

ANALISIS SENTIMEN PENGGUNA DOMPET DIGITAL DANA PADA KOLOM KOMENTAR GOOGLE PLAY STORE DENGAN METODE KLASIFIKASI SUPPORT VECTOR MACHINE

Wahyu Eko Saputro, Haris Yuana, Wahyu Dwi Puspitasari
Program Studi Teknik Informatika S1, Fakultas Teknologi Informasi
Universitas Islam Balitar, Jalan Majapahit Blitar, Indonesia
wahyuekosaputro511@gmail.com

ABSTRAK

Dompot digital dana merupakan sebuah alat transaksi elektronik yang di gunakan untuk berbagai macam pembayaran, respon pro dan kontra oleh masyarakat terhadap aplikasi dompet digital DANA tersebut banyak di luncurkan di kolom komentar yang ada di google play store. Opini para pengguna aplikasi ini dapat di klasifikasikan menjadi komentar bersentimen positif dan negatif. Tujuan penelitian ini adalah untuk mengetahui sentimen positif dan negatif dari pengguna aplikasi dompet digital DANA. Maka dari itu peneliti menggunakan algoritma Support Vector Machine (SVM) dalam analisis sentimen tentang aplikasi dompet digital DANA. Hasil penelitian menunjukan bahwa sebanyak 35% pengguna aplikasi DANA memiliki sentiment positif sedangkan sebanyak 65% pengguna aplikasi DANA memiliki sentiment negatif berdasarkan pengujian klasifikasi SVM memiliki akurasi sebesar 80%, precision sebesar 84.06% untuk sentimen negatif dan 74.08% untuk sentiment positif, serta recall sebesar 87.02% untuk sentimen negatif, serta recall sebesar 69.21% untuk sentiment positif.

Kata kunci: Dompot digital, Support Vector Machine, web scraping, sentimen

1. PENDAHULUAN

Kebutuhan transaksi atau jual beli baik *offline* maupun *online* di era modern seperti ini semakin meningkat begitupun teknologi pada zaman sekarang membuat kehidupan semakin mudah dan juga di era *covid-19* ini kita rentan terkena virus baik dari tertular virus ataupun dari kurangnya prokes yang kita gunakan maka dari itu perkembangan teknologi juga berperan penting terhadap penangkalan penyebaran virus *covid-19* dengan adanya teknologi yang mengarah pada inovasi finansial salah satunya adalah alat pembayaran digital atau dompet digital. Dompot digital sendiri digunakan untuk memenuhi kebutuhan transaksi dan preferensi penggunaanya, seperti pembayaran makanan, alat transportasi, pembayaran belanja baik online maupun di toko ritel, membayar tagihan, dan lain sebagainya [1].

DANA merupakan salah satu dompet digital pendatang baru di Indonesia. Sebelum dompet digital DANA terdapat dompet digital yang lain seperti OVO, GoPay, LinkAja, dan sebagainya. Aplikasi dompet digital DANA pertama kali dihadirkan kepada publik pada tanggal 5 Desember 2018 yang didirikan oleh Elang Sejahtera Mandiri termasuk anak usahanya PT Elang Mahkota Teknologi Tbk (EMTEK) dan kongsi Bersama Ant Financial. Aplikasi DANA didesain mengikuti perkembangan zaman dimana aplikasi DANA menjadikan transaksi nontunai dan non kartu secara digital, cepat, dan praktis.

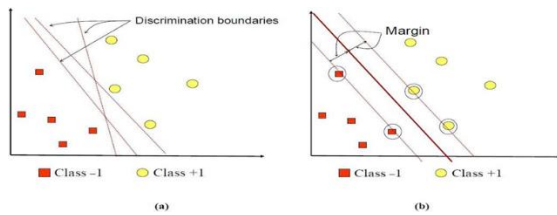
Adanya beberapa aplikasi dompet digital membuat si pengguna aplikasi memiliki tanggapan positif dan negatif baik itu dari segi fitur aplikasi dari segi keamanan kemudahan dan lain-lain terhadap DANA salah satunya pada *Google Play Store*. Maka dari itu penelitian ini akan menganalisis sentimen para

pelanggan terhadap dompet digital DANA. Analisis sentimen atau opinion mining merupakan perpaduan dari *data mining* dan *text mining*, suatu teknik untuk menganalisa pendapat, sentimen, evaluasi, sikap, penilaian dan emosi seseorang apakah penerim atau penulis berkenan dengan suatu topik, produk, layanan, organisasi, individu, ataupun kegiatan tertentu. Dengan analisis sentimen kita dapat memahami apakah isi suatu kalimat bersifat positif atau negatif. Sentimen mengacu pada satu fokus topik dimana pernyataan suatu topik tersebut mungkin akan berbeda makna.

2. TINJAUAN PUSTAKA

2.1. Support Vector Machine (SVM)

Support Vector Machine (SVM) pertama kali diperkenalkan oleh seseorang berkebangsaan Uni Soviet yang bernama Vapnik pada tahun 1992 sebagai metode *learning machine* yang bekerja dengan prinsip *Structural Risk Minimization (SRM)* yang bertujuan untuk menemukan *hyperplane* terbaik yang memisahkan dua buah *class* pada *input space* [2]. Metode ini menggunakan hipotesis berupa fungsi – fungsi linier dalam sebuah ruang fitur yang berdimensi tinggi, dengan mengimplementasikan *learning* bias berasal dari teori pembelajaran statistik. Tingkat akurasi pada model yang di hasilkan oleh proses peralihan dengan SVM sangat bergantung terhadap fungsi kernel dan parameter yang digunakan [3].



Gambar 1. Hyperplane svm terbaik

Pada prinsip dasarnya pembelajaran SVM digunakan untuk menentukan *support vector* yang hanya bergantung pada *dot product* dari data ruang fitur, yaitu ϕ_i dan ϕ_j . Pada umumnya transformasi ϕ tidak diketahui dan sangat sulit dipahami. Dengan demikian perhitungan *dot product* dapat diganti dengan fungsi lain yaitu *kernel* $K(x_i, x_j)$ yang secara tujuan untuk fungsi transformasi ϕ , itulah yang dimaksud sebagai *kernel trick*. *Kernel trick* dapat diformulasikan sebagai berikut:

$$K(x_i, x_j) = \phi_j(x_i) \cdot \phi_i(x_j) \quad [1]$$

Pada umumnya terdapat empat jenis fungsi kernel yang dapat digunakan, yaitu:

$$1. \text{ Kernel Linier } K(x, x_k) = x_k^T x \quad [2]$$

$$2. \text{ Kernel Polynomial } K(x, x_k) = x_k^T x^d + 1 \quad [3]$$

$$3. \text{ Kernel Radial Basis Function (RBF)}$$

$$4. K(x, x_k) = \exp\{-\|x - x_k\|_2^2 / \sigma^2\} \quad [4]$$

$$5. \text{ Kernel Sigmoid } K(x, x_k) = \tanh[kx_k^T x + \theta] \quad [5]$$

2.2. Dompot Digital DANA

Dana adalah penyedia layanan keuangan berbasis digital yang berkantor di Jakarta. Dana pada prinsipnya berperan sebagai pembayaran digital untuk menggantikan pembayaran secara fisik [4]. Didirikan pada tahun 2017 oleh Vincent Iswara, dengan memiliki empat lisensi Dana telah terdaftar di Bank Indonesia, berikut empat lisensi yang didapat yaitu sebagai uang elektronik, dompet digital, kirim uang, dan Likuiditas Keuangan Digital (LKD) [5].

2.3. Confusion Matrix

Confusion Matrix adalah metode yang sering digunakan untuk perhitungan *accuracy* dan *precision* [6]. Fungsi *precision* yaitu kemampuan mengevaluasi suatu sistem untuk menemukan peringkat yang relevan, dan dia artikan sebagai presentase dokumen yang di *retrieve* dan benar-benar relevan terhadap *query* [7].

Tabel 1. Tabel Confusion Matrix

Data Aktual	Prediksi	
	Negative	Positive
	Negative	True Negatif
Positive		True Positif

Berdasarkan table diatas ada beberapa istilah penilaian yang ada pada metode *Confusion Matrix*. *True Negative* merupakan data negatif yang secara benar diklasifikasikan sebagai data negatif (TN), *False*

Negative merupakan data negatif yang salah diklasifikasikan sebagai data positif (FN), *False Positive* merupakan data positif yang salah diklasifikasikan sebagai data negatif (FP) dan *True Positive* merupakan data yang benar diklasifikasikan sebagai data positif (TP). *Confusion Matrix* terdapat perhitungan yang digunakan untuk mengukur nilai akurasi, presisi berikut merupakan perhitungan untuk mendapatkan presentase penilaiannya. Akurasi merupakan jumlah proporsi yang benar [8], sedang presisi merupakan tingkat kepastian dalam perbandingan antara sampel positif dengan kelas positif [9].

$$\text{Akurasi} = \frac{TP+TN}{TP+TN+FN+FP} \quad [6]$$

$$\text{Presisi} = \frac{TP}{TP+FP} \quad [7]$$

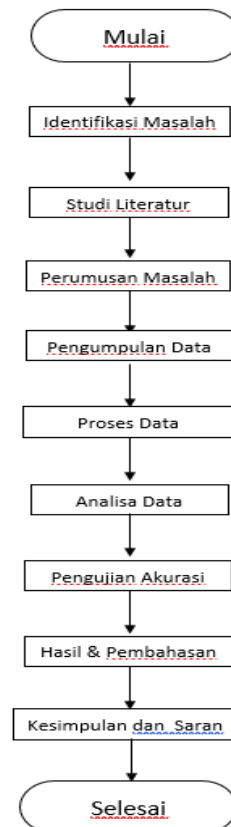
2.4. Analisis Sentimen

Analisis sentimen adalah sentimen dari teks subjektif memproses, menganalisis, meringkas dan proses inferensial. Saat ini analisis sentimen dibagi menjadi penggunaan klasifikasi pembelajaran berdasarkan aturan dan klasifikasi pembelajaran mesin [10].

3. METODE PENELITIAN

Metode penelitian pada jurnal ini melakukan sampel pada mahasiswa di Universitas Islam Balitar yang menggunakan Dompot Digital DANA.

Adapun tahapan yang dilakukan oleh peneliti adalah sebagai berikut:



Gambar 2. Tahapan penelitian

3.1. Identifikasi Masalah

Pada Proses identifikasi masalah di lakukan dengan studi masalah yang dihadapi oleh penyedia layanan dompet digital DANA tentang bagaimana opini masyarakat terhadap penyedia layanan aplikasi dompet digital DANA tersebut.

3.2. Studi Literatur

Studi Literatur berkenaan dalam pencarian dokumen yang terkait dengan permasalahan dalam penelitian yang dilakukan. Dokumen yang dicari pada sumber-sumber referensi penelitian ilmiah seperti researchgate dan sumber-sumber lain yang memiliki dokumen penelitian terkait.

3.3. Perumusan Masalah

Rumusan masalah dari proses identifikasi masalah yang sudah dilakukan, ada beberapa rumusan masalah yang didapatkan tentang bagaimana opini masyarakat terhadap perusahaan Dompet digital DANA. Beberapa rumusan masalah yang muncul yaitu: bagaimana cara mendapatkan data opini masyarakat? bagaimana analisa opini masyarakat yang didapatkan? bagaimana hasil dari proses analisa yang dilakukan?

3.4. Pengumpulan Data

Pengumpulan data didapatkan dari ulasan pengguna yang sudah diberikan kepada aplikasik terkait di *google play store*, dengan menggunakan metode *web scraping* yakni suatu kemampuan untuk pengambilan data dalam jumlah besar dari sebuah website, selanjutnya data yang sudah dipilih akan di menjadi deretan tabel (*spreadsheet*) yang di kumpulkan di dalam *microsoft excel*,

3.5. Proses Data

Berikut adalah tahapan proses data sebagai berikut:

3.5.1. Tokenization

Pada tahap *tokenization* ini suatu data yang telah terkumpul akan dipecah kedalam beberapa bagian karakter atau batas kata yang telah ditentukan, sering juga disebut dengan token karena dari proses ini dapat menjadi sebuah karakter/kata yang memiliki arti khusus.

3.5.2. Transform case

Untuk meminimalkan kesalahan pada proses *tokenize* maka dilakukan transformasi huruf yang awalnya kapital/kecil menjadi huruf besar dan sebaliknya huruf kecil menjadi huruf besar agar dalam tahap *filter tokens* terdapat keseragaman huruf dengan tujuan meminimalkan kesalahan proses *tokenize*.

3.5.3. Filter tokens (by length)

Yaitu tahap menghilangkan suatu kata/karakter dengan panjang huruf yang sudah ditentukan. Disini penulis memberikan batas yaitu minimal 2 karakter dan maksimal 25 karakter. Jadi suatu kata yang

memiliki satu karakter dan lebih dari 25 karakter akan otomatis dihilangkan.

3.5.4. Filter Stopwords

Merupakan tahap penghapusan terhadap suatu kata yang tidak memiliki arti biasanya berupa kata keterangan, kata sambung, dll. Penulis sudah menyiapkan data pendukung yaitu data stoplist yang berisi kata yang memiliki kemungkinan besar tidak memiliki arti dan tidak akan memberi pengaruh dalam tahapan perediksi.

3.5.5. Stemming

Stemming bertugas untuk menghilangkan imbuhan setiap kata sehingga suatu kata yang belum memiliki arti menjadi kata dasar, dan dalam tahap *stemