

Analisis Sentimen Pengguna Aplikasi MyPertamina Dengan Menggunakan Algoritma Naïve Bayes

I Gusti Bgs Darmika Putra^{a1}, Cokorda Rai Adi Pramarthar^{a2}

^aProgram Studi Informatika, Universitas Udayana
Bali, Indonesia

¹gstbgsdarmika@gmail.com

²cokorda@unud.ac.id

Abstrak

MyPertamina adalah aplikasi layanan keuangan digital yang di miliki oleh PT Pertamina dan anggota Badan Usaha Milik Negara (BUMN), dengan menggunakan aplikasi ini pengguna dapat membeli beberapa produk buatan Pertamina secara cashless atau nontunai Tujuan dari aplikasi MyPertamina yaitu untuk mendata masyarakat yang telah membeli BBM bersubsidi. Dalam membangun aplikasi, diperlukan keahlian khusus serta banyak hal yang harus dipertimbangkan, termasuk ulasan-ulasan pengguna setelah menggunakan aplikasi karena dapat menentukan kualitas dari aplikasi yang diciptakan apakah bagus atau justru sebaliknya. Dengan analisis sentiment, pengembang aplikasi dapat dengan mudah untuk mengklasifikasikan ulasan karena data ulasan akan dikelompokkan menjadi ulasan positif atau ulasan negative. Tujuan dari penelitian ini, yaitu untuk mengetahui sentimen pengguna aplikasi MyPertamina yang terdapat pada ulasan Google Playstore dengan menerapkan metode Naïve Bayes untuk mengklasifikasikan sentiment dan menggunakan metode tf-idf yang digunakan untuk mencari representasi nilai dari tiap-tiap dokumen dari suatu kumpulan data training (training set). Dari hasil penelitian memperoleh hasil akurasi sebesar 72% dengan menggunakan 200 data ulasan pengguna aplikasi MyPertamina yang terdiri dari 100 ulasan positif dan 100 ulasan negative.

Keywords: Analisis Sentimen, Naïve Bayes, Multinomial Naïve Bayes, TF-IDF, Ulasan Aplikasi, Aplikasi Mypertamina,

1. Pendahuluan

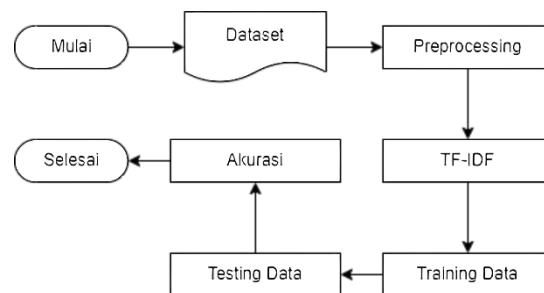
Perkembangan teknologi ini membuat semua aspek kehidupan tidak dapat terlepas dari penggunaan teknologi, baik secara langsung maupun secara tidak langsung. Perkembangan teknologi memunculkan berbagai jenis aplikasi digital yang dapat menunjang kebutuhan manusia untuk membuat pekerjaan manusia menjadi lebih mudah, cepat dan efisien. Pada tahun 2012, terdapat sekitar 700.000 aplikasi digital yang tersedia untuk pengguna android serta terdapat sekitar 25 juta aplikasi yang telah diunduh dari Google Playstore atau toko aplikasi utama Android [1]. Aplikasi yang ada di Google Playstore memiliki berbagai macam fungsi dan manfaat bagi pengguna yang dapat menunjang kebutuhan pengguna, salah satunya yaitu aplikasi MyPertamina. Aplikasi ini merupakan aplikasi layanan keuangan digital yang di miliki oleh PT Pertamina dan anggota Badan Usaha Milik Negara (BUMN) dengan menggunakan aplikasi ini pengguna dapat untuk membeli beberapa produk buatan Pertamina, termasuk BBM, secara cashless atau nontunai. Tujuan dari aplikasi MyPertamina yaitu untuk mendata masyarakat yang telah membeli BBM bersubsidi. Alih-alih untuk mempermudah pengguna dalam melakukan pembayaran pembelian BBM melalui aplikasi, malah justru kebalikannya yaitu membuat pengguna kesulitan dalam menggunakan aplikasi hal ini dapat dilihat dari ulasan yang diberikan oleh pengguna yang telah menggunakan aplikasi MyPertamina. Hal ini membuat tuntutan terhadap mutu pelayanan aplikasi MyPertamina harus di tingkatkan agar pengalaman pengguna dapat lebih baik lagi.

