

SENTIMENT ANALYSIS OF COMMENTS ON GOOGLE PLAY STORE, TWITTER AND YOUTUBE TO THE MYPERTAMINA APPLICATION WITH SUPPORT VECTOR MACHINE

Aji Nursalim^{*1}, Rice Novita²

^{1,2}Information Systems Study Program, Faculty of Science and Technology, UIN Suska Riau, Indonesia
Email: ¹11950311518@students.uin-suska.ac.id, ²rice.novita@uin-suska.ac.id

(Article received: May 21, 2023; Revision: May 30, 2023; published: Desember 23, 2023)

Abstract

Application is an important requirement in a business because it makes work more efficient thereby increasing the results of the company, Pertamina as a supplier of fuel oil (BBM) in Indonesia provides the latest innovations by launching the myPertamina application for purchasing BBM which raises public opinion, and conveys its aspirations in social media. Text mining is a way to group community comments because text mining has an analysis that focuses on analyzing a comment that is extracted into information. The purpose of this study was to determine public sentiment towards the use of myPertamina by classifying comments using the Support Vector Machine (SVM) algorithm and finding the best kernel among linear, polynomial and RBF. In this study, data was taken from three social media, namely Google Play Store with 18.000 data, Twitter with 20.000 data and YouTube with 6.400 data with a total of 44.400 data. Sentiment is carried out by giving positive and negative classes, the accuracy obtained from sentiment is carried out for Google Play Store data of 95%, Twitter 76% and YouTube 99% and it is known that the best svm kernel in this study is the RBF kernel which outperforms the linear and polynomial kernels.

Keywords: *application, myPertamina, sentiment, SVM.*

ANALISIS SENTIMEN KOMENTAR PADA GOOGLE PLAY STORE, TWITTER DAN YOUTUBE TERHADAP APLIKASI MYPERTAMINA DENGAN SUPPORT VECTOR MACHINE

Abstrak

Aplikasi merupakan kebutuhan yang penting dalam sebuah bisnis karena membuat pekerjaan menjadi lebih efisien sehingga meningkatkan hasil dari perusahaan, Pertamina selaku penyuplai bahan bakar minyak (BBM) di Indonesia memberikan inovasi terbaru dengan meluncurkan aplikasi myPertamina untuk pembelian BBM yang menimbulkan opini masyarakat, dan menyampaikan aspirasinya di sosial media. Text mining merupakan salah satu cara untuk mengelompokkan komentar masyarakat karena text mining memiliki analisa yang berfokus untuk menganalisa suatu komentar yang diekstrak menjadi sebuah informasi. Tujuan penelitian ini untuk mengetahui sentiment masyarakat terhadap penggunaan myPertamina dengan pengklasifikasian komentar menggunakan algoritma *Support Vector Machine* (SVM) dan mencari kernel terbaik diantara *linear*, *polynomial* dan *RBF*. Dalam penelitian ini data diambil dari tiga media sosial yaitu Google Play Store sebanyak 18.000 data, Twitter sebanyak 20.000 data dan YouTube sebanyak 6.400 data dengan total sebanyak 44.400 data. Sentimen dilakukan dengan memberikan kelas positif dan negatif, akurasi yang di dapat dari sentiment yang dilakukan untuk data google play store sebesar 95%, twitter 76% dan youtube 99% serta diketahui kernel svm terbaik dalam penelitian ini adalah kernel RBF yang mengungguli kernel *linear* dan *polynomial*.

Kata kunci: *aplikasi, myPertamina, sentiment, SVM.*

1. PENDAHULUAN

Indonesia merupakan salah satu dari lima negara dengan penduduk paling banyak di dunia, dengan banyaknya mobilitas penduduk di Indonesia

menimbulkan banyaknya rakyat yang menggunakan kendaraan dalam setiap kegiatan aktivitasnya. Berdasarkan data kendaraan yang diterbitkan laman korlantas.polri.go.id total kepemilikan kendaraan di Indonesia sebanyak 150.786.747 unit. Bahan Bakar

Minyak (BBM) ialah bahan bakar yang digunakan oleh rakyat Indonesia, banyaknya kendaraan di negara ini mengakibatkan besarnya permintaan BBM [1].

BBM sangat penting bagi masyarakat [2]. Indonesia memiliki kewenangan dalam menetapkan harga jual BBM. Pemerintah selalu menjaga harga BBM agar tetap stabil dan terjangkau untuk masyarakat Indonesia, namun tingginya konsumsi BBM masyarakat dan naiknya harga BBM Internasional mengakibatkan harga BBM dalam negeri harus disesuaikan dengan harga BBM internasional agar keberlangsungan fiscal negara tidak terancam dan tetap aman [3].

Pertamina selaku Perusahaan yang memproduksi serta mensuplai bahan bakar di Indonesia telah menerbitkan inovasi terbaru dalam segi pembayaran untuk memastikan kemudahan bagi konsumen dalam bertransaksi untuk pembelian berbagai produk Pertamina [4]. Untuk mengikuti program yang sedang diumumkan pemerintah yakni *cashless* ataupun transaksi non tunai, Pertamina menghimbau masyarakat agar menggunakan e-payment dari aplikasi MyPertamina yang bekerja sama dengan Link Aja[5].

