ANALISIS SENTIMEN PENGGUNA APLIKASI KARIER.MU MENGGUNAKAN ALGORITMA BERNOULLI NAÏVE BAYES CLASSIFIER DENGAN METODE N-GRAM

Tasya Nurpadilah, Ultach Enri, E. Haodudin Nurkifli
Universitas Singaperbangsa Karawang; Jl. HS. Ronggo Waluyo, Puseurjaya, Telukjambe Timur, Karawang, Jawa Barat 41361; Telp. (0267) 641177
2010631170122@student.unsika.ac.id

ABSTRAK

Kekurangan keterampilan di pasar kerja global telah menjadi isu mendesak bagi pengusaha, terutama terkait keterampilan teknis dan soft skills. Aplikasi mobile seperti Karier.mu menawarkan solusi dengan menyediakan platform bagi pencari kerja untuk mengembangkan keterampilan yang relevan. Aplikasi ini telah mendapatkan lebih dari 500.000 pengguna dan berperan penting dalam program Kartu Prakerja untuk meningkatkan daya saing tenaga kerja Indonesia. Meskipun Karier.mu diterima dengan baik, masih terdapat tantangan dalam menganalisis umpan balik pengguna secara mendalam. Penelitian ini bertujuan untuk menganalisis sentimen pengguna aplikasi Karier.mu menggunakan algoritma Bernoulli Naïve Bayes dengan metode n-gram. Data dikumpulkan melalui ulasan pengguna dan diproses melalui tahapan teks preprocessing, pelabelan sentiment dengan pendekatan lexicon, dan ekstraksi fitur menggunakan TFIDF N-Gram. Metode SMOTE diterapkan untuk menangani ketidakseimbangan kelas dalam data. Hasil penelitian menunjukkan bahwa model mampu mengklasifikasikan sentimen dengan akurasi 96%. Hasil dari analisis ini diharapkan memberikan wawasan bagi pengembang aplikasi dalam meningkatkan kualitas layanan dan pengalaman pengguna, serta membantu perusahaan beradaptasi dengan kebutuhan pasar kerja yang kompetitif.

Kata kunci : *Analisis Sentimen, Karier.mu; Lexicon; Bernoulli Naïve Bayes, N-Gram; SMOTE.*

1. PENDAHULUAN

Kekurangan keterampilan di pasar kerja global telah menjadi isu yang semakin mendesak, sebagaimana diungkapkan dalam laporan "Talent Shortage Survey 2020" oleh ManpowerGroup. Laporan tersebut menunjukkan bahwa banyak pengusaha menghadapi tantangan dalam menemukan tenaga kerja yang memiliki keterampilan teknis, seperti informasi teknologi (IT), teknik, dan ilmu pengetahuan, serta keterampilan soft seperti kreativitas dan komunikasi. Dalam beberapa tahun terakhir, industri dan organisasi telah mengalami perubahan teknologi, komersial, dan budaya yang pesat. Oleh karena itu, para pencari kerja harus mampu beradaptasi dan terus meningkatkan keterampilannya. Namun banyak pencari kerja yang tidak memiliki kemampuan tersebut sehingga membuat mereka kurang mampu bersaing di pasar kerja [1].

