evaluations is the advantage of this research. The investigation produced a classification assessment of the Naive Bayes algorithm in Google evaluations of the AdaKami application.

A. Pendahuluan

Pada revolusi internet yang tumbuh pesat, teknologi digital sangat membantu manusia. Pada dasarnya manusia akan membutuhkan ekonomi yang cukup untuk mencapai kebutuhan mendasar agar tercapai kepuasan dan kesejahterannya. Hal tersebut dimanfaatkan teknologi digital di Indonesia cukup besar dan memberikan dampak bagi beberapa sektor, salah satunya di sektor bisnis yang memunculkan perdagangan *online* atau *e-commerce*. Tidak hanya memberikan dampak bagi industri perdagangan, tetapi juga memberikan dampak terhadap industri keuangan Indonesia [1].

Pinjaman online semakin banyak dikenal di beberapa tahun terakhir, dimana pinjaman online sangat cepat dan mudah digunakan masyarakat untuk mengakses kredit. Bagaimana, dasar teori dapat menunjukkan bahwa pinjaman *online* dapat menimbulkan dampak negatif pada kehidupan manusia, dimana mampu menimbulkan kekhawatiran finansial dan tekanan psikologis (Ryu, S., & Fan, L, 2023). Hal ini berlaku pada biaya kredit yang tinggi seperti *payday loan*, yang mampu menjebak pengguna dalam bentuk utang. Penyebab masyarakat menggunakan pinjaman online dikarenakan mudahnya di akses aplikasi tersebut [2].

Munculnya pinjaman *online* di Indonesia, OJK mendata *fintech peer-to-peer lending* yang telah terdaftar per 31 Agustus 2021 sebanyak 106 perusahaan. Adapun salah satu perusahaan pinjaman berbasis *online* yang teregistrasi di OJK yaitu PT Pembinaan Digital Indonesia (AdaKami), layanan *fintech Peer-to-Peer* (P2P) 824 yang mempunyai izin KEP-128/D.05/2019 di bidang Pembiayaan *Fintech* [3].

Analisis sentimen salah satu teknik populer yang digunakan para ilmuwan untuk mengklasifikasikan ulasan positif atau negatif. Digunakannya analisis sentimen untuk menganalisis ulasan pengguna terhadap aplikasi pinjaman *online* di *Google Play Store*. Algoritma *Naive Bayes* merupakan pilihan populer untuk menganalisis suatu sentimen dikarenakan kesederhanaan dan kecepatannya. Dalam metode ini, setiap ulasan diubah menjadi vektorisasi, dan pengklasifikasian *Naive Bayes* dapat memprediksi sentimen baru. Terdapat beberapa penelitian telah menggunakan metode ini untuk menganalisis ulasan pengguna terhadap aplikasi *mobile fintech* dan layanan P2P lending di Indonesia [4]. Dimana setelah menggunakan metode ini peneliti mendapat keakuratan 85% hingga 90% [5]. Dengan menganalisa sentimen dari ulasan pengguna, peneliti dapat memiliki wawasan tentang perspektif masyarakat terhadap pinjaman *online*.

Pengumpulan data menggunakan teknik *web scraping* dengan pengambilan data dari ulasan pengguna aplikasi AdaKami pada situs *google play store*. Teknik *web scraping* merupakan pengambilan data atau informasi dengan skala besar dengan praktis. Pada umumnya dibangun menggunakan bahasa markup misalnya HTML atau XHTML. Tujuan digunakannya teknik *web scarping* tersebut adalah memperoleh data untuk diolah dan dianalisa [6]. Penelitian lainnya juga dilakukan oleh Artanti Inez Tanggraeni, melkior N. N Sitokdana dengan judul Analisis sentimen aplikasi *e-Government* pada *google play* menggunakan algoritma *Naive Bayes* Penelitian ini membahas mengenai lembaga pemerintahan yang telah menerapkan e-gov. mengambil data set pada ulasan *google playstore* menggunakan *extension tool* dari chrome dalam bentuk excel sebanyak 674 ulasan