Dari berbagai fasilitas yang PT. Pertamina berikan melalui MyPertamina, masih didapati persoalan yang ada pada masyarakat seperti proses membuka aplikasi mypertamina yang sulit, banyaknya SPBU yang tidak support dengan MyPertamina, sering gagalnya proses pendaftaran akun dan transaksi yang sukses pada MyPertamina tetapi poin konsumen tetap sama yang mengakibatkan menurunnya kepercayaan pada MyPertamina dan konsumen kurang mendapatkan keuntungan dari adanya MyPertamina [4].

Uji coba pembatasan pembelian BBM bersubsidi telah dilakukan oleh pemerintah dengan menerapkan MyPertamina untuk mengontrol pembelian BBM bersubsidi mulai juli 2022, agar dapat meringankan beban keuangan negara. Pada September 2019 diluncurkan terdapat 179.854 konsumen aktif, kemudian meningkat signifikan pada tahun 2021 sebanyak 5.975.478 namun terjadi penurunan pada bulan Januari-April 2022 sebanyak 4.923.805 konsumen aktif [6]. konsumen memberikan ulasan tentang aplikasi MyPertamina pada kolom komentar Google Play Store, Youtube maupun Twitter. Ulasan tersebut dikemas dalam bentuk komentar positif, dan negatif. Komentar yang didapat pada google play store sebanyak 18.000, twitter 20.000 dan youtub 6.400. Dari data komentar ini dapat dianalisis menggunakan *text mining* [7].

Text mining merupakan suatu proses dengan tujuan menemukan informasi atau tren terbaru yang belum terungkap [8], juga dapat diartikan sebagai suatu teknik untuk mengambil informasi dari sejumlah data *unsupervised* pada suatu topik tertentu yang berkualitas tinggi hingga didapat permasalahan data-data dalam teks [9]. Pada *Text Mining* terdapat

beberapa tahapan pengolahan sumber data diantaranya *Text Preprocessing*. Dalam *Text Preprocessing* terdapat beberapa tahapan seperti *Cleaning, Case Folding, Stopword, Tokenize, dan Stemming* [10].

Dalam analisis sentimen dengan metode penerapan *machine learning* ada beberapa metode yang sering digunakan seperti Decision Tree, KNN, NB dan SVM [11]. *Machine learning* merupakan salah satu cabang ilmu dari kecerdasan buatan (Intelligence), metode *machine learning* memanfaatkan bahasa pemrograman agar komputer dapat memproses data historis dalam jumlah besar dan mengidentifikasi pola data [12], untuk melakukan analisis pada sekumpulan data besar (*Big Data*) dapat diproses menggunakan *machine learning* untuk mendapatkan suatu pola tertentu [13]. Jika ingin mengklasifikasikan sebuah data maka dapat digunakan algoritma SVM.

Adapun kelebihan dari algoritma SVM yaitu mampu mengatasi permasalahan regresi dan klasifikasi secara linear ataupun non-linear serta akurasi yang dimiliki cenderung lebih baik [14]. Penelitian yang dilakukan Rokhman dkk, dalam penelitian-nya digunakan data sentimen aplikasi gojek, pada penelitian ini digunakan metode *Decision Tree* dan *Support Vectore Machine*, untuk *Decision Tree* diperoleh hasil akurasi sebesar 89.80% sedangkan untuk SVM diperoleh hasil sebesar 90.20% [15]. Sementara itu Ichwan dkk melakukan penelitian tingkat kematangan buah manga berdasarkan fitur warna menggunakan algoritma KNN dan SVM, berdasarkan hasil proses pengujian yang dilakukan didapat akurasi sebesar 83.3% untuk metode KNN sedangkan untuk metode SVM didapat akurasi sebesar 87.5% [16]. SVM memiliki beberapa tipe kernel diantaranya *linear, polynomial* dan RBF [17].

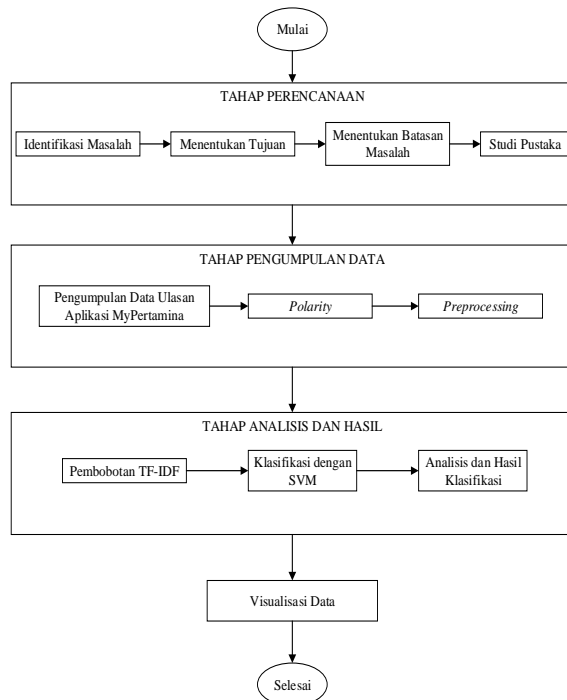
Penelitian mengenai tipe kernel svm yang dilakukan oleh feta dan ginanjar yang membandingkan fungsi kernel support vector machine mengenai penyakit tanaman kedelai, diketahui kernel terbaik yaitu RBF dengan akurasi 83% [18]. Adapun penelitian yang dilakukan Liani dkk, yang membandingkan kernel dari *support vector machine* untuk klasifikasi skripsi berdasarkan abstrak, mendapatkan hasil kernel terbaik yaitu *linear* dengan akurasi 81% [19]. Sementara itu Aulia dkk melakukan penelitian untuk membandingkan kernel dari *support vector machine* pada sentimen vaksinasi covid19, didapat kernel terbaik yaitu sigmoid dan linear dengan akurasi 87% [20].