Dalam era digital saat ini, aplikasi mobile telah menjadi alat yang tak terpisahkan dalam pencarian dan pengembangan karir Aplikasi mobile seperti *Karier.mu* muncul sebagai solusi inovatif yang tidak hanya membantu pencari kerja menemukan pekerjaan, tetapi juga mengembangkan keterampilan yang relevan dengan kebutuhan pasar. *Karier.mu* menawarkan kursus di berbagai bidang, didukung oleh Program Kartu Prakerja yang bertujuan untuk meningkatkan daya saing tenaga kerja Indonesia. Meskipun aplikasi ini telah diterima dengan baik, masih terdapat tantangan dalam menganalisis lebih dari 60.000 ulasan pengguna. Analisis sentimen menjadi penting karena sekitar 50% pengguna internet mengandalkan ulasan sebelum memutuskan menggunakan suatu produk atau layanan. Analisis sistematis terhadap ulasan pengguna *Karier.mu* dapat memberikan wawasan berharga bagi pengembang untuk mengatasi kekurangan dan meningkatkan kualitas layanan. Penelitian sebelumnya oleh Taslim et al. (2023) menunjukkan bahwa algoritma Bernoulli Naive Bayes (BNB) konsisten dengan akurasi 80%, presisi 88%, dan recall 68%, lebih unggul dibandingkan Gaussian dan Multinomial Naive Bayes. Sementara itu, penelitian oleh Nur Arifin et al. (2021) menyimpulkan bahwa kombinasi unigram dan bigram dalam algoritma SVM memberikan akurasi 70%, lebih baik dibandingkan penggunaan unigram atau bigram secara terpisah yang hanya mencapai 60% [2].

Berdasarkan latar belakang tersebut, maka analisis sentimen perlu dilakukan. Penelitian ini berfokus pada analisis ulasan pengguna aplikasi Karier.mu pada layanan Google Play dengan menggunakan algoritma *Bernoulli Naïve Bayes Classifier.*

Penelitian ini menggunakan algoritma *Bernoulli Naïve Bayes Classifier* dengan metode N-Gram yaitu unigram dan bigram karena algoritma ini dikenal sebagai metode sederhana dan cepat, yang memiliki performa yang tinggi dalam klasifikasi teks, sehingga cocok untuk analisis sentimen dan dinilai memiliki akurasi klasifikasi yang tinggi untuk analisis sentimen serta pelabelan data yang menggunakan Kamus *Lexicon* serta untuk mengantisipasi ketidakseimbangan data maka diterapkan Teknik SMOTE untuk penyeimbang kelas sentimen.

2. TINJAUAN PUSTAKA

2.1. Analisis Sentimen

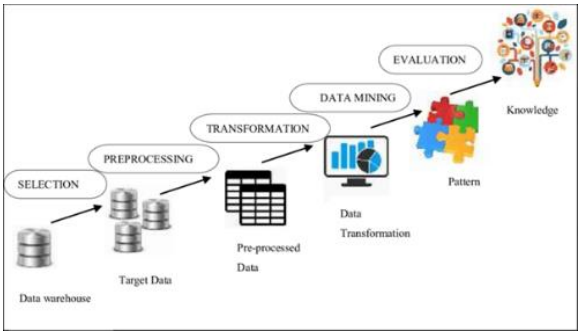
Analisis sentimen atau biasa disebut dengan opinion mining adalah bidang penelitian yang terus berkembang seiring dengan kemajuan teknologi media sosial, terutama aplikasi yang ada pada google play store. Hal ini karena google play store menyediakan berbagai macam aplikasi dan memberikan ruang untuk pengguna dalam memberikan ulasan sebuah aplikasi. Analisis sentimen merupakan proses komputasional untuk mengekstrak, memahami dan mengolah opini atau sentimen dari teks, secara otomatis untuk mendapatkan informasi sentimen yang terdapat pada kalimat opini [3].

2.2. Karier.mu

Karier.mu adalah platform teknologi untuk pengembangan dan peningkatan karir oleh Sekolah.mu, berkolaborasi untuk menyediakan program pelatihan bagi semua orang. Didirikan oleh Najelaa Shihab dan tim pada tahun 2019, tujuan utama karier.mu adalah menghubungkan orang-orang dengan mentor terbaik, membantu mereka mewujudkan mimpi dan karier impian di masa depan. Karier.mu menawarkan berbagai program dengan pelayanan berkualitas yang bermanfaat bagi semua orang, serta akses pelatihan digital yang terintegrasi dengan materi yang dirancang oleh para ahli dari berbagai bidang.

