

Analisis Sentimen Berbasis Aspek Terhadap Ulasan Pengguna Aplikasi Pegadaian Digital Menggunakan Algoritma Naïve Bayes

Syfyani Alya Fauziyyah¹, Faqih Hamami², Rachmadita Andreswari³

^{1,2,3}Universitas Telkom, Bandung, Indonesia

E-mail: ¹syfyani@student.telkomuniversity.ac.id,

²faqihhamami@telkomuniversity.ac.id, ³andreswari@telkomuniversity.ac.id

Abstract

Pegadaian. PT. Pegadaian's form of transformation is the launch of Pegadaian Digital application. The application aims to facilitate the community and improve the service of products owned by PT. Pegadaian. Based on the monitoring as of 20 October 2022, the Pegadaian Digital application received 3.5 points on a scale of 5. This score is low because it contains many negative reviews. Therefore, it is necessary to analyse the review section of the application to increase the score. The method that can be used to analyse it is aspect-based sentiment analysis. Aspects are those that relate to the experience felt by users, namely aspects of learnability, efficiency, errors, and satisfaction. Sentiment analysis requires an optimal algorithm, one of which is Naïve Bayes. This algorithm was chosen because it is known as a simple but efficient algorithm when processing large amounts of data. This research uses two test scenarios, the first scenario using different ratios and base parameters and the second scenario using the addition of smoothing parameters. The result of this research is that the model with the ratio of 80:20 and the addition of smoothing is the best model for sentiment analysis because it produces the best performance value, with an accuracy value of 92%, precision of 80%, recall of 70% and f1-score of 73%.

Keywords: *Pegadaian Digital, Aspect-Based Sentiment Analysis, Naïve Bayes, User Review*

Abstrak

Bentuk transformasi dari PT. Pegadaian adalah diluncurkannya aplikasi Pegadaian Digital. Aplikasi tersebut memiliki tujuan untuk memudahkan masyarakat serta meningkatkan pelayanan dari produk yang dimiliki PT. Pegadaian. Berdasarkan pantauan per tanggal 20 Oktober 2022, aplikasi Pegadaian Digital mendapatkan poin 3,5 dari skala 5. Skor tersebut merupakan skor yang kecil karena mengandung banyaknya ulasan negatif. Oleh karena itu, perlu dilakukannya analisis pada bagian ulasan aplikasi supaya rating tersebut meningkat. Cara yang dapat digunakan untuk menganalisisnya yaitu dengan menggunakan sentiment analisis berbasis aspek. Dengan aspek yang mengacu pada pengalaman yang dirasakan pengguna, yaitu aspek learnability, efficiency, error, dan satisfaction. Ketika melakukan analisis sentiment dibutuhkan algoritma yang optimal, salah satunya yaitu Naïve Bayes. Algoritma ini dipilih karena dikenal sebagai algoritma yang sederhana tapi efisien ketika memproses data dengan jumlah yang besar. Penelitian ini menggunakan dua skenario pengujian, dengan skenario pertama menggunakan rasio berbeda dan parameter dasar dan skenario kedua menggunakan penambahan parameter smoothing. Hasil dari penelitian ini yaitu model dengan rasio 80:20 dan penambahan smoothing menjadi model terbaik untuk dilakukan analisis sentiment karena menghasilkan nilai performansi yang paling baik, dengan nilai akurasi sebesar 92%, presisi 80%, recall 70% dan f1-score 73%.

Kata Kunci: *Pegadaian Digital, Analisis Sentimen Berbasis Aspek, Naïve Bayes, Review Pengguna*

1. Pendahuluan

Seiring berkembangnya teknologi dan informasi, akan mendorong perubahan perilaku masyarakat dimana hampir semua aktivitas telah menggunakan teknologi, salah satunya adalah internet. Berdasarkan laporan Asosiasi Penyelenggara Jasa Internet Indonesia (APJII) tahun 2019-2020 (Q2) menunjukkan bahwa 73% dari 266 juta jiwa masyarakat Indonesia telah terhubung internet, yang mana mengalami peningkatan dari tahun 2018 yaitu sebesar 64,8% [1]. Peningkatan ini memberikan dampak yang besar pada berbagai bidang misalnya bidang pendidikan, ekonomi, sosial, politik, maupun budaya. Perkembangan pengguna internet yang sangat pesat menjadi salah satu faktor bagi sebuah perusahaan untuk melakukan transformasi digital.