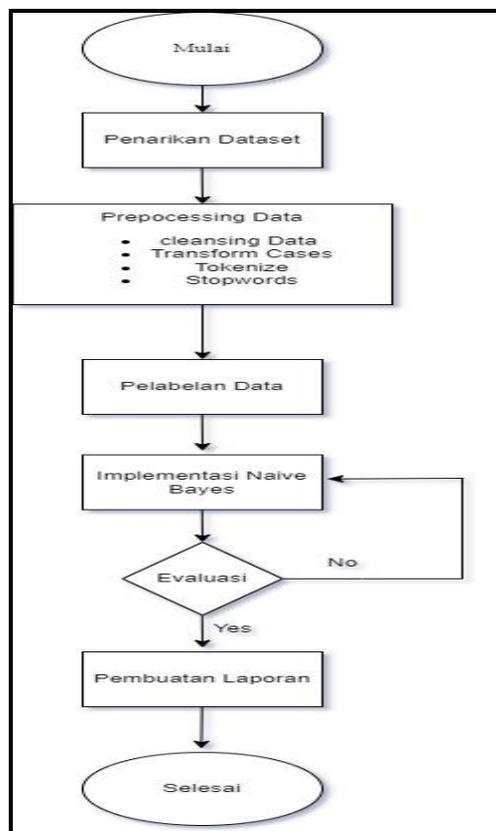
dengan sentimen positif dan negatif. Kemudian melakukan *text processing* data setelah bersih data melakukan pembobotan dengan fitur TF-Idf, membuat data latih dan uji, melakukan klasifikasi *Naive Bayes* dan evaluasi mendapatkan nilai akurasi sebesar 89%, presisi 83%, dan recall sebesar 87% [4][7].

Penelitian sebelumnya telah dilakukan oleh Topan Dedi Ramadhan, Deden Wahidin, Elsa Elvira Awal yang berjudul klasifikasi sentimen terhadap pinjaman *online* (pinjol) menggunakan algoritma *Naive Bayes* menghasilkan Analisis sentimen dengan mengklasifikasi algoritma *Naive Bayes* dengan data set sebanyak 650 data, 80% sebagai data training dan 20% digunakan sebagai data testing dengan hasil akurasi sebesar 75% [8].

Oleh sebab itu, penelitian ini dilakukan untuk mengetahui hasil analisis sentimen pada aplikasi AdaKami dengan memanfaatkan ulasan dari aplikasi *google playstore* untuk mengetahui hasil sentimen positif dan negatif dari ulasan *google playstore*. dan mengetahui tingkat akurasi menggunakan algoritma *Naive Bayes* agar aplikasi AdaKami dapat meningkatkan kemajuan dikemudian hari.

B. Metode Penelitian

adapun flowchart dari penelitian yang dilakukan dapat dilihat pada gambar 1 berikut ini :



Gambar 1. Metodologi penelitian

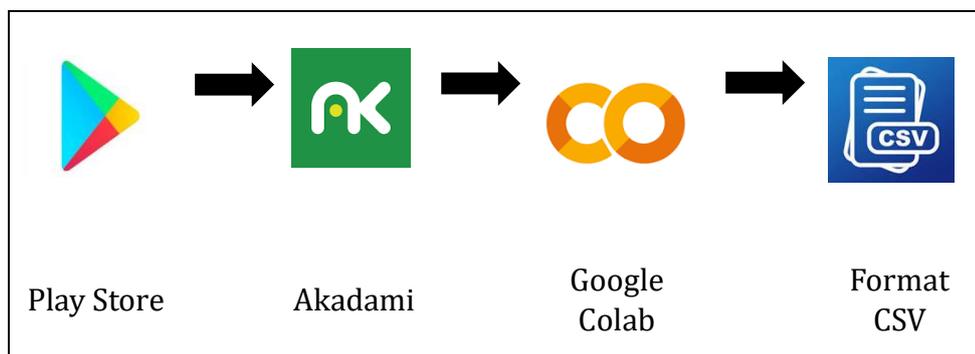
1. Adapun alur dari penelitian ini tahapan pertama dilakukan penarikan data dengan menggunakan *google colab* dibantu dengan library *google_play_scraper*.

2. Tahap selanjutnya *preprocessing* dataset dengan tujuan untuk membersihkan data dan mempermudah mesin dalam memproses data, tahapan yang dilakukan dengan *preprocessing* terdiri dari : *cleaning data*, *transformcases*, *tokenize*, dan *stopword*.
3. Setelah melewati tahap *preprocessing* selanjutnya dilakukan tahap pelabelan secara manual.
4. Implementasi *Naive Bayes* dikukan dengan cara pembagian data menjadi 2 kelas yaitu negatif dan positif berdasarkan hasil dari ulasan pengguna aplikasi AdaKami.
5. Tahapan terakhir dari penelitian ini yaitu evaluasi untuk mengetahui nilai hasil *accuracy*, *precision*, dan *recall*. Bertujuan menghasilkan performa algoritma *Naive Bayes* dalam melakukan analisis sentimen dibantu dengan menerapkan *confusion matrix*.

C. Hasil dan Pembahasan

1. Penarikan Data

Peneliti melakukan penarikan data melalui ulasan aplikasi AdaKami di *google playstore* menggunakan bantuan *google collab* dan dilanjutkan dengan aplikasi *rapidminer*. Teknik penarikan data dilakukan dengan menggunakan *web scraping*. *Web Scraping* adalah teknik untuk mengambil informasi dari *World Wide Web* (WWW) dan menyimpannya dalam format file atau database untuk penggunaan atau analisis lebih lanjut. Terkadang juga disebut sebagai ekstraksi *web* atau permanenan *web* [9]. Setelah mendapatkan data yang dihasilkan menggunakan bantuan *google collab* data lalu disimpan dalam bentuk file *csv* yang akan dijadikan dataset untuk penelitian ini.



Gambar 2. Proses penarikan data

Gambar 2 merupakan tahapan peneliti dalam melakukan pengambilan data yang di dapat pada ulasan playstore, peneliti melakukan kurang lebih selama 4 bulan dimulai dari oktober 2023 sampai januari 2024. Gambar 3 menunjukkan hasil scrapping data melalui bantuan *google colab*, peneliti melakukan *scraping* data ulasan pengguna aplikasi AdaKami sebanyak 1000 data ulasan, setelah dilakukan *scraping* data peneliti mendapatkan data sebanyak 998 data. Setelah proses

scrapping selesai, kemudia file *csv* akan di download dan dilanjutkan ke tahap proses *scrapping*.

```

from google_play_scraper import Sort, reviews

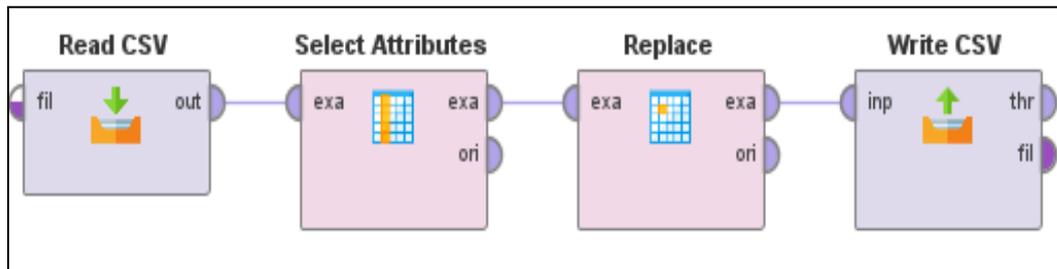
result, continuation_token = reviews(
    'com.adakami.dana.kredit.pinjaman&hl=',
    lang='id',
    country='id',
    sort=Sort.MOST_RELEVANT,
    count=1000,
    filter_score_with=None

```

Gambar 3. Proses Scrapping Data ulasan AdaKami

2. Preprocessing Data

Preprocessing adalah tahapan awal dalam mempersiapkan data agar data menjadi lebih terstruktur, *preprocessing* melakukan seleksi data yang akan di proses pada setiap dokumen [10]. Proses *preprocessing* atau pembersihan merupakan langkah dalam suatu penelitian yang bertujuan untuk membersihkan data sebelum dianalisis lebih lanjut. *Preprocessing* penelitian ini sangat membantu dalam meningkatkan kualitas data, pada tahap *preprocessing* terdapat beberapa tahapan diantaranya *cleansing*, *tokenizing*, *transform cases*, dan *filter stopwords*. *Tools* yang digunakan dalam *preprocessing* yaitu *rapidminer*.



Gambar 4. Proses *cleansing*

Gambar 4 merupakan suatu proses *cleansing* yang digunakan dalam pembersihan dataset untuk menghilangkan karakter seperti tanda baca, karakter asing dan lain sebagainya.

Tabel 1. Hasil *cleansing*

Sebelum	Sesudah
---------	---------

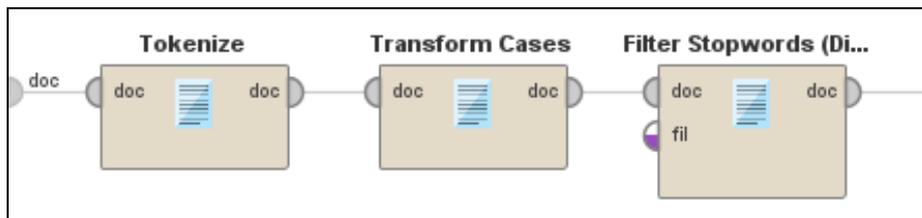
Fast respon sangat membantu sekali disaat butuh dana cepat tks AdaKami '??'??

Fast respon sangat membantu sekali disaat butuh dana cepat tks AdaKami

Amazing baru pertama sudah di kasih 26jt lebih thank you so much AdaKami??»'??' cuman bunganya tinggi~>>>

Amazing baru pertama sudah di kasih jt lebih thank you so much AdaKami cuman bunganya tinggi

Setelah mendapatkan hasil cleaning, dataset masih perlu melanjutkan tahapan-tahapan berikutnya seperti : *tokenize, transform case, dan filter stopwords*.



Gambar 5. Operator preprocessing

Gambar 5 merupakan tahapan kedua dalam proses preprocessing yaitu *tokenizing*, tahapan ini dilakukan dengan cara memecah kalimat menjadi beberapa kata.

Tabel 2. Hasil Tokenize

Sebelum	Sesudah
Sangat membantu tapi sayang mau pinjam lagi dana mendadak ngga bisa limit terkunci	Sangat, membantu, tapi, sayang, mau, pinjam, lagi, dana, mendadak, ngga, bisa, limit, terkunci
Proses mudahcepat dan limit besar sangat membantu	Proses, mudah, cepat, dan, limit, besar, sangat, membantu

Pada tahap ketiga dilakukan proses *transformcases*. Berfungsi untuk pemrosesan teks mengubah kalimat yang sebelumnya kalimat terdapat *uppercase* akan dirubah menjadi *lowercase* atau perubahan pada kalimat dimana huruf kapital yang ada pada dataset menjadi huruf kecil.

Tabel 3. Hasil transformcase

Sebelum	Sesudah
---------	---------

Sangat, membantu, tapi, sayang, mau, pinjam, lagi, dana, mendadak, ngga, bisa, limit, terkunci	sangat, membantu, tapi, sayang, mau, pinjam, lagi, dana, mendadak, ngga, bisa, limit, terkunci
Proses, mudah, cepat, dan, limit, besar, sangat, membantu	proses, mudah, cepat, dan, limit, besar, sangat, membantu

Tahapan terakhir pada proses *preprocessing* adalah *filter stopwords*. Tahapan ini digunakan untuk menghapus kata yang tidak penting atau kata yang sering muncul namun tidak memiliki nilai dalam sentimen. Penghapusan kata dalam proses ini dilakukan dengan menggunakan kamus yang terdapat pada *www.Kagel.com* dengan menggunakan bahasa Indonesia.

Tabel 4. Hasil *Filter Stopwords*

Sebelum	Sesudah
Sangat, membantu, tapi, sayang, mau, pinjam, lagi, dana, mendadak, ngga, bisa, limit, terkunci	membantu, sayang, pinjam, dana, mendadak, limit, terkunci
Proses, mudah, cepat, dan, limit, besar, sangat, membantu	Proses, mudah, cepat, limit, besar, membantu

3. Labeling

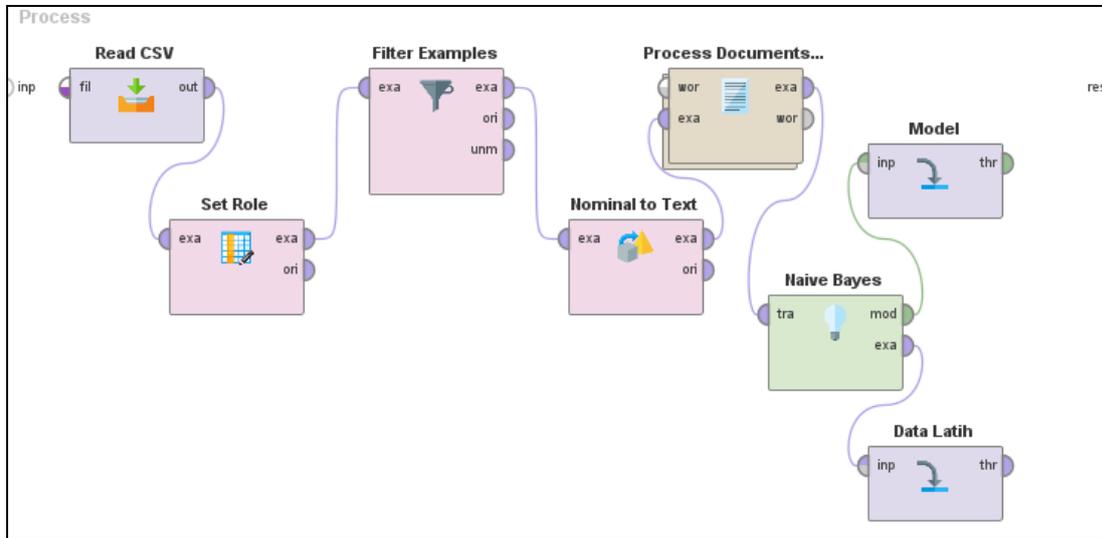
Setelah melewati tahap *preprocessing*, tahapan yang dilakukan yaitu proses *labelling* atau pelabelan. Pada tahap pelabelan ini peneliti melakukan tahap pelabelan secara manual dengan membagi dua data yaitu data latih dan data uji. Dari hasil data yang telah di *scraping* sebanyak 1000 data, kemudian data berubah sebanyak 998 data ulasan pengguna yang sudah melewati tahap *preprocessing* selanjutnya diambil sebanyak 700 data sebagai data latih dan 298 sebagai data uji. Sebanyak 700 ulasan dilabeli menjadi dua yaitu positif dan negatif secara manual menurut ulasan yang diberikan oleh pengguna aplikasi AdaKami Sementara itu sebanyak 298 ulasan yang tidak dilabeli menjadi data uji dalam klasifikasi menggunakan algoritma *Naive Bayes*

Tabel 5. Hasil *Labelling*

Sentimen	Ulasan
Positif	Sukses
Negatif	Gagal
Positif	Rekomendet
Negatif	Ribet

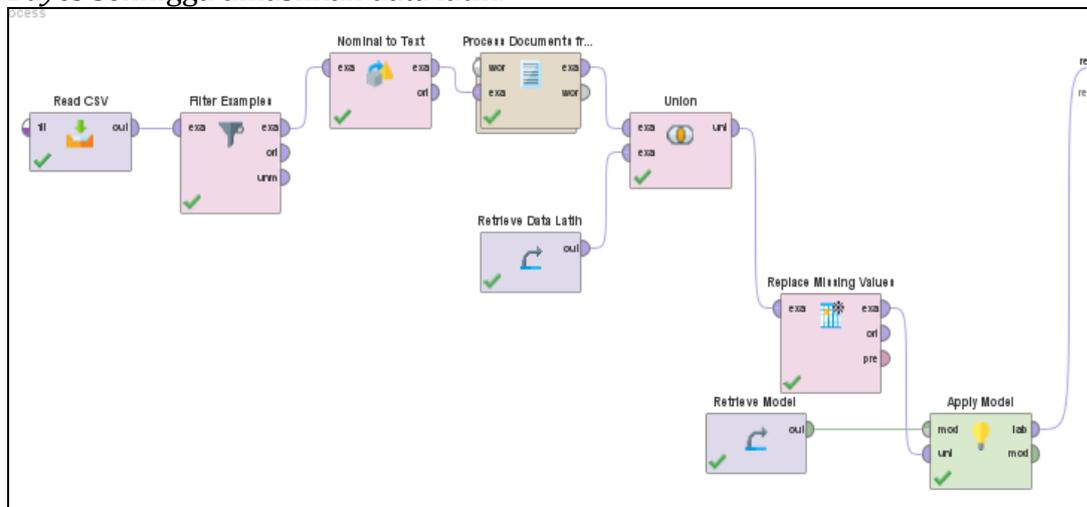
4. Implementasi *Naive Bayes*

Sebelum dilakukannya implementasi *Naive Bayes* dataset dilakukan pembagian data *training* dan *testing*. Dataset yang telah berhasil di *scraping* sebanyak 1000 data, proses *preprocessing* menghasilkan dataset sebanyak 998. Dataset dibagi menjadi 2 sebanyak 700 dijadikan data *training* dan 298 dijadikan data *testing*. Selanjutnya masuk pada tahapan impementasi algoritma *Naive Bayes* sebagai model.



Gambar 6. Proses pembuatan data latih

Pada gambar 6 merupakan implementasi *Naive Bayes* sebagai model. Tahapan pertama dalam implementasi *Naive Bayes* dengan menggunakan *read CSV* untuk membaca dataset. *Filter example* untuk memilah data yang belum diberikan label. Selajutnya, *nominal to text* untuk mengubah data berbentuk nominal menjadi bentuk teks. Proses document berisi *tokenize*, *transform case* dan *filter stopwords*. Kemudian data melakukan proses pembelajaran pada mesin menggunakan *Naive Bayes* sehingga dihasilkan data latih.



Gambar 7. Implementasi *Naive Bayes*

setelah didapatkan data latih, selanjutnya dilakukan mendapatkan hasil prediksi dari model. Selanjutnya data masuk pada operator *filter example* digunakan untuk menyaring data yang belum dilabeli untuk *nominal text* dan

jumlah kemuculan terbanyak. Dan kata ‘bintang’ mendapatkan jumlah kemunculan kata yang paling sedikit.

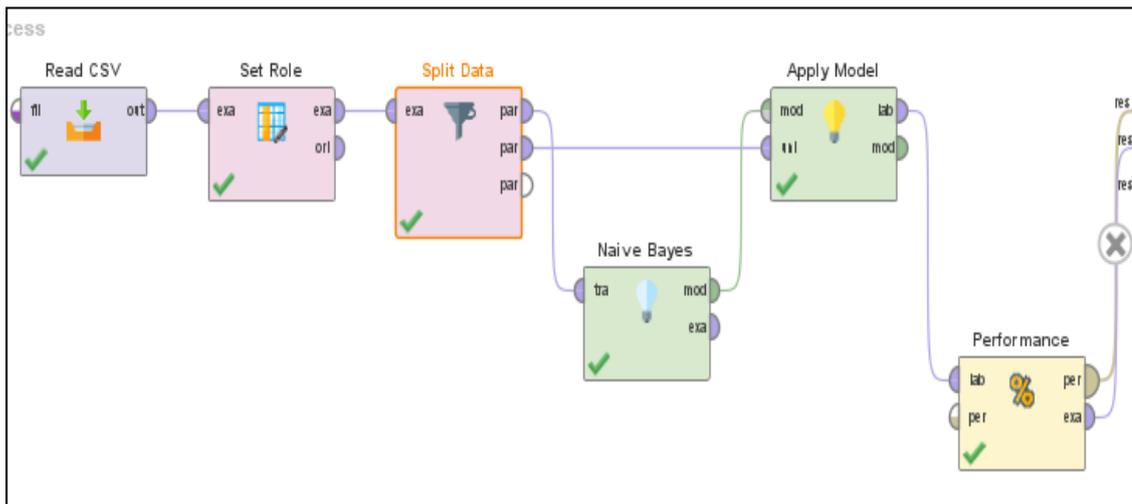
5. Evaluasi

Proses akhir dari tahap penelitian ini terdapat poin evaluasi. Dimana evaluasi mendapatkan hasil dari model *Naive Bayes*. Sehingga pada tahap pengklasifikasian data dapat memperoleh nilai tingkat akurasi dari algoritma *Naive Bayes*. Di dalam penelitian ini menerapkan persamaan *confusion matrix*, dimana fungsi dari *confusion matrix* sendiri yaitu untuk mengetahui *accuracy*, *precision*, dan *recall* [11].

Tabel 6. *Confusion matrix*

		Predicted Class	
		Positive	Negative
Actual Class	Positive	TP	FP
	Negative	FN	TN

Istilah yang digunakan di dalam *confusion matrix* terdapat *True Positive* (TP) dimana *actual* sama *predicted class*-nya positif. *False Positive* (FP) dimana *actual class*-nya negatif dan *predicted class*-nya positif. *False Negative* (FN) dimana *actual class*-nya positif dan *predicted class*-nya negatif. Dan *True Negative* (TN) dimana *actual* dan *predicted class*-nya negatif. Pada gambar10 terdapat operator-operator untuk mengimplementasikan dalam perhitungan *confusion matrix* pada *rapid miner*.



Gambar 10. Proses *confusion matrix*

Proses dari perhitungan *confusion matrix* dengan membaca *read CSV* untuk menambahkan file yang akan diproses. *set role* digunakan untuk membaca kolom prediksi yang digunakan sebagai label. *Split data* digunakan untuk membagi data 70 % : 30% dan dihubungkan pada model *Naive Bayes* dan *apply model*. Selanjutnya data disambungkan dengan *performance* untuk mengetahui presentase nilai akurasi sebesar 78.93% seperti yang ditunjukkan pada gambar dibawah ini

accuracy: 78.93%			
	true Positif	true Negatif	class precision
pred. Positif	104	28	78.79%
pred. Negatif	35	132	79.04%
class recall	74.82%	82.50%	

Gambar 11. Hasil accurracy klasifikasi

Dengan menerapkan persamaan *confusion matrix* peneliti mendapatkan hasil perhitungan *accuracy*, *precision*, dan *recall* dari hasil klasifikasi diatas adalah sebagai berikut.

$$Accuracy = \frac{TP+FN}{TP+TN+FP+FN} = \frac{104+132}{104+132+28+35} = \frac{236}{299} = 0,7892 = 78,92\%$$

$$Precision = \frac{TP}{TP+FP} = \frac{104}{104+28} = \frac{104}{132} = 0,7879 = 78,79\%$$

$$Recall = \frac{TP}{TP+FN} = \frac{104}{104+35} = \frac{104}{139} = 0,7482 = 74,82\%$$

Berdasarkan hasil perhitungan diatas peneliti dapat menentukan hasil klasifikasi pada ulasan aplikasi AdaKami di *playstore* dengan *score accuracy* sebesar 78,92%, *precision* 78,79%, dan *recall* 74,82%.

D. Simpulan

Hasil penelitian yang telah dilakukan diatas dengan menggunakan data dari ulasan aplikasi AdaKami di *google playstore* menggunakan algoritma *Naive Bayes* mendapatkan 998 data yang digunakan untuk penelitian. Dengan membagi data sebanyak 700 data sebagai data latih dan 298 data sebagai data uji. Dengan jumlah 998 data peneliti mendapatkan hasil sebanyak 451 ulasan positif dan 547 ulasan negatif. Maka dari itu dapat disimpulkan penelitian ini yang menggunakan metode *Naive Bayes* dan *confusion matrix* di tahap evaluasi mendapatkan hasil nilai *accuracy* sebesar 78,92%, *precision* sebesar 78,79%, dan *recall* sebesar 74,82%. Penelitian diatas memprediksi bahwa berdasarkan hasil dari klasifikasi yang telah dilakukan diatas memiliki skor akurasi yang cukup tinggi. Dapat disimpulkan bahwa penelitian ini melakukan analisis sentimen ulasan pengguna aplikasi AdaKami mendapati hasil negatif lebih dominan. Penelitian ini menghasilkan bahwa pengguna aplikasi AdaKami berisi tentang ketidakpuasan terhadap pengguna AdaKami.

E. Referensi

- [1] Sherty Veronika and A. Y. Mafruhah, "Pengaruh Pertumbuhan Ekonomi, Investasi dan Inflasi terhadap Pengangguran Terdidik di Provinsi Jawa Barat," *J Ris Ilmu Ekon dan Bisnis*, pp. 139-146, 2022, doi: 10.29313/jrieb.vi.1290.

- [2] Nabilla Nurdina and Ima Amailah, "Preferensi Masyarakat Kota Bandung Terhadap Pinjaman Online," *J Ris Ilmu Ekon dan Bisnis*, pp. 33–38, 2023, doi: 10.29313/jriieb.v3i1.1908.
- [3] U. Z. Faridhun, "Review of Sharia Economic Law on Billing Online Loans: Adakami Top Fintech Study," *Int Conf Islam Stud*, no. March, 2022.
- [4] N. W. A. POHAN, I. BUDI, and R. R. SURYONO, "Borrower Sentiment on P2P Lending in Indonesia Based on Google Playstore Reviews," vol. 172, no. Siconian 2019, pp. 17–23, 2020, doi: 10.2991/aisr.k.200424.003.
- [5] R. R. Putra, M. E. Johan, and E. R. Kaburuan, "A naïve bayes sentiment analysis for fintech mobile application user review in Indonesia," *Int J Adv Trends Comput Sci Eng*, vol. 8, no. 5, pp. 1856–1860, 2019, doi: 10.30534/ijatcse/2019/07852019.
- [6] R. Wahyudi and G. Kusumawardana, "Analisis Sentimen pada Aplikasi Grab di Google Play Store Menggunakan Support Vector Machine," *J Inform*, vol. 8, no. 2, pp. 200–207, 2021, doi: 10.31294/ji.v8i2.9681.
- [7] A. I. Tanggraeni and M. N. N. Sitokdana, "Analisis Sentimen Aplikasi E-Government pada Google Play Menggunakan Algoritma Naïve Bayes," *JATISI (Jurnal Tek Inform dan Sist Informasi)*, vol. 9, no. 2, pp. 785–795, 2022, doi: 10.35957/jatisi.v9i2.1835.
- [8] T. D. Ramadhan, D. Wahiddin, and E. E. Awal, "Klasifikasi Sentimen Terhadap Pinjaman Online (Pinjol) Menggunakan Algoritma Naive Bayes," *Sci Student J Information, Technol Sci*, vol. IV, no. 1, pp. 82–87, 2023, [Online]. Available: www.tripadvisor.com
- [9] Y. A. Hafiz and E. Sudarmilah, "Implementasi Web Scraping Pada Portal Berita Online Pendahuluan Informasi semakin pesat dan penting seiring dengan semakin canggihnya era digital (Fathir et al ., 2023). Portal berita online adalah salah satu sumber informasi yang dapat memberikan manfa," vol. 12, no. 1, pp. 55–60, 2023.
- [10] L. Hermawan and M. Bellaniar Ismiati, "Pembelajaran Text Preprocessing berbasis Simulator Untuk Mata Kuliah Information Retrieval," *J Transform*, vol. 17, no. 2, p. 188, 2020, doi: 10.26623/transformatika.v17i2.1705.
- [11] Ernianti Hasibuan and Elmo Allistair Heriyanto, "Analisis Sentimen Pada Ulasan Aplikasi Amazon Shopping Di Google Play Store Menggunakan Naive Bayes Classifier," *J Tek dan Sci*, vol. 1, no. 3, pp. 13–24, 2022, doi: 10.56127/jts.v1i3.434.