2.3. Knowledge Discovery in Database (KDD

Knowledge Discovery in Database adalah metode pengumpulan pengetahuan database yang telah ada. Suatu database mempunyai tabel-tabel yang saling berhubungan. Hasil yang diperoleh dapat digunakan sebagai dasar pengetahuan untuk pengambilan keputusan. Istilah *Knowledge Discovery in Database* (KDD) dalam data mining sering digunakan secara bergantian untuk menggambarkan proses penggalian informasi tersembunyi dalam database berukuran besar. Proses KDD memiliki beberapa tahapan diantaranya: Data selection, Pre-processing, Transformation.



Gambar 1. Proses tahapan KDD

2.4. Lexicon

Lexicon Based adalah pendekatan yang menggunakan kamus sentimen yang berisi kata-kata positif dan negatif yang mencocokkan kata-kata dalam sebuah kalimat untuk menentukan derajat polaritasnya. Klasifikasi sentimen *lexicon* merupakan klasifikasi berdasarkan kata positif dan kata negatif pada ulasan aplikasi yang telah dibersihkan [5].

Kelebihan dari metode *lexicon* adalah data berupa kata-kata dalam kalimat akan langsung dibandingkan dengan kamus kata-kata pendapat yang terdapat dalam kosakata. Jika suatu kalimat mengandung kata-kata yang berbentuk opini, maka kalimat tersebut termasuk kalimat opini.

2.5. N-Gram

N-Gram adalah fitur yang digunakan pada tahap pengolahan teks pada proses klasifikasi sentiment. N-Gram digunakan untuk menggabungkan kata sifat yang sering digunakan atau memberi bobot pada kata-kata penting dalam dokumen untuk mengekspresikan emosi atau perasaan [6].

2.6. Naïve Bayes Classifier

Algoritma Naïve Bayes Classifier termasuk kepada metode klasifikasi berdasarkan teorema bayes karena menggunakan Teknik *probabilistic* dan *statistic* yang dikemukakan oleh Thomas Bayes sebagai seorang ilmuwan asal Inggris, fungsi dari teori tersebut yaittu untuk memprediksi peluang masa depan berdasarkan apa yang terjadi di masa lalu [7]. Algoritma *Naïve Bayes Classifier* terbagi menjadi 3 yaitu, *Bernoulli Naïve Bayes*, *Multinomial Naïve Bayes, dan Gaussian Naïve Bayes* [8]. Metode yang akan digunakan pada penelitian ini yaitu *Bernoulli Naïve Bayes Classifier*, yang artinya metode ini bekerja pada konsep biner bahwa apakah istilah itu terjadi pada dokumen tersebut atau tidak.

2.7. SMOTE

SMOTE adalah teknik pengambilan sampel ulang data yang banyak digunakan untuk mengatasi masalah overfitting pada oversample dan hilangnya informasi karena undersambling. Ketidakseimbangan kelas terjadi ketika jumlah sampel pada suatu kelas tertentu lebih banyak atau lebih sedikit dibandingkan dengan kelas lainnya.SMOTE telah menjadi metode yang digunakan untuk mempelajari data yang tidak seimbang sejak tahun 2002, dengan menggunakan sampel sintetis dari kelompok minoritas dan sumber lainnya [9].

Teknik SMOTE digunakan untuk mengatasi masalah ketimpangan kelas. Teknik ini mensitesis sampel baru dari kelas minoritas untuk menyeimbangkan kumpulan data dengan membuat instance kelas minoritas baru dengan membentuk kombinasi cembung dari sampel tetangga. Kemudian, tingkat pengambilan sampel terlebih yang diperlukan dipilih secara acak. Contoh minoritas baru ini ditambahkan ke data pelatihan dan pengklasifikasi dilatih dengan data tambahan. Teknik SMOTE secara umum lebih akurat dibandingkan metode sampling konvensional pada metode SMOTE [11].

2.8. Confusion Matrix

Confusion matriks adalah alat yang digunakan untuk mengevaluasi sistem klasifikasi yang menghasilkan tabel matriks untuk membandingkan hasil prediksi model dengan kelas data sebenarnya [12].

Dibawah ini merupakan contoh dari perhitungan akurasi dengan *Confusion Matrix* menggunaka dua kelas yaitu positif dan negative seperti yang terlihat pada Tabel 1.

Tabel 1. Confusion Matrix

Classification	***Prediction Class***	
	Class - Yes	***Class - No***
Class - Yes	A (*True Positive*)/TP	B (*False Negatif*)/FN
Class - No	C (*False Positive*)/FP	D (*True Negatif*)/TN

Keterangan :

a. True Positive (TP) = kondisi data yang diprediksi positif dan nilai sebenarnya positif.
b. False Positive (FP) = kondisi data yang diprediksi positif dan nilai sebenarnya negatif.
c. False Negative (FN) = kondisi data yang diprediksi negatif dan nilai sebenarnya positif.
d. True Negative (TN) = kondisi data yang diprediksi negatif dan nilai sebenarnya negatif.

Berikut merupakan rumus untuk menghitung accuracy, precision, recall, f1- score.

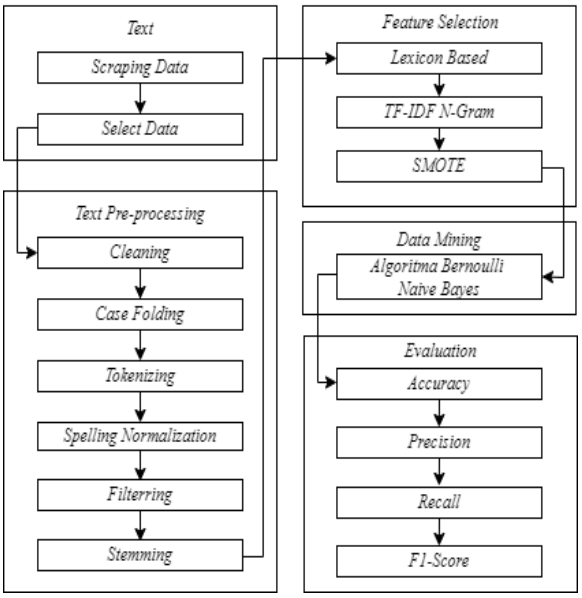
$$Accuracy = \frac{TP+TN}{TP+FN+TP+TN} \quad (1)$$

$$Precision = \frac{TP}{TP+FP} \quad (2)$$

$$Recall = \frac{TP}{TP+FN} \quad (3)$$

$$F1\text{-}Score = 2x\frac{precision\ x\ recall}{precision+\ recall} \quad (4)$$

3. METODE PENELITIAN



Gambar 2. Metodologi Penelitian

Dalam penelitian ini, rancangan penelitian disesuaikan dengan metodologi yang digunakan yaitu *Knowledge Discovery in Database* (KDD). Penelitian ini menggunakan 5 tahapan garis besar proses penelitian dengan siklus pada gambar 2

3.1. Text

Tahap *text* bertujuan untuk mengakuisisi data objek penelitian dengan teknik tertentu. Setelah *dataset* didapatkan, kemudian dilakukan seleksi untuk mendukung proses pada tahap selanjutnya. Pada tahap *text* ini terbagi menjadi dua bagian, yaitu *Scraping Data* dan *Data Selection*.

a. Scraping Data
   Pada tahapan ini, dilakukan akuisisi data atau pengumpulan data. Data yang digunakan pada penelitian ini berupa data ulasan pengguna aplikasi Karier.mu yang ditulis pada layanan Google Play pada periode Juli 2021 hingga Februari 2024. Data didapatkan dengan melalui proses *scraping* menggunakan *tools Instant Data Scraper*.
b. Data Selection
   Tahap ini merupakan proses proses seleksi data seperti kolom, baris, dan atribut untuk mengurangi kata yang kurang relevan, serta untuk mempermudah proses klasifikasi data.

3.2. Text Pre-processing

Tahap *text pre-processing* dilakukan untuk mengubah data teks yang tidak terstruktur atau acak menjadi data yang terstruktur.

a. Cleaning
   Proses yang dilakukan pada tahap *cleaning* pada penelitian ini adalah membersihkan data teks untuk karakter hastag, menghilangkan angka, menambah spasi setelah koma, menghilangkan tanda baca, meghilangkan karakter yang memanjang pada Kumpulan data hingga menghilangkan spasi tambahan yang ada didalamnya karena atribut-atribut tersebut dianggap tidak pantas.
b. Case Folding
   Yaitu proses mengubah karakter dalam dokumen menjadi huruf kecil.
c. Tokenizing
   Yaitu proses memotong setiap kata pada dokumen teks dengan spasi atau tanda baca untuk membuat penggalan kata atau token yang akan menjadi entitas berharga yang berguna untuk langkah selanjutnya.
d. Normalization
   Pada proses Spelling Normalization, dokumen teks diperbaiki dengan menormalkan kata-kata yang tidak sesuai ejaan menggunakan database normalisasi kata sehingga dokumen diperbaiki sesuai dengan perbaikan ejaan
e. Filtering
   Proses *Filtering* yang dilakukan pada penelitian ini meliputi menghilangkan kata-kata yang tidak penting dari proses klasifikasi. Pada tahap ini diterapkan metode stop-word, yaitu metode

menghilangkan kata-kata non-deskriptif yang sering muncul dalam dokumen hingga tersisa kata-kata penting dan bermakna.

f. Stemming
Yaitu tahap mengganti kata menjadi kata dasar dengan aturan tertentu untuk meringankan database.

3.3. Text Transformation

Setelah tahap pre-processing selesai, kemudian data dilabeli dengan Menggunakan pendekatan lexicon untuk menentukan sentiment dari setiap teks. Kelas sentimen yang digunakan yaitu sentimen positif dan negatif. Langkah berikutnya pada tahap Transformation dilakukan seleksi fitur menggunakan TF-IDF (Term Fre Inverse Dokument Frekuensi). Jika hasil pada proses pelabelan data menghasilkan kelas yang tidak seimbang, maka dilakukannya penerapan SMOTE untuk mengatasi hal tetsebut dengan cara menambah sampel sintetis pada kelas minoritas guna meningkatkan performa model dala mengenali pola dari data minoritas. Teknik pemilihan fitur dengan TF-IDF dicapai dengan mempertimbangkan kemunculan kata-kata dalam dokumen.

Proses ini juga bertujuan untuk mengurangi jumlah variable dan data yang sebenarnya tidak diperlukan. Ekstrasi TF-IDF depat digunakan untuk mencari nilai kemunculan term suatu kata dalam sebuah dokumen. Penerapan N-Gram pada ekstrasi TF-IDF membuat algoritma perhitungan TF-IDF berdasarkan kriteria menjadi lebih efisien. Selain itu, penggunaan N-Gram untuk mengeksrak fitur TF-IDF menjadi akurat dan menghindari kesalahan komputasi. Dalam pencarian ini, menghasilkan nilai TF-IDF untuk semua dokumen [10].

3.4. Data Mining

Pada penelitian ini digunakan metode klasifikasi data yang telah diolah sebelumnya dengan menggunakan algoritma *Bernoulli Naïve Bayes Classifier* ke dalam dua *class* yaitu sentiment posistif dan sentimen negatif. Pembagian *dataset* diterapkan Teknik *percentage split.* Terdapat 3 skenario yang akan digunakan. Skenario tersebut diantaranya:

a. Skenario pertama : 90% *data training* dan 10% *data testing*
b. Skenario kedua : 80% *data training* dan 20% *data testing*
c. Skenario ketiga : 70% *data training* dan 30% *data testing*

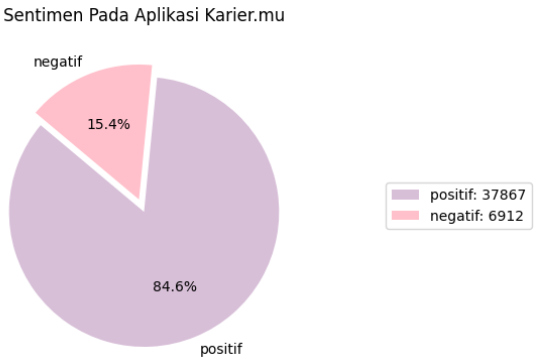
Data yang sudah dibagi menjadi *data training* dan *data testing* selanjutnya akan dicari nilai Accuracy, precision, recall, *F1-Score*, *true positif, true negatif, false positif, false negatif*. Pada proses ini menggunakan Bahasa pemrograman Python.

3.5. Evaluation

Pada tahap ini dilakukan evaluasi untuk megukur validitas hasil klasifikasi dengan menghitung nilai *accuracy*, *precision*, dan *recall*. Parameter yang digunakan untuk evaluasi adalah *accuracy*, *precision*, dan *recall* yang dihitung dari matriks konfusi dengan 4 nilai sebagai acuan dalam perhitungan, yaitu *True Positif rate* (TP *rate*), *True Negatif rate* (TN *rate*), *False Positif rate* (FP *rate*), dan *False Negatif rate* (FN *rate*). Evaluasi dilakukan dengan membandingkan hasil klasifikasi akurasi menggunakan seleksi fitur TF-IDF N-Gram.

4. HASIL DAN PEMBAHASAN

Penelitian ini menggunakan algoritma Benoulli Naïve Bayes Classifier dengan metode n-gram. Metodologi yang digunakan yaitu *Knowledge Discovery in Database (KDD)*, yang memuat diantaranya *Text, Pre-processing, transformation, data mining*, dan *evaluation*. Teknik pengumpulan data yang digunakan adalah *scraping data* menggunakan Google Colaboratory. Pada rentang waktu Juli 2021 hingga Februari 2024, berhasil mengumpulkan total 45.450 data. Setelah proses pre-processing, data menjadi 44.779 data yang terdiri dari 37.867 data positif dan 6.912 data negative.

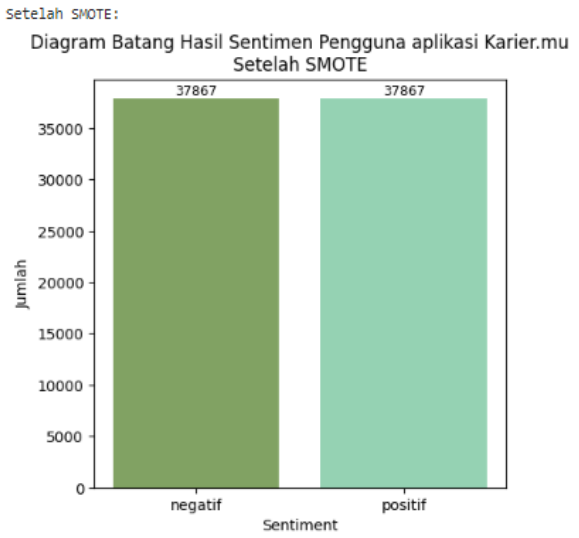


Gambar 3. Hasil Pelabelan Sentimen

Sentimen dengan kamus *lexicon* maka dilakukan pembobotan kata dengan mengimplementasikan dengan dilakukan seleksi fitur menggunakan TF-IDF (*Term Fre Inverse Dokument Frekuensi*). Ekstrasi TF-IDF dapat digunakan untuk mencari nilai kemunculan term suatu kata dalam sebuah dokumen. Penerapan N-Gram pada ekstrasi TF-IDF membuat algoritma perhitungan TF-IDF berdasarkan kriteria menjadi lebih efisien. Pada proses ini menggunakan N-Gram yaitu unigram dan bigram. Fungsi yang digunakan yaitu TfidfVectorizer yang terdapat pada library sklearn. Sebelumnya, dataset akan di bagi menjadi data train dan data test. Data training yaitu data yang digunakan untuk melatih suatu model dalam mempelajari pola atau karakteristik sebuah data, sedangkan data testing yaitu data yang digunakan untuk menguji sebuah model yang sebelumnya telah di latih. Berikut merupakan hasil dari proses pembobotan TF-IDF Ngram yang ditetapkan pada *data train.*

Pada hasil pembobotan sentimen, data sentiment positif dan negative tidak seimbang seperti yang ditujukan pada Gambar 3 diatas, maka dari itu dilakukannya penerapan Teknik SMOTE untuk

menyeimbangkan data dengan hasil seperti diagram dibawah ini.



Gambar 4. Hasil penerapan SMOTE

Setelah proses tersebut selesai, maka dilakukan klasifikasi data yang digunakan dengan mengimplementasikan model dari algoritma Bernoulli naïve bayes dengan menggunakan *library* sklearn.naive_bayes yang terdapat pada pyrhon. Dalam penelitian ini di gunakan proses pengolahan data menggunakan algoritma Bernoulli naïve bayes dengan 3 skenario yang terdiri dari 90:10, 80:20, dan 70:30. Berikut adalah pembagian data training dan data testing.

Tabel 2. Pembagian Data Training dan Data testing

Persentase Data		**Jumlah Data**	
Training	**Testing**	**Training**	**Testing**
90%	10%	68062	4478
80%	20%	60536	8956
70%	30%	52982	13434

Hasil dari penerapan algoritma Bernoulli naïve bayes tersebut, menghasilkan confusion matrix seperti pada Tabel 2 dibawah ini.

Tabel 2. Hasil evaluasi Confusion matrix

Evaluasi	**Rasio 90:10**	**Rasio 80:20**	**Rasio 70:30**
Accuracy	96%	92%	92%
Precision	100%	100%	100%
Recall	96%	91%	91%
F1-Score	98%	95%	95%

Pada penelitian ini skenario dengan rasio 90:10 dengan menunjukkan hasil akurasi sebesar 96%, precision seebesar 100%, recall sebesar 96%, serta f1-score sebesar 98% yang merupakan hasil scenario yang peling tinggi dari scenario lainnya.

5. KESIMPULAN

Analisis sentimen pengguna aplikasi Karier.mu*menggunakan algoritma Bernoulli Naïve Bayes dan metode N-Gram dilakukan melalui pengumpulan data di Google Colab dengan Python, melibatkan 44.779 ulasan yang dikumpulkan dari Juli 2021 hingga Februari 2024. Data ini melalui pelabelan otomatis dengan Insert Lexicon, menghasilkan 37.867 data sentimen positif dan 6.912 data sentimen negatif. Teknik Synthetic Minority Over-sampling Technique (SMOTE) digunakan untuk mengatasi ketidakseimbangan kelas pada data sentimen, dengan tujuan memberikan pemahaman mendalam tentang pandangan dan sikap pengguna terhadap aplikasi Karier.mu.

Temuan menunjukkan bahwa pendekatan ini efektif dalam mengidentifikasi pola opini, yang didominasi oleh sentimen positif, mencerminkan tingkat kepuasan tinggi dari pengguna aplikasi. Evaluasi algoritma Bernoulli Naïve Bayes dilakukan pada tiga skenario pembagian data: 90:10, 80:20, dan 70:30, menggunakan confusion matrix dan SMOTE untuk mengukur kinerja model. Skenario 90:10 menunjukkan hasil terbaik dengan akurasi sebesar 96%, precision 100%, recall 96%, dan F1-score 98%, menunjukkan bahwa metode N-Gram pada algoritma Bernoulli Naïve Bayes efektif dalam analisis sentimen aplikasi ini.

DAFTAR PUSTAKA

[1] Rismawati, D., Hadian, D., Manik, E., & Titi, T. (2021). Pengaruh Kompensasi Dan Motivasi Kerja Terhadap Kinerja Karyawan. *Majalah Bisnis & IPTEK*, *14*(2), 83–93. https://doi.org/10.55208/bistek.v14i2.234

[2] N. Arifin, U. Enri, and N. Sulistiyowati, "Penerapan Algoritma Support Vector Machine (SVM) dengan TF-IDF N-Gram untuk Text Classification," *STRING (Satuan Tulisan Riset dan Inovasi Teknologi)*, vol. 6, no. 2, pp. 123-135, Dec. 2021.

[3] Mutmainah, R. F. (2020). *Aplikasi E-Wallet Dana Pada Situs Google Play Menggunakan Naive Bayes Classifier Analisis Sentimen Ulasan Pengguna Aplikasi E-Wallet Dana Pada Situs Google Play Menggunakan Naive Bayes Classifier*.

[4] W. Gata and Purnomo, "Akurasi Text Mining Menggunakan Algoritma K-Nearest Neighbour pada Data Content Berita SMS," *Jurnal*, vol. 6, no. 1, pp. 1-10, 2017, ISSN: 2089-5615.

[5] Al Khadafi, M., Kurnia Paranitha Kartika, & Filda Febrinita. (2022). Penerapan Metode Naïve Bayes Classifier Dan Lexicon Based Untuk Analisis Sentimen Cyberbullying Pada Bpjs. *JATI (Jurnal Mahasiswa Teknik Informatika)*, *6*(2), 725–733. https://doi.org/10.36040/jati.v6i2.5633

[6] Indarbensyah, P. P. E., & Rochmawati, N. (2021). Penerapan N-Gram menggunakan

Algoritma Random Forest dan Naïve Bayes Classifier pada Analisis Sentimen Kebijakan PPKM 2021. *Journal of Informatics and Computer Science (JINACS)*, *2*(04), 235–244. https://doi.org/10.26740/jinacs.v2n04.p235-244

[7] Watratan, A. F., Puspita, A., & Moeis, D. (2022). Implementasi Algoritma Naive Bayes Untuk Memprediksi Tingkat Penyebaran Covid-19 Di Indonesia. *Jural Riset Rumpun Ilmu Teknik*, *1*(1), 38–46. https://doi.org/10.55606/jurritek.v1i1.127

[8] Singh, M., Bhatt, M. W., Bedi, H. S., & Mishra, U. (2020). *Performance of bernoulli's naive bayes classifier in the detection of fake news*. Elsevier.

[9] Ananda, D., & Suryono, R. R. (2024). Analisis Sentimen Publik Terhadap Pengungsi Rohingya di Indonesia dengan Metode Support Vector Machine dan Naïve Bayes. *Jurnal Media Informatika Budidarma*, *8*(April), 748–757. https://doi.org/10.30865/mib.v8i2.7517

[10] Gifari, O. I., Adha, M., Freddy, F., & Durrand, F. F. S. (2022). Analisis Sentimen Review Film Menggunakan TF-IDF dan Support Vector Machine. *Journal of Information Technology*, *2*(1), 36–40. https://doi.org/10.46229/jifotech.v2i1.330

[11] M. S. Shahab, A. Junaidi, and A. N. Sihananto, "Klasifikasi citra plankton dengan algoritma hibrida convolutional neural network dan extreme learning machine," *JITET (Jurnal Informatika dan Teknik Elektro Terapan)*, vol. 12, no. 3S1, pp. 1-6, 2024. DOI: 10.23960/jitet.v12i3S1.5219.

[12] Z. A. Sriyanti, D. S. Y. Kartika, dan A. R. E. Najaf, "Implementasi model BERT pada analisis sentimen pengguna Twitter terhadap aksi boikot produk Israel," *Jurnal Informatika dan Teknik Elektro Terapan*, vol. 12, no. 3, Sep. 2024, doi: 10.23960/jitet.v12i3.4743.