Klasifikasi sentiment bertujuan untuk mengatasi masalah ini dengan cara otomatis mengklasifikasikan ulasan dari pengguna dengan mengelompokkan menjadi ulasan positif atau ulasan negatif [2]. Terdapat beberapa penelitian sebelumnya mengenai ulasan aplikasi yang terdapat pada Google Playstore diantaranya, penelitian yang dilakukan oleh [3] pada tahun 2020 dengan judul Algoritma Klasifikasi Naïve Bayes untuk Analisa Sentimen Aplikasi Shopee. Hasil penelitian ini menghasilkan nilai akurasi sebesar 71.50% dan Nilai AUC (Area Under Curve) sebesar 0,500. Penelitian selanjutnya oleh [4] pada tahun 2019 dengan judul Analisis Sentimen Aplikasi Transportasi Online Krl Access Menggunakan Metode Naïve Bayes. Penelitian ini menggunakan metode NBC yang memberikan hasil akurasi hingga 84.00% untuk data uji review opini positif berbahasa Indonesia pada pemilihan aplikasi transportasi darat pada smartphone.

Merancang suatu aplikasi diperlukan keahlian khusus serta banyak hal yang harus dipertimbangkan dalam proses perancangannya, namun dalam merancang suatu aplikasi diperlukan juga untuk memperhatikan ulasan-ulasan yang diberikan dari pengguna terhadap aplikasi yang telah diciptakan karena dengan melihat ulasan tersebut pengembang aplikasi tau hal apa saja yang membuat aplikasi buruk atau sebaliknya, sehingga memudahkan pengembang aplikasi dalam menambahkan dan memperbaiki fitur dari aplikasi yang diciptakan. Dalam penelitian ini, penulis melakukan analisis sentimen pada aplikasi MyPertamina untuk mengetahui apakah ulasan yang diberikan oleh pengguna termasuk ke ulasan positif atau negatif. Hasil akhir yang akan dihasilkan adalah tingkat akurasi yang dicapai dalam melakukan analisis sentimen menggunakan metode Naïve Bayes.

2. Metode Penelitian

Pada penelitian tentang analisis sentimen pengguna aplikasi MyPertamina ini, penulis melakukan metode eksperimen, dengan tahapan berikut:



Gambar 1. Tahap Penelitian

2.1 DataSet

Pada penelitian ini, pengguna menggunakan data-data komentar yang ada pada tahun 2022. Data yang dikumpulkan sebanyak 200 data. Data komentar yang terdiri dari data komentar positif sebanyak 100 data, dan data komentar negatif sebanyak 100 data. Data ini bersumber dari <https://play.google.com>.

2.2 Preprocessing

Pada tahapan preprocessing yang penulis gunakan untuk dataset yang dimiliki adalah:

a. Tokenization

Pada proses ini, dilakukan pemotongan string input berdasarkan tiap kata yang menyusunnya dan kata yang memiliki tanda baca serta simbol yang bukan huruf dihilangkan.

b. Filtering (Stopword Removal)

Proses penghapusan atau pembuangan kata-kata yang sering ditampilkan dalam dokumen. Proses menghilangkan kata-kata tersebut adalah dengan melakukan pengecekan pada kamus *stopword*. Jika kata tersebut ada di dalam kamus *stopword*, maka kata tersebut akan dihilangkan. Stopword dapat berupa kata depan, kata sambung, kata seru, dan lain sebagainya. Contoh *stopword* dalam bahasa Indonesia adalah “yang”, “dan”, “di”, “dari”, dll.

c. Stemming

Stemming merupakan tahapan mengubah kata menjadi bentuk dasar dengan menghilangkan semua imbuhan. Terdapat dua pendekatan dalam melakukan stemming yaitu dengan pendekatan kamus dan pendekatan aturan. Pada penelitian ini menggunakan bantuan Sastrawi Python untuk mendapatkan kata dasar. merupakan library python yang dapat digunakan untuk mendapatkan kata dasar dari kata yang kita inputkan. Sastrawi Python sendiri bergantung pada kamus kata dasar dari www.kateglo.com. Algoritma yang digunakan oleh library ini adalah algoritma Nazief dan Andriani, di mana algoritma ini merupakan salah satu algoritma yang cukup populer untuk melakukan stemming kata dalam bahasa Indonesia. [5].

Tabel 1. Hasil Preprocessing

Data	Aplikasi bagus, cuman kok gak ad jawaban Pertamina ya,
Tokenization	aplikasi, bagus, cuman, kok, gak, ad, jawaban, pertamina, ya
Stopwords Removal	aplikasi, bagus, cuman, kok, gak, ad, jawaban, pertamina, ya
Stemming	aplikasi, bagus, cuman, gak, ad, pertamina, ya

2.3 Term Frequency Invers Document Frequency (TF-IDF)

TF-IDF merupakan metode yang merupakan integrasi antar term frequency (TF), dan inverse document frequency (IDF). Fungsi metode tf-idf adalah untuk mencari representasi nilai dari tiap-tiap dokumen dari suatu kumpulan data training (*training set*) dimana nantinya dibentuk suatu vektor Antara dokumen dengan kata (*documents with terms*) yang kemudian untuk kesamaan antar dokumen dengan *cluster* akan ditentukan oleh sebuah *prototype* vektor yang disebut juga dengan *cluster centroid* [6].

2.4 Naïve Bayes

Naïve Bayes adalah salah satu algoritma klasifikasi yang sederhana dan sering digunakan untuk klasifikasi dokumen. Dalam proses membangun sistem pengklasifikasi menggunakan NB terdapat 2 tahapan yang dilakukan. Tahap pertama adalah proses pelatihan (training) dan tahap yang kedua adalah proses pengujian (testing) [7].

a. Multinomial Naïve Bayes

Multinomial Naïve Bayes adalah salah satu metode khusus dari metode Naïve Bayes yang menggunakan probabilitas bersyarat. Probabilitas bersyarat dapat dilakukan dengan menggunakan frekuensi kemunculan suatu kata dalam suatu kelas (*raw term frequency*) [8].

$$P(c|term\ dokumen\ d) = P(c) \times P(t1|c) \times P(t2|c) \times P(t3|c) \dots \times P(tn|c)$$

Probabilitas kelas c sebelumnya ditentukan oleh rumus:

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

Multinomial Naïve Bayes untuk nilai input x ditunjukkan pada persamaan berikut:

$$P(tn|c) = \frac{W(c,tn)+1}{(\sum W' < VW'(c,tn) + B')}$$

Keterangan :

$W(c, tn)$: Nilai pembobotan tfidf atau W dari term t di kategori c

$\sum W' < VW'ct$: Jumlah total W dari keseluruhan term yang berada di kategori c.

B' : Jumlah W kata unik (nilai idf tidak dikali dengan tf) pada seluruh dokumen.

b. Skenario Eksperimen

Perhitungan akurasi dari sistem yang berguna untuk mengukur kinerja sistem yang telah lakukan dengan menggunakan rumus:

$$Akurasi = \frac{Db}{N} \times 100\%$$

3. Hasil dan Pembahasan

Metode yang diusulkan menggunakan 200 data ulasan dari pengguna aplikasi MyPertamina dengan 100 data ulasan positif dan 100 data ulasan negative, data ini bersumber dari

<https://play.google.com>. Terdapat 20 % dataset digunakan untuk data uji dan sisanya digunakan sebagai data latih. Terdapat tahapan preprocessing data sebelum data ulasan pengguna digunakan analisis data. Proses ini berguna untuk memastikan kualitas data baik dan lebih bersih sehingga bisa digunakan untuk pengolahan selanjutnya. Terdapat beberapa tahapan dalam proses preprocessing yaitu: (1) Pertama, memisahkan kalimat ulasan menjadi beberapa kata dan mengubah seluruh data menjadi huruf kecil, (2) Menghilangkan tanda baca yang ada di setiap ulasan pengguna. (3) Kemudian penghapusan atau pembuangan kata-kata yang sering ditampilkan atau kata-kata yang tidak relevan menggunakan Sastrawi. Setelah melalui tahap preprocessing, hasilnya dari proses preprocessing akan dibobot menggunakan rumus *Term Frequency Invers Document Frequency* (TFIDF). Setelah menentukan bobot dari suatu term (kata) menggunakan tf-idf maka Langkah yang terakhir yaitu n klasifikasi menggunakan metode Naïve Bayes dan Multinomial Naïve Bayes.

Tabel 2. Hasil akurasi Naïve Bayes dan Multinomial Naïve Bayes

Accuracy	72 %
Precision	88 %
Recall	64 %
f1_score	74 %

Pada percobaan pertama, kami menggunakan 100 data ulasan yang diberikan oleh pengguna aplikasi MyPertamina yang terdiri dari 50 ulasan positif dan 50 ulasan negative namun akurasi dari percobaan pertama ini masih kurang baik yaitu 65 % dikarenakan data ulasan tersebut masih kurang bersih pada saat tahapan preprocessing data yang mengakibatkan akurasi dari mesin kurang maksimal. Kemudian untuk percobaan kedua kami menggunakan 200 data ulasan pengguna aplikasi yang terdiri dari 100 ulasan positif dan 100 ulasan negative, dari hasil percobaan kedua di dapatkan akurasi dari algoritma naïve bayes sebesar 72% sesuai dengan tabel 2 yang tertera diatas.

4. Kesimpulan

Metode Naïve Bayes dan Multinomial Naïve Bayes memberikan hasil akurasi hingga 72% untuk data uji ulasan pengguna terhadap aplikasi MyPertamina yang bersumber dari <https://play.google.com>. Metode Naïve Bayes dan Multinomial Naïve Bayes memberikan hasil yang tepat dalam mengklasifikasikan ulasan apakah termasuk dalam sentiment positif atau termasuk dalam sentiment negatif. Metode tf-idf digunakan untuk mencari representasi nilai dari tiap-tiap dokumen dari suatu kumpulan data training (*training set*).

Dari penelitian ini, kedua metode tersebut dapat memberikan hasil yang tepat dalam mengklasifikasikan ulasan pada aplikasi MyPertamina yang menentukan apakah ulasan tersebut termasuk kedalam sentiment positif atau negatif. Kedepannya dapat dilakukan perbaikan dengan melakukan tahapan preprocessing data yang lebih baik lagi agar data yang digunakan lebih bersih sehingga mendapatkan akurasi yang lebih baik lagi.

Referensi

- [1] R. Ridjaludin, I. A. Ratnamulyani, and A. A. Kusumadinata, "Pengaruh Penggunaan Layanan Aplikasi Digital Google Play Dalam Smartphone Terhadap Pemenuhan Kebutuhan Informasi Mahasiswa," *J. Komun.*, vol. 2, no. 2, pp. 135–146, 2017, doi: 10.30997/jk.v2i2.229.
- [2] R. Sari, "Analisis Sentimen Review Restoran menggunakan Algoritma Naive Bayes berbasis Particle Swarm Optimization," *J. Inform.*, vol. 6, no. 1, pp. 23–28, 2019, doi: 10.31311/ji.v6i1.4695.
- [3] S. Masripah and L. D. Utami, "Algoritma Klasifikasi Naïve Bayes untuk Analisa Sentimen Aplikasi Shopee," *Swabumi*, vol. 8, no. 2, pp. 114–117, 2020, doi: 10.31294/swabumi.v8i2.8444.
- [4] S. Nurwahyuni, "Analisis Sentimen Aplikasi Transportasi Online Krl Access Menggunakan Metode Naive Bayes," *Swabumi*, vol. 7, no. 1, pp. 31–36, 2019, doi: 10.31294/swabumi.v7i1.5575.
- [5] G. Setiawan, H. N. Palit, and E. Setyati, "Aspect Based Sentiment Analysis pada Layanan Umpan Balik Universitas dengan Menggunakan Metode Naïve Bayes dan Latent Semantic Analysis," *J. Infra*, vol. 7, no. 1, pp. 170–174, 2019.
- [6] N. K. Widyasanti, I. K. G. Darma Putra, and N. K. Dwi Rusjyanthi, "Seleksi Fitur Bobot Kata dengan Metode TFIDF untuk Ringkasan Bahasa Indonesia," *J. Ilm. Merpati (Menara Penelit.*

- [7] *Akad. Teknol. Informasi*, vol. 6, no. 2, p. 119, 2018, doi: 10.24843/jim.2018.v06.i02.p06.
D. P. Artanti, A. Syukur, A. Prihandono, and D. R. I. M. Setiadi, "Analisa Sentimen Untuk Penilaian Pelayanan Situs Belanja Online Menggunakan Algoritma Naïve Bayes," pp. 8–9, 2018.
- [8] I. G. C. P. Yasa, N. A. Sanjaya ER, and L. A. A. Rahning Putri, "Sentiment Analysis of Snack Review Using the Naïve Bayes Method," *JELIKU (Jurnal Elektron. Ilmu Komput. Udayana)*, vol. 8, no. 3, p. 333, 2020, doi: 10.24843/jlk.2020.v08.i03.p16.

This page is intentionally left blank.