Salah satu perusahaan yang telah melakukan transformasi digital adalah PT Pegadaian. PT Pegadaian adalah salah satu perusahaan di bawah naungan Badan Usaha Milik Negara (BUMN) yang bergerak di bidang keuangan non-profit yang berkaitan dengan pinjaman gadai, pinjaman non-gadai, jasa, dan kerja sama. Pegadaian telah melakukan transformasi sejak 1 Oktober 2018 [2]. Salah satu transformasi yang dilakukan oleh PT Pegadaian adalah dengan diluncurkannya aplikasi Pegadaian Digital yang mana memiliki tujuan untuk memudahkan masyarakat dan meningkatkan pelayanan dari produk-produk yang dimiliki oleh PT Pegadaian. Berdasarkan pantauan tanggal 20 Oktober 2022, aplikasi Pegadaian Digital memiliki poin 3,5 dari skala 5. Poin tersebut merupakan nilai yang cukup kecil bagi perusahaan BUMN karena pada aplikasi terdapat ulasan negatif yang relatif banyak.

Menurut penelitian yang dilakukan oleh *Apptentive* menunjukkan bahwa 50% pengguna tidak mempertimbangkan aplikasi dengan *rating* bintang 3 dan turun menjadi 85% untuk *rating* bintang 2. Selain itu, 77% orang membaca setidaknya 1 ulasan sebelum mengunduh aplikasi gratis dan 80% untuk aplikasi berbayar [3]. Oleh karena itu, perlu dilakukan analisis pada ulasan aplikasi agar *rating* aplikasi Pegadaian Digital dapat meningkat. Cara yang dapat digunakan untuk menganalisis data ulasan tersebut adalah dengan menggunakan analisis sentimen. Analisis sentimen dapat mengklasifikasikan sebuah opini menjadi positif, negatif, atau netral [4]. Untuk mendapatkan hasil yang lebih detail, diperlukan analisis sentimen yang melihat beberapa aspek yang terdapat pada ulasan aplikasi. Metode analisis sentimen tersebut biasanya disebut dengan metode *Aspect Based Sentiment Analysis* atau dapat disingkat menjadi ABSA. ABSA merupakan analisis sentimen yang bertujuan untuk mengidentifikasi polaritas sentimen dari aspek-aspek yang ada pada setiap kalimat [5]. Aspek-aspek yang dipilih mengacu pada pengalaman yang dirasakan oleh pengguna ketika menggunakan aplikasi, secara garis besar aspek-aspek tersebut adalah *learnability*, *efficiency*, *error*, dan *satisfaction*. Hasil yang akan didapatkan merupakan informasi yang dapat digunakan oleh perusahaan untuk memperbaiki dan meningkatkan rating aplikasi.

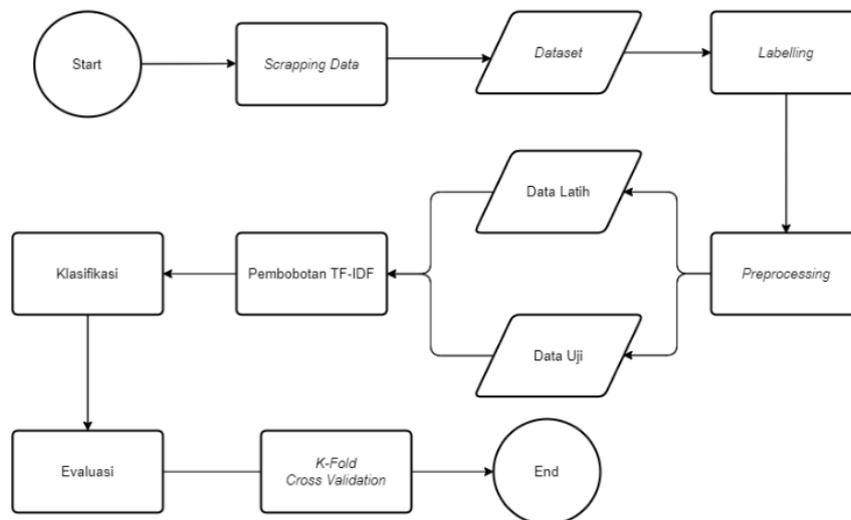
Ketika menggunakan analisis sentimen, dibutuhkan algoritma yang optimal, salah satunya yaitu *Naïve Bayes*. *Naïve Bayes* merupakan algoritma klasifikasi yang didasarkan pada *Teorema Bayes* yang dikenal sebagai algoritma yang sederhana namun efisien ketika mengolah data dalam jumlah besar [6]. Hal ini dapat diperkuat dari penelitian yang dilakukan oleh Parasati dkk. dengan judul "Analisis Sentimen Berbasis Aspek pada Ulasan Pelanggan Rumah Makan Bakso President Malang dengan Metode *Naïve Bayes Classifier*". Penelitian tersebut menggunakan data ulasan dari situs *TripAdvisor* sebanyak 943 data dan *Google Review* sebanyak 1.209 data dengan perbandingan rasio antara data latih (*training data*) dan data uji (*test data*) yaitu 50% : 50%, menghasilkan nilai akurasi aspek makanan sebesar 88%, aspek layanan sebesar 76%, serta aspek atmosfer sebesar 84% [7]. Tujuan dari penelitian ini adalah untuk mengetahui model analisis sentimen berbasis aspek pada ulasan aplikasi pegadaian digital dengan menggunakan algoritma *Naïve Bayes* dan mengetahui tingkat akurasi analisis sentimen berbasis aspek pada ulasan aplikasi pegadaian digital. Penelitian ini dilakukan dengan harapan dapat memberikan

manfaat kepada perusahaan berupa penyebab rendahnya rating aplikasi sehingga dapat melakukan evaluasi terhadap aplikasi Pegadaian Digital.

2. Metodologi Penelitian

2.1. Sistematisasi Penyelesaian Masalah

Sistem penyelesaian masalah yang digunakan dalam penelitian ini adalah KDD (*Knowledge Discovery in Database*). KDD adalah proses menemukan informasi yang berguna dalam data yang melibatkan algoritma untuk mengidentifikasi pola dalam data [8]. Penelitian ini menerapkan proses KDD yang telah disesuaikan dengan kebutuhan. Gambar di bawah ini merupakan ilustrasi tahapan yang akan dilakukan dalam penelitian ini.



Gambar 1. Sistematisasi Penyelesaian Masalah

Berdasarkan Gambar 1, penelitian ini dimulai dengan melakukan pengumpulan data pada aplikasi Pegadaian Digital yang tersedia di Google Play Store dan melakukan pelabelan data (*labelling*). Kemudian dilanjutkan *preprocessing* yang mana dilakukan *data cleaning*, standarisasi kata, tokenisasi, *stopword removal*, dan *stemming*. Proses selanjutnya yaitu dilakukan pembagian data (*split data*) berupa data latih dan data uji dan dilanjutkan dengan pembobotan kata dengan TF-IDF. Setelah dilakukan pembobotan kata, selanjutnya yaitu melakukan klasifikasi menggunakan algoritma *Multinomial Naïve Bayes*. Terakhir yaitu dilakukannya evaluasi model menggunakan *confusion matrix* dan *k-fold cross validation*.

2.2. Pengumpulan Data

Dataset yang digunakan dalam penelitian ini adalah *dataset* ulasan aplikasi Pegadaian Digital. Data dikumpulkan dengan melakukan *scraping* pada tanggal 13 Maret 2023 menggunakan *library google_play_scraper* pada *python* dengan jumlah data yang terkumpul sebanyak 22.000 data, namun tidak semua data dari *dataset* tersebut digunakan dalam penelitian ini. Data yang digunakan untuk penelitian ini memiliki kriteria yaitu, data ulasan aplikasi versi 5.0 hingga 5.4. Hal ini dikarenakan pada versi tersebut aplikasi memiliki perubahan yang cukup signifikan mulai dari perubahan UI, perbaikan, dan penambahan beberapa fitur sehingga data tersebut lebih relevan untuk dijadikan pertimbangan perusahaan dalam melakukan pengembangan aplikasi. Karena kriteria tersebut, maka data yang didapatkan untuk penelitian ini sebanyak 3.547 data.

2.3. Labelling

Setelah *dataset* dikumpulkan melalui proses *scraping*, langkah pertama yang harus dilakukan adalah melakukan pelabelan data atau *labelling*. Proses pelabelan data dilakukan secara manual oleh tiga orang mahasiswa Sistem Informasi Universitas Telkom dan dua orang pengguna aplikasi Pegadaian Digital dengan aspek yang didasarkan pada faktor *usability*. Menurut Nielsen (1994) *usability* mencakup semua area sistem dari aplikasi yang berinteraksi langsung dengan pengguna [9]. Aspek-aspek tersebut antara lain *learnability*, *efficiency*, *memorability*, *error*, dan *satisfaction*. Namun, pada penelitian ini hanya menggunakan aspek *learnability*, *efficiency*, *error*, dan *satisfaction*. Karena merujuk pada penelitian sebelumnya, aspek *learnability* dan *memorability* dianggap sebagai aspek yang memiliki interpretasi yang sama [9]. Pelabelan ulasan menggunakan label 0 yang berarti ulasan tersebut tidak mengandung aspek, 1 yang berarti ulasan tersebut mengandung aspek positif, dan 2 yang berarti ulasan tersebut mengandung aspek namun memiliki nilai negatif.

2.4. Preprocessing

Setelah melakukan pelabelan data, proses selanjutnya adalah *preprocessing*. *Dataset* yang telah diberikan label masih berupa data mentah (*raw data*) dimana data tersebut masih belum terstruktur yang dapat menyebabkan ketidaktepatan model yang dihasilkan sehingga proses *preprocessing* perlu dilakukan. Tahapan *preprocessing* meliputi sebagai berikut:

- a) *Data cleaning*, adalah proses pembersihan data dari *noise* yang tidak berguna dengan cara menghapus atribut yang tidak terpakai, mengisi nilai yang kosong, mengubah tipe data, menghapus karakter dan angka, penghapusan spasi, dan *case folding*.
- b) Standarisasi kata, adalah proses untuk mengubah atau mengganti kata-kata yang disingkat atau kata-kata yang memiliki kesalahan ketik pada saat penulisan ulasan.
- c) Tokenisasi, adalah proses segmentasi teks dengan membagi ulasan menjadi beberapa segmen dalam bentuk token.
- d) *Stopword removal*, adalah proses penghapusan kata-kata yang tidak memiliki arti penting namun sering muncul dalam ulasan.
- e) *Stemming*, adalah proses mengubah kata yang memiliki imbuhan menjadi kata dasar.

2.5. Split Data

Setelah *dataset* melewati *preprocessing* selanjutnya yaitu *split data*. Pada proses ini data dibagi menjadi dua bagian, yaitu data latih (*data training*) dan data uji (*data testing*). Pembagian data tersebut menggunakan rasio 60:40, 70:30, 80:20, dan 90:10.

2.6. TF-IDF

Setelah data dibagi menjadi data latih dan data uji, langkah selanjutnya adalah menghitung pembobotan kata berdasarkan TF-IDF. *Term* akan dihitung bobot frekuensinya dalam dokumen (TF) kemudian dilanjutkan dengan menghitung nilai IDF-nya. Berikut ini adalah persamaan untuk mencari nilai TF-IDF:

$$tfidf(t, d) = tf(t, d) \cdot idf(t) \quad (1)$$

$$idf(t) = \log \frac{(1+n)}{(1+df(t))} + 1 \quad (2)$$

Keterangan:

- $tf(t, d)$ = Jumlah kata t dalam dokumen (d)
 $idf(t)$ = Invers frekuensi dokumen
 n = Jumlah total dokumen
 $df(t)$ = jumlah kata t dalam dokumen

2.7. Implementasi Algoritma

Setelah melakukan pembobotan menggunakan TF-IDF, langkah selanjutnya adalah mengimplementasikan algoritma *Multinomial Naive Bayes* untuk mendapatkan model. Berikut ini adalah langkah-langkah dalam implementasi.

- a) Menghitung probabilitas setiap kelas dalam suatu aspek, yaitu kelas 0 (tidak ada aspek), 1 (positif), dan 2 (negatif). Berikut merupakan persamaan untuk menghitung probabilitas setiap kelas:

$$\hat{P}(c) = \frac{N_c}{N} \quad (3)$$

Keterangan:

$\hat{P}(c)$ = Probabilitas pada kelas atau *prior* suatu dokumen yang terdapat dalam kelas c

N_c = Jumlah semua dokumen pada kelas

N = Jumlah semua dokumen secara keseluruhan

- b) Menghitung nilai probabilitas setiap *term* yang terdapat pada dokumen. Berikut merupakan persamaannya:

$$\hat{P}(t|c) = \frac{W_{ct} + 1}{(\sum_{w' \in v} W_{ct'}) + B'} \quad (4)$$

Keterangan:

W_{ct} = Bobot TF-IDF *term* t pada dokumen

$\sum_{w' \in v} W_{ct'}$ = Jumlah bobot TF-IDF semua *term* pada kelas c

B = Jumlah IDF semua *term*

- c) Menghitung nilai probabilitas sentimen suatu dokumen berdasarkan nilai probabilitas *term* sebelumnya (data latih). Berikut merupakan persamaannya:

$$P(c|d) = P(c) \prod_{n=1}^i P(t_n|c) \quad (5)$$

Keterangan:

$P(c|d)$ = Probabilitas dokumen masuk terhadap kelas c

$P(c)$ = Probabilitas *prior* kelas c

$P(t_n|c)$ = Probabilitas bersyarat *term* t_n pada dokumen kelas c

2.8. Hyperparameter Tuning

Setelah mengimplementasikan algoritma, langkah selanjutnya adalah melakukan *hyperparameter tuning* dengan tujuan untuk mendapatkan parameter yang terbaik untuk digunakan. Tujuannya adalah untuk mendapatkan model algoritma yang optimal. Rasio yang digunakan dalam penerapan *hyperparameter tuning* adalah rasio terbaik yang didapatkan dari skenario pengujian pertama. Sedangkan parameter yang diujikan adalah parameter *smoothing* dengan rentang $0 < \alpha \leq 1$.

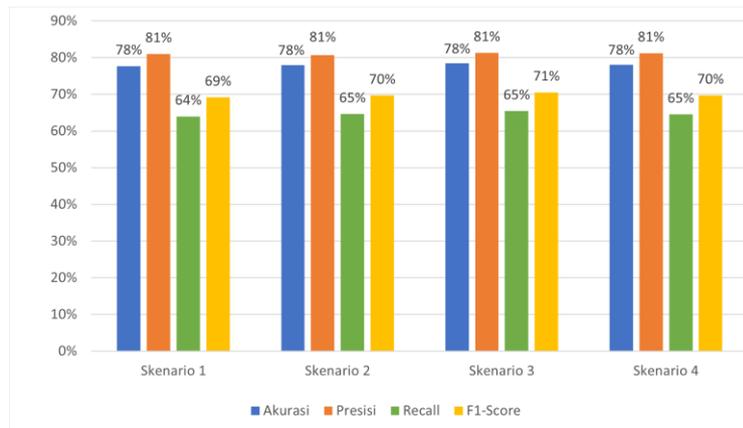
2.9. Evaluasi

Tahap terakhir adalah evaluasi model. Pada tahap ini, model dievaluasi dengan menggunakan *confusion matrix* untuk mendapatkan nilai akurasi, presisi, *recall*, dan *f1-score* serta *k-fold cross validation* untuk mendeteksi apakah terdapat *overfitting* pada model.

3. Hasil dan Pembahasan

3.1. Skenario Pengujian Pertama

Pada skenario pengujian pertama, *dataset* akan diuji dengan menggunakan rasio pembagian data yang berbeda dan menggunakan parameter dasar algoritma *multinomial naïve bayes*. Rasio yang digunakan adalah 60:40, 70:30, 80:20, dan 90:10. Tujuan dari pengujian rasio data yang berbeda adalah untuk mengetahui rasio mana yang memiliki performansi terbaik saat melakukan analisis sentimen menggunakan algoritma *multinomial naïve bayes*. Berikut ini adalah hasil skenario pengujian pertama.

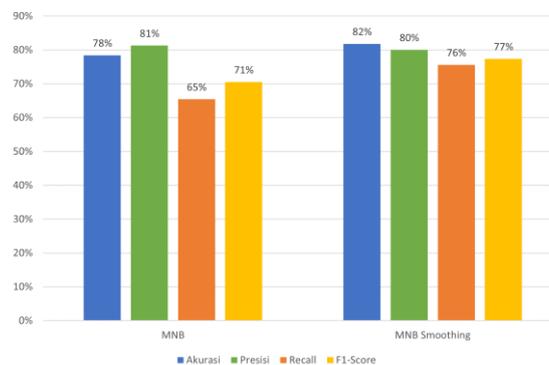


Gambar 2. Hasil Skenario Pengujian Pertama

Berdasarkan Gambar 2, menunjukkan bahwa skenario 3 dengan rasio 80:20 memiliki performansi yang paling baik jika dibandingkan dengan skenario lainnya. Rata-rata performansi pada skenario 3 yaitu 78% akurasi, 81% presisi, 65% *recall*, dan 71% *f1-score*.

3.2. Skenario Pengujian Kedua

Pada skenario pengujian kedua, *dataset* akan diuji dengan menggunakan rasio 80:20 yang merupakan rasio terbaik dari skenario pengujian pertama. Pada pengujian ini ditambahkan *smoothing* pada algoritma *multinomial naïve bayes* dengan parameter yang didapatkan dari hasil *GridSearchCV*. Parameter *smoothing* yang didapatkan adalah $\alpha = 0,1$. Berikut merupakan hasil skenario pengujian kedua.



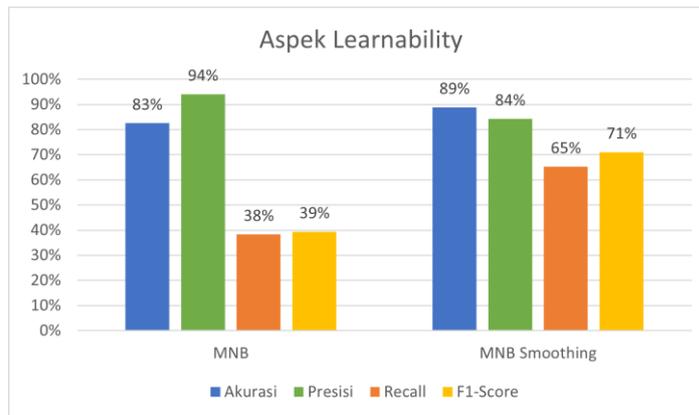
Gambar 3. Hasil Skenario Pengujian Kedua

Berdasarkan Gambar 3, dapat dilihat bahwa penambahan parameter *smoothing* mempengaruhi performansi model yang dihasilkan. Performa model ketika menggunakan parameter dasar memiliki nilai rata-rata akurasi 78%, presisi 81%, *recall* 65%, dan *f1-*

score: 71%. Sedangkan ketika ditambahkan parameter *smoothing*, nilai rata-rata akurasi, *recall*, dan *f1-score* mengalami peningkatan, yaitu 82% untuk akurasi, 76% untuk *recall* dan 77% untuk *f1-score*. Namun, nilai presisi mengalami penurunan ketika menggunakan *smoothing* yaitu menjadi 80%. Hal ini mengindikasikan bahwa model tidak dapat memprediksi kelas positif dengan benar.

3.3. Evaluasi Aspek

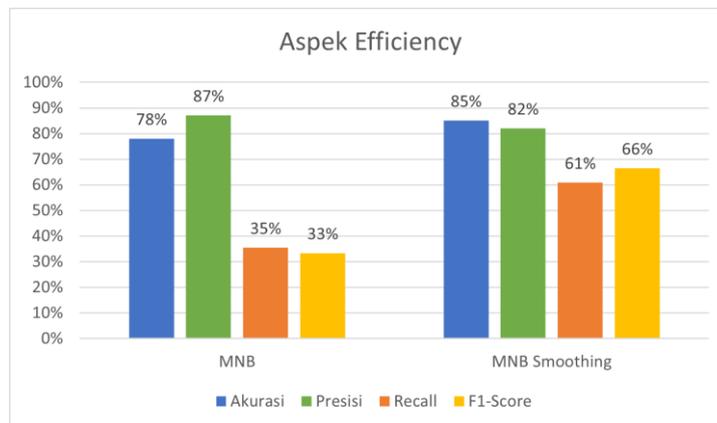
3.3.1. Aspek Learnability



Gambar 4. Performansi Aspek *Learnability*

Berdasarkan Gambar 4, menunjukkan bahwa pada aspek *learnability*, performansi dengan menggunakan parameter *smoothing* lebih baik jika dibandingkan dengan parameter dasar yaitu dengan nilai akurasi 89%, *recall* 65%, dan *f1-score* 71%. Sedangkan presisi lebih baik ketika menggunakan parameter dasar yaitu 94%.

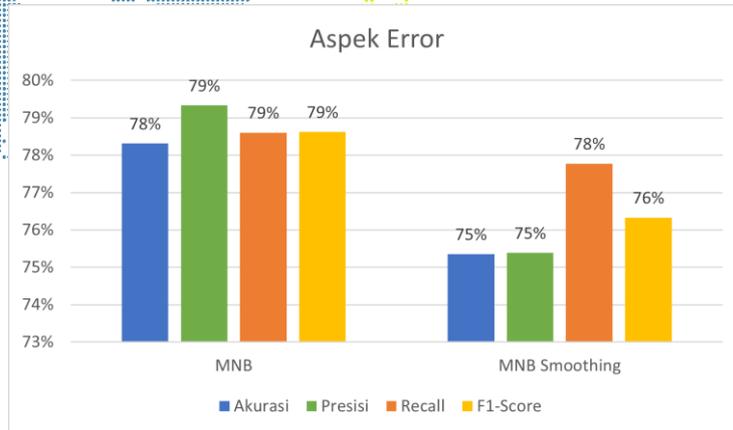
3.3.2. Aspek Efficiency



Gambar 5. Performansi Aspek *Efficiency*

Berdasarkan Gambar 5 menunjukkan bahwa pada aspek *efficiency*, performansi dengan menggunakan parameter *smoothing* lebih baik jika dibandingkan dengan parameter dasar yaitu dengan nilai akurasi sebesar 85%, *recall* 61%, dan *f1-score* 66%. Sedangkan nilai presisi lebih baik ketika menggunakan parameter dasar yaitu 87%.

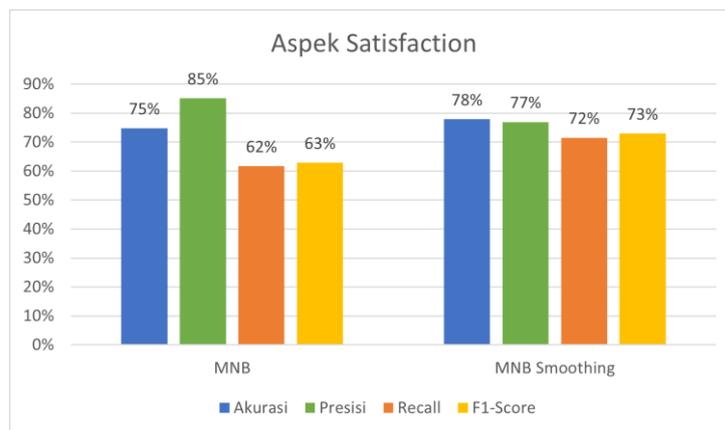
3.3.3. Aspek Error



Gambar 6. Performansi Aspek Error

Berdasarkan Gambar 6 menunjukkan bahwa performansi pada aspek *error* lebih baik menggunakan parameter dasar yaitu dengan nilai akurasi sebesar 78%, presisi 79%, *recall* 79%, dan *f1-score* sebesar 79%.

3.3.4. Aspek Satisfaction



Gambar 7. Performansi Aspek Satisfaction

Berdasarkan Gambar 7 menunjukkan bahwa aspek *satisfaction* memiliki performansi lebih baik ketika menggunakan parameter *smoothing* yaitu dengan akurasi sebesar 78%, *recall* sebesar 72%, dan *f1-score* sebesar 73%. Untuk nilai presisi lebih baik ketika menggunakan parameter dasar yaitu 85%.

3.4. Evaluasi K-Fold Cross Validation

Evaluasi menggunakan *k-fold cross validation* bertujuan untuk menguji performansi model secara objektif dan untuk mengetahui kemungkinan terjadinya *overfitting* pada model yang dihasilkan. Pengujian *k-fold cross validation* menggunakan *5-fold* ($k=5$) dengan menggunakan skenario pengujian kedua yaitu rasio perbandingan 80:20 serta penggunaan parameter dasar dan *smoothing* dalam pengujian. Berikut ini adalah hasil dari *k-fold cross validation* skenario pengujian kedua.

Tabel 1. Evaluasi K-Fold Cross Validation

Pengujian	Nilai Akurasi	
	Akurasi Biasa	Akurasi K-Fold Cross Validation

Penguujian	Nilai Akurasi	
	Akurasi Biasa	Akurasi K-Fold Cross Validation
Parameter Dasar	0,7842	0,7793
Parameter <i>Smoothing</i>	0,8180	0,8188

Berdasarkan Tabel IV-1 menunjukkan bahwa hasil evaluasi menggunakan *k-fold cross validation* dengan *5-fold* memiliki rata-rata akurasi yang tidak jauh berbeda dengan akurasi yang diperoleh model dari hasil pengujian. Hal tersebut mengindikasikan bahwa model dengan menggunakan parameter dasar dan parameter *smoothing* tidak mengalami *overfitting*.

4. Kesimpulan

Penelitian ini berfokus pada analisis sentimen berbasis aspek terhadap ulasan pengguna aplikasi Pegadaian Digital. Penelitian ini membahas aspek *learnability*, *efficiency*, *error*, dan *satisfaction*. *Dataset* yang digunakan untuk penelitian ini adalah data ulasan aplikasi yang diambil pada *google play store*. Total *dataset* yang digunakan adalah 3.547 data yang kemudian diberikan label dan dibersihkan melalui *preprocessing*, dilanjutkan dengan pembagian data, dan pembobotan TF-IDF. Setelah dilakukan pembobotan TF-IDF, dilanjutkan dengan klasifikasi menggunakan algoritma *multinomial naïve bayes* dan diakhiri dengan evaluasi. Skenario yang digunakan adalah dengan membagi *dataset* menjadi rasio yang berbeda dengan parameter dasar dan rasio terbaik dengan penambahan parameter *smoothing*. Hasil yang didapatkan adalah rasio 80:20 merupakan rasio terbaik pada skenario pengujian pertama dengan rata-rata akurasi 78%, presisi 81%, *recall* 65%, dan % *f1-score* 71. Serta akan lebih baik lagi jika rasio tersebut menggunakan parameter *smoothing* dengan $\alpha = 0,1$. Hasil yang didapatkan ketika menggunakan parameter *smoothing* pada algoritma *multinomial naïve bayes* yaitu dengan rata-rata akurasi sebesar 82%, *recall* sebesar 76%, dan *f1-score* sebesar 77%. Namun, nilai presisi menurun ketika menggunakan parameter *smoothing* menjadi 80% yang mengindikasikan bahwa model tidak dapat memprediksi kelas positif dengan benar. Untuk hasil evaluasi menggunakan *k-fold cross validation* dengan *5-fold* membuktikan bahwa model tidak mengalami *overfitting* ketika menggunakan parameter dasar maupun parameter *smoothing*.

Daftar Pustaka

- [1] Arindra Meodia, "Riset APJII: 73,7 persen masyarakat Indonesia terhubung internet," *ANTARANEWS*, Nov. 09, 2020. <https://www.antaraneWS.com/berita/1829936/riset-apjii-737-persen-masyarakat-indonesia-terhubung-internet#mobile-nav> (accessed Jul. 30, 2023).
- [2] "Mantul, Pegadaian Dinobatkan Sebagai Salah Satu Perusahaan yang Berhasil Menerapkan Tranformasi Bisnis," *Pegadaian*, Jul. 19, 2021.
- [3] "How ratings and reviews affect consumers decision to download apps," *Business of Apps*, Jan. 22, 2020.
- [4] Taeho Jo, "Text Mining," 2019. [Online]. Available: <http://www.springer.com/series/11970>
- [5] C. Zong, R. Xia, and J. Zhang, "Text Data Mining," 2021.
- [6] S. Raschka, "Naive Bayes and Text Classification I - Introduction and Theory," Oct. 2014, [Online]. Available: <http://arxiv.org/abs/1410.5329>
- [7] W. Parasati, F. Abdurrachman Bachtiar, and N. Y. Setiawan, "Analisis Sentimen Berbasis Aspek pada Ulasan Pelanggan Restoran Bakso President Malang dengan Metode Naïve Bayes Classifier," 2020. [Online]. Available: <http://j-ptiik.ub.ac.id>

- [8] R. Sharda, D. Delen, E. Turban, J. E. Aronson, T.-P. Liang, and David. King, *Business intelligence analytics, and data science: a managerial perspective*. 2018.
- [9] H. Mustakim and S. Priyanta, "Aspect-Based Sentiment Analysis of KAI Access Reviews Using NBC and SVM," *IJCCS (Indonesian Journal of Computing and Cybernetics Systems)*, vol. 16, no. 2, p. 113, Apr. 2022, doi: 10.22146/ijccs.68903.