Penelitian ini menggunakan algoritma SVM untuk melakukan analisis sentimen mengenai aplikasi mypertamina berdasarkan tiga sumber data yaitu data dari google play store, twitter dan youtube. Proses pengerjaan dalam penelitian ini menggunakan phyton, yang nantinya akan dicari kernel terbaik diantara tiga tipe kernel SVM yaitu *linear, polynomial* dan RBF serta akan didapat hasil akurasi

algoritma SVM dan analisis sentimen yang telah dikelompokkan dalam kelas negatif dan positif.

2. METODE PENELITIAN

Alur yang dibangun pada penelitian ini memiliki beberapa proses, berikut proses dari penelitian dapat dilihat pada gambar 1.



Gambar 1. Metodologi penelitian

Proses pertama dalam penelitian ini adalah mengumpulkan sumber data, pengambilan data dilakukan pada periode Juli hingga Desember 2022.

Data tersebut berupa komentar yang diambil dengan metode *crawling* data. Data yang dikumpulkan merupakan data komentar dari google play store, twitter dan youtube mengenai aplikasi mypertamina. Jumlah komentar yang dikumpulkan untuk data dari google play store sebanyak 18.000 komentar, twitter sebanyak 20.000 komentar dan youtube sebanyak 6.400 komentar yang nantinya dapat digunakan untuk data penelitian. Berikutnya memberikan label pada data yang sudah didapat untuk menentukan apakah data tersebut memiliki sentimen positif ataupun negatif. Proses pelabelan ini dilakukan dengan bantuan dari seorang pakar bahasa.

Setelah dilakukan pelabelan, data yang dikumpulkan akan dilakukan tahapan *preprocessing* yang meliputi *cleaning*, *case folding*, *tokenizing*, *stopword*, *normalisasi* dan *stemming* agar data menjadi terstruktur. Selanjutnya melakukan pembobotan pada setiap kata yang ada pada data dengan metode TF-IDF. Setelah melewati tahapan TF-IDF selanjutnya akan diterapkan algoritma SVM, dalam proses SVM terdapat beberapa kernel, tiap kernel akan memiliki akurasi yang berbeda, pada tahap ini akan dicari kernel terbaik dari tiga kernel

yaitu *linear*, *polynomial* dan RBF. Proses akhir dalam penelitian ini yaitu mengklasifikasikan masing-masing sumber data (google play store, twitter dan youtube) menggunakan SVM sehingga didapatkan hasil akurasi, sentimen dan visualisasi kata.

2.1. Text Mining

Text mining merupakan sebuah proses untuk menemukan informasi terbaru yang sebelumnya tidak diketahui [21], *text mining* dapat mengekstrak informasi dari sumber-sumber yang berbeda. Dengan *text mining* kita dapat mengetahui isi pada sekumpulan teks walaupun tidak membaca semua isi teks tersebut sehingga dapat mempersingkat waktu [10].

2.2. Term Frequency-Inverse Document Frequency (TF-IDF)

Pemberian bobot pada kata bertujuan untuk mengetahui nilai dari sebuah kata berdasarkan frekuensi kemunculan kata. *Term Frequency-Inverse Document Frequency* (TF-IDF) merupakan metode yang digunakan untuk menghitung bobot dari setiap kata yang telah diekstrak. Model pembobotan TF-IDF adalah metode yang mengintegrasikan model *term frequency* (TF) dan *inverse document frequency* (IDF) [22]. Persamaannya dapat dilihat sebagai berikut:

$$TF = 0.5 + 0.5 \frac{tf}{\max(tf)} \quad (1)$$

$$IDF = \ln \frac{N}{df} + 1 \quad (2)$$

$$TF - IDF = TF \times IDF \quad (3)$$

Keterangan :

- df = jumlah kata pada dokumen
- ln = logaritma natural
- N = total dokumen
- Tf = banyaknya data muncul dalam sebuah dokumen
- max(tf) = panjang data sebuah dokumen.

2.3. Support Vector Machine (SVM)

Vapnik memperkenalkan Support Vector Machine (SVM) untuk pertama kalinya pada tahun 1992 sebagai metode yang efisien untuk klasifikasi dalam permasalahan nonlinear [23]. Pada masalah nonlinear SVM melakukan perhitungan menggunakan mapping non linear untuk mentransformasikan training data awal ke dimensi yang lebih tinggi [24]. Dalam pemodelan klasifikasi, SVM memiliki kematangan konsep yang lebih dan jelas secara Matematis dibanding teknik klasifikasi lainnya. SVM memiliki kelebihan diantaranya untuk menemukan *hyperplane* terbaik yang memisahkan dua kelas, *hyperplane* pemisah terbaik antara kedua

class dapat ditemukan dengan mengukur margin dari *hyperplane* dan mencari titik maksimalnya [15].

3. HASIL DAN PEMBAHASAN

Jumlah data mentah yang digunakan dalam penelitian ini adalah 44.400 data komentar yang berasal dari google play store, twitter dan youtube. Data diambil dengan metode *crawling* data menggunakan bahasa pemrograman Python. Data yang telah dikumpulkan selanjutnya akan memasuki tahap *preprocessing* agar data menjadi terstruktur, kemudian dilakukan pelabelan serta pembobotan pada setiap kata menggunakan TF-IDF dan barulah dataset bisa dipakai untuk pengujian algoritma SVM untuk mencari akurasi dan kernel terbaik.

3.1. Data Awal

Data awal yang digunakan pada penelitian ini bersumber dari tiga sosial media yaitu google play store, twitter dan youtube. Data yang diambil berupa komentar dan proses pengambilan data dilakukan menggunakan bahasa pemrograman python. Sampel data awal dapat dilihat pada tabel 1.

Tabel 1. Sampel Data Awal

| Sumber | Username | Komentar |
|------------|----------------|--|
| play store | Saiful saiful | Bagus sekali mypertamina |
| | Zulvatun | Aplikasi susah...ribed juga harus di ulang ulang...#ribed |
| twitter | kabarS4 | MyPertamina cara agar subsidi tepat sasaran.! |
| | paandaaak | Ini min setelah melakukan claim di mypertamina malah ngeleg eror mulu ni aplikasi. |
| youtube | Ahmad sairi | Mantap saya setuju dengan peraturan ini, semoga cepat terlaksana |
| | sakila nurohma | Rumit kaya hidupmu pak |

3.2. Pelabelan

Pelabelan dilakukan untuk mengetahui apakah komentar tersebut bermakna positif ataupun negatif. Sebelumnya data mentah yang didapat dari hasil *crawling* dibersihkan dari data yang duplikat dan data yang tidak berhubungan dengan mypertamina, sampel hasil pelabelan terdapat pada tabel 2.

Tabel 2. Sampel Pelabelan Data

| Sumber | Komentar | Sentimen |
|------------|--|----------|
| play store | Bagus sekali mypertamina | positif |
| | Aplikasi susah...ribed juga harus di ulang ulang...#ribed | negatif |
| twitter | MyPertamina cara agar subsidi tepat sasaran.! | positif |
| | Ini min setelah melakukan claim di mypertamina malah ngeleg eror mulu ni aplikasi. | negatif |
| youtube | Mantap saya setuju dengan peraturan ini, semoga cepat terlaksana | positif |
| | Rumit kaya hidupmu pak | negatif |

3.3. Pre Processing

Preprocessing merupakan proses mengubah data mentah kedalam bentuk yang mudah dipahami, data harus melalui tahapan *preprocessing* terlebih dahulu sebelum memasuki tahapan klasifikasi dengan algoritma. Data awal yang di ambil dengan teknik *crawling* masih berupa *unsupervised* (data tidak terstruktur) maka *preprocessing* dibutuhkan agar data tersebut dapat digunakan.

3.3.1. Case Folding

Case folding digunakan untuk merubah seluruh huruf menjadi lowercase. Sampel hasil *case folding* terdapat pada tabel 3.

Tabel 3. Sampel hasil *case folding*

| Sumber | Komentar |
|------------|--|
| play store | bagus sekali mypertamina |
| | aplikasi susah...ribed juga harus di ulang ulang...#ribed |
| twitter | mypertamina cara agar subsidi tepat sasaran.! |
| | ini min setelah melakukan claim di mypertamina malah ngeleg eror mulu ni aplikasi. |
| youtube | mantap saya setuju dengan peraturan ini, semoga cepat terlaksana |
| | rumit kaya hidupmu pak |

3.3.2. Cleaning

Cleaning merupakan proses penghapusan kata-kata yang tidak diperlukan dan juga tanda baca (*delimiter*) dan simbol-simbol lainnya yang tidak penting [25]. Sampel hasil *cleaning* terdapat pada tabel 4.

Tabel 4. Sampel hasil *cleaning*

| Sumber | Komentar |
|------------|---|
| play store | bagus sekali mypertamina |
| | aplikasi susah ribed juga harus di ulang ulang ribed |
| twitter | mypertamina cara agar subsidi tepat sasaran |
| | ini min setelah melakukan claim di mypertamina malah ngeleg eror mulu ni aplikasi |
| youtube | mantap saya setuju dengan peraturan ini semoga cepat terlaksana |
| | rumit kaya hidupmu pak |

3.3.3. Tokenizing

Tahap berikutnya adalah *tokenizing* yang dilakukan untuk memecah kalimat menjadi beberapa kata. Sampel hasil *tokenizing* terdapat pada tabel 5.

Tabel 5. Sampel hasil *tokenizing*

| Sumber | Komentar |
|------------|--|
| play store | bagus, sekali, mypertamina |
| | aplikasi, susah, ribed, juga, harus, di, ulang, ulang, ribed, |
| twitter | mypertamina, cara, agar, subsidi, tepat, sasaran, ini, min, setelah, melakukan, claim, di, mypertamina, malah, ngeleg, eror, mulu, ni, aplikasi, |
| | mantap, saya, setuju, dengan, peraturan, ini, semoga, cepat, terlaksana, |
| youtube | rumit, kaya, hidupmu, pak, |

3.3.4. Stopword

Stopword digunakan untuk menghapus kata yang mengandung informasi rendah. Sampel hasil *stopword* terdapat pada tabel 6.

Tabel 6. Sampel hasil *stopword*

| Sumber | Komentar |
|------------|--|
| play store | bagus mypertamina aplikasi susah ribed ulang ulang ribed mypertamina subsidi sasaran |
| twitter | min claim mypertamina ngeleg eror mulu ni aplikasi |
| youtube | mantap setuju peraturan semoga cepat terlaksana rumit hidupmu |

3.3.5. Normalisasi

Normalisasi dilakukan untuk mengubah kalimat yang tidak baku kedalam kalimat yang baku. Sampel normalisasi terdapat pada tabel 7.

Tabel 7. Sampel hasil normalisasi

| Sumber | Komentar |
|------------|--|
| play store | bagus mypertamina aplikasi susah ribed ulang ulang ribed mypertamina subsidi sasaran |
| twitter | min claim mypertamina ngeleg eror mulu nih aplikasi |
| youtube | mantap setuju peraturan semoga cepat terlaksana rumit hidupmu |

3.3.6. Stemming

Stemming dilakukan untuk menghilangkan imbuhan pada awalan kata maupun akhiran kata sehingga kembali menjadi kata dasar. Sampel stemming terdapat pada tabel 8.

Tabel 8. Sampel hasil *stemming*

| Sumber | Komentar |
|------------|--|
| play store | bagus mypertamina aplikasi susah ribed ulang ulang ribed mypertamina subsidi sasar |
| twitter | min claim mypertamina ngeleg eror mulu nih aplikasi |
| youtube | mantap setuju atur moga cepat laksana rumit hidup |

3.4. Tf-Idf

Tahap TF-IDF dilakukan dengan menggunakan dataset yang telah dibersihkan dan diberi label pada proses sebelumnya [26]. Metode ini bertujuan untuk memberikan nilai pada kata dalam sebuah dokumen dan mencari kata-kata yang menjadi kunci dari dokumen tersebut.[27]. Hasil TF-IDF terdapat pada tabel 9.

Tabel 9. Sampel hasil tfidf

| Kata | Frekuensi | Nilai TF-IDF |
|----------|-----------|---------------------|
| aplikasi | 4647 | 0.35333051803771154 |
| susah | 4612 | 0.3719187466468923 |
| ribet | 4040 | 0.4438159288452775 |
| ulang | 3573 | 0.2696290493886085 |
| aneh | 1658 | 0.49185931788941595 |

3.5. Klasifikasi SVM

Tahap Setelah dataset melalui tahapan preprocessing, pelabelan dan TF-IDF, kemudian akan dilakukan klasifikasi dengan menerapkan model SVM. dalam proses klasifikasi SVM kombinasi parameter dilakukan dari mulai mencari tipe kernel terbaik diantara *linear*, *polinomial* dan RBF. Hasil pengujian tipe kernel SVM terdapat pada tabel 10.

Tabel 10. Hasil pengujian kernel

| | Akurasi | | |
|------------|---------|------------|-------|
| | linear | polynomial | rbf |
| Play Store | 0.943 | 0.934 | 0.946 |
| Twitter | 0.736 | 0.738 | 0.762 |
| Youtube | 0.986 | 0.989 | 0.989 |

Berdasarkan pengujian yang dilakukan dari tiga tipe kernel SVM yakni linear, polynomial dan rbf terhadap ketiga data. Akurasi terbaik yakni pada tipe kernel rbf dengan akurasi masing-masing 0.946, 0.762 dan 0.989. Setelah didapat kernel terbaik proses selanjutnya adalah melakukan validasi menggunakan perhitungan *confusion matrix*.

Tabel 11. Hasil *confusion matrix*

| Sumber data | True positive | True negative | False positive | False negative |
|-------------|---------------|---------------|----------------|----------------|
| PlayStore | 58 | 1276 | 64 | 12 |
| Twitter | 773 | 124 | 91 | 194 |
| Youtube | 0 | 1039 | 11 | 0 |

Berdasarkan Tabel 11 diketahui bahwa total ketepatan prediksi data play store adalah 1410, yang dihasilkan dari true positif = 58, true negatif 1276 dan jumlah eror sebanyak 76. Kemudian untuk data twitter total ketepatan prediksinya adalah 1200, yang dihasilkan dari true positif = 773, true negatif 124 dan jumlah eror sebanyak 285. Sedangkan untuk data youtube total ketepatan prediksinya adalah 1050, yang dihasilkan dari true positif = 0, true negatif = 1039 dan jumlah eror sebanyak 11, yang merupakan hasil kesalahan prediksi dari kelas sebenarnya, baik positif, maupun negatif.

3.6. Evaluasi

Setelah melewati proses validasi dengan algoritma SVM, evaluasi yang dilakukan dengan metode *confusion matrix* akan menghasilkan nilai *precision*, *recall*, *F1-score* dan akurasi dari SVM model. Terdapat pada tabel 12.

Tabel 12. Hasil evaluasi

| Sumber | Precision | Recall | F1-Score | Akurasi |
|------------|-----------|--------|----------|---------|
| play store | 0.95 | 0.99 | 0.97 | 0.95 |
| Twitter | 0.80 | 0.89 | 0.84 | 0.76 |
| Youtube | 0.99 | 1.00 | 0.99 | 0.99 |

Pada data play store evaluasi dilakukan terhadap 1410 data uji, 30% dari total data dengan nilai *precision* 0.95, nilai *recall* 0.99, nilai *f1-score* 0.97 dan nilai akurasi 0.95 atau 95%. Serta hasil sentimen data google play store mengenai penggunaan aplikasi mypertamina adalah negatif Untuk data twitter

evaluasi dilakukan terhadap 1200 data uji, 30% dari total data dengan nilai *precision* 0.80, nilai *recall* 0.89, nilai *f1-score* 0.84 dan nilai akurasi 0.76 atau 76%. Serta hasil sentiment data twitter mengenai penggunaan aplikasi mypertamina adalah positif. Kemudian untuk data youtube evaluasi dilakukan terhadap 1050 data uji, 30% dari total data dengan nilai *precision* 0.99, nilai *recall* 1.00, nilai *f1-score* 0.99 dan nilai akurasi 0.99 atau 99%. Serta hasil sentimen data youtube mengenai penggunaan aplikasi mypertamina adalah negatif.

3.7. Visualisasi Kata

Setelah Berikut merupakan hasil dari visualisasi kata yang memanfaatkan *wordcloud* dari *tools* google colab dengan bahasa pemrograman python yang mencakup data komentar positif dan negatif.



Gambar 2. Visualisasi kata data play store

Berdasarkan Gambar 2 dapat dilihat bahwa kata yang sering muncul pada komentar data play store yaitu kata “aplikasi”, “daftar”, “ribet”, “pertamina” dan “bahan bakar”.



Gambar 3. Visualisasi kata data twitter

Berdasarkan Gambar 3 dapat dilihat bahwa kata yang sering muncul pada komentar data twitter yaitu kata “bahan bakar”, “bakar minyak”, “mypertamina”, “aplikasi” dan “minyak subsidi”.

Berdasarkan Gambar 4 dapat dilihat bahwa kata yang sering muncul pada komentar data youtube yaitu kata “baru”, “ribet”, “atur”, “rakyat” dan “bakar minyak”.



Gambar 4. Visualisasi kata data youtube

4. DISKUSI

Pada penelitian ini dilakukan pengujian terhadap data komentar masyarakat terkait dengan aplikasi MyPertamina yang diambil dari tiga media sosial yaitu Google Play Store, Twitter dan Youtube. Pengujian dilakukan dengan algoritma Support Vector Machine, dari hasil pengujian didapat akurasi untuk data Google Play Store sebesar 95%, Twitter sebesar 76% dan Youtube sebesar 99%. Penelitian ini juga melakukan pengujian terhadap tiga kernel SVM yaitu *linear*, *polynomial* dan RBF, dari hasil pengujian menggunakan ketiga data diketahui kernel terbaik adalah RBF dengan akurasi masing-masing data sebesar 0.946, 0.762 dan 0.989.

Penelitian ini memiliki hasil yang berbeda dengan penelitian sebelumnya yang membandingkan kernel SVM, dimana pada penelitian sebelumnya yang dilakukan oleh Liani dkk didapat kernel terbaik yaitu *linear*, sedangkan pada penelitian ini didapat kernel terbaiknya adalah RBF. sumber data yang digunakan pada penelitian ini juga berbeda dengan penelitian sebelumnya serta jumlah data yang digunakan juga lebih banyak.

5. KESIMPULAN

Berdasarkan hasil dari pembahasan yang telah dilakukan, maka didapat hasil pengujian akurasi dengan menggunakan Algoritma SVM untuk data google play store adalah 95% dengan hasil sentimen negatif. Untuk data twitter didapat akurasi 76% dengan hasil sentimen positif. Sedangkan untuk data youtube mendapatkan akurasi 99% dengan hasil sentimen negatif.

Pengujian kernel yang dilakukan dari tiga tipe kernel SVM yakni *linear*, *polynomial* dan rbf terhadap ketiga data. Akurasi terbaik yakni pada tipe kernel rbf dengan akurasi masing-masing 0.946 atau 94.6% untuk data google playstore, 0.762 atau 76.2% untuk data twitter dan 0.989 atau 98.9% untuk data youtube.

Hasil dari visualisasi kata yang dilakukan dengan memanfaatkan *wordclouds* dari *tools* google colab, untuk data google playstore kata yang paling sering muncul adalah “aplikasi”, untuk data twitter adalah “bahan bakar” dan untuk data youtube adalah “baru”.

DAFTAR PUSTAKA

- [1] G. R. Hrp and N. Aslami, "Analisis Dampak Kebijakan Perubahan Publik Harga BBM terhadap Perekonomian Rakyat Indonesia," *J. Ilmu Komputer, Ekon. dan Manaj.*, vol. 2, no. 1, pp. 1464–1474, 2022.
- [2] R. I. Liperda, I. K. Hardianti, I. N. Widyah, A. Rahmadini, N. A. Fadjri, and R. R. Agustin, "Simulasi-Optimasi Sistem Transportasi Penentuan Kebutuhan Truk Tangki Pada Proses Distribusi BBM: Studi Kasus TBBM Plumpang," *JISI J. Integr. Sist. Ind.*, vol. 9, no. 2, p. 92, 2022, doi: 10.24853/jisi.9.2.92-102.
- [3] A. A. Writings, C. Hendriyani, and F. Damayanti, "MyPertamina Application To Increase Consumer Engagement," *Int. J. Bus. Rev.*, vol. 5, no. 1, pp. 65–73, 2022, [Online]. Available: <https://ejournal.upi.edu/index.php/thejobsreview/article/view/48470>
- [4] R. M. Ibrahim and N. N. K. Moeliono, "Pengaruh manfaat, kepercayaan, efikasi diri, kemudahan penggunaan, keamanan terhadap persepsi konsumen pada my pertamina (Studi pada penggunaan my pertamina kota Bandung)," *J. Ilm. Mhs. Ekon. Manaj. Accred. SINTA*, vol. 4, no. 2, pp. 396–413, 2020, [Online]. Available: <http://jim.unsyiah.ac.id/ekm>
- [5] G. S. Heryadi, "Strategi Promosi Loyalty Program My Pertamina yang Dilakukan Oleh PT Pertamina (PERSERO) Tahun 2018," *Visi Komun.*, vol. 17, no. 01, pp. 23–44, 2018.
- [6] A. A. Sinurat, C. Hendriyani, and F. Damayanti, "MyPertamina Application To Increase Consumer Engagement," *Int. J. Bus. Rev. (The Jobs Rev.)*, vol. 5, no. 1, pp. 65–74, 2022, doi: 10.17509/tjr.v5i1.48470.
- [7] A. W. Syaputri, "Analisis Sentimen Pada Ulasan Hotel Grand Elite Di Website Traveloka Menggunakan Algoritma K-Nearest Neighbor," *J. Ekon. Vol. 18, Nomor 1 Maret 201*, vol. 2, no. 1, pp. 41–49, 2020.
- [8] C. Dreisbach, T. A. Koleck, P. E. Bourne, and S. Bakken, "A systematic review of natural language processing and text mining of symptoms from electronic patient-authored text data," *Int. J. Med. Inform.*, vol. 125, no. February, pp. 37–46, 2019, doi: 10.1016/j.ijmedinf.2019.02.008.
- [9] F. Ratnawati, "Implementasi Algoritma Naive Bayes Terhadap Analisis Sentimen Opini Film Pada Twitter," *INOVTEK Polbeng - Seri Inform.*, vol. 3, no. 1, p. 50, 2018, doi: 10.35314/isi.v3i1.335.
- [10] V. A. Flores, L. Jasa, and L. Linawati, "Analisis Sentimen untuk Mengetahui Kelemahan dan Kelebihan Pesaing Bisnis Rumah Makan Berdasarkan Komentar Positif dan Negatif di Instagram," *Maj. Ilm. Teknol. Elektro*, vol. 19, no. 1, p. 49, 2020, doi: 10.24843/mite.2020.v19i01.p07.
- [11] A. M. Zuhdi, E. Utami, and S. Raharjo, "Analisis Sentiment Twitter Terhadap Capres Indonesia 2019 dengan Metode K-NN," *J. Inf. Politek. Indones. Surakarta*, vol. 5, no. 2, pp. 1–7, 2019.
- [12] E. Retnoningsih and R. Pramudita, "Mengenal Machine Learning Dengan Teknik Supervised Dan Unsupervised Learning Menggunakan Python," *Bina Insa. Ict J.*, vol. 7, no. 2, p. 156, 2020, doi: 10.51211/biict.v7i2.1422.
- [13] P. P. P. A. N. . F. I. R.H. Zer, B. H. Hayadi, and A. R. Damanik, "Pendekatan Machine Learning Menggunakan Algoritma C4.5 Berbasis Pso Dalam Analisa Pemahaman Pemrograman Website," *J. Inform. dan Tek. Elektro Terap.*, vol. 10, no. 3, 2022, doi: 10.23960/jitet.v10i3.2700.
- [14] D. A. Otchere, T. O. Arbi Ganat, R. Gholami, and S. Ridha, "Application of supervised machine learning paradigms in the prediction of petroleum reservoir properties: Comparative analysis of ANN and SVM models," *J. Pet. Sci. Eng.*, vol. 200, no. August 2020, p. 108182, 2021, doi: 10.1016/j.petrol.2020.108182.
- [15] K. A. Rokhman, B. Berlilana, and P. Arsi, "Perbandingan Metode Support Vector Machine Dan Decision Tree Untuk Analisis Sentimen Review Komentar Pada Aplikasi Transportasi Online," *J. Inf. Syst. Manag.*, vol. 3, no. 1, pp. 1–7, 2021, doi: 10.24076/joism.2021v3i1.341.
- [16] M. Ichwan, I. A. Dewi, and Z. M. S, "Klasifikasi Support Vector Machine (SVM) Untuk Menentukan TingkatKemanisan Mangga Berdasarkan Fitur Warna," *MIND J.*, vol. 3, no. 2, pp. 16–23, 2019, doi: 10.26760/mindjournal.v3i2.16-23.
- [17] D. L. Rianti, Y. Umaidah, and A. Voutama, "Tren Marketplace Berdasarkan Klasifikasi Ulasan Pelanggan Menggunakan Perbandingan Kernel Support Vector Machine," *STRING (Satuan Tulisan Ris. dan Inov. Teknol.)*, vol. 6, no. 1, p. 98, 2021, doi: 10.30998/string.v6i1.9993.
- [18] N. R. Feta and A. R. Ginanjar, "Komparasi Fungsi Kernel Metode Support Vector Machine Untuk Pemodelan Klasifikasi Terhadap Comparison of the Kernel Function of Support Vector Machine Method for Modeling Classification of Soybean Plat Disease," *J. Ilm. Ilmu Komputer, Sains dan*

- Teknol. Terap.*, vol. 1, no. 1, pp. 33–39, 2019.
- [19] A. Liani, “Analisis Perbandingan Kernel Algoritma Support Vector Machine dalam Mengklasifikasikan Skripsi Teknik Informatika berdasarkan Abstrak,” *JOINS (Journal Inf. Syst.*, vol. 5, no. 2, pp. 240–249, 2020, doi: 10.33633/joins.v5i2.3715.
- [20] T. M. Permata Aulia, N. Arifin, and R. Mayasari, “Perbandingan Kernel Support Vector Machine (Svm) Dalam Penerapan Analisis Sentimen Vaksinasi Covid-19,” *SINTECH (Science Inf. Technol. J.*, vol. 4, no. 2, pp. 139–145, 2021, doi: 10.31598/sintechjournal.v4i2.762.
- [21] M. Pejic-Bach, T. Bertoncel, M. Meško, and Ž. Krstić, “Text mining of industry 4.0 job advertisements,” *Int. J. Inf. Manage.*, vol. 50, no. July, pp. 416–431, 2020, doi: 10.1016/j.ijinfomgt.2019.07.014.
- [22] W. A. Luqyana, I. Cholissodin, and R. S. Perdana, “Analisis Sentimen Cyberbullying pada Komentar Instagram dengan Metode Klasifikasi Support Vector Machine,” *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 2, no. 11, pp. 4704–4713, 2018, [Online]. Available: <http://j-ptiik.ub.ac.id>
- [23] Y. Al-amrani, M. Lazaar, and K. Eddine, “Sentiment Analysis Using Hybrid Method of,” *J. Theor. Appl. Inf. Technol.*, vol. 96, no. 7, pp. 1886–1895, 2018, [Online]. Available: www.jatit.org
- [24] S. Sumayah, F. Sembiring, and W. Jatmiko, “Analysis of Sentiment of Indonesian Community on Metaverse Using Support Vector Machine Algorithm,” *J. Tek. Inform.*, vol. 4, no. 1, pp. 143–150, 2023.
- [25] S. Mulyani, R. Novita, P. Studi, and S. Informasi, “IMPLEMENTATION OF THE NAIVE BAYES CLASSIFIER ALGORITHM FOR CLASSIFICATION OF COMMUNITY SENTIMENT ABOUT DEPRESSION ON IMPLEMENTASI ALGORITMA NAIVE BAYES CLASSIFIER UNTUK,” vol. 3, no. 5, 2022.
- [26] B. Herwijayanti, D. E. Ratnawati, and L. Muflikhah, “Klasifikasi Berita Online dengan menggunakan Pembobotan TF-IDF dan Cosine Similarity,” *Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 2, no. 1, pp. 306–312, 2018, [Online]. Available: <https://j-ptiik.ub.ac.id/index.php/j-ptiik/article/view/796>
- [27] W. S. U. Saragih, N. A. Hasibuan, and ..., “Penerapan Text Mining Dengan Menggunakan Metode TF-IDF Untuk Menentukan Genre Dari Komik,” *KOMIK (Konferensi ...)*, vol. 4, pp. 191–199, 2020, doi: 10.30865/komik.v4i1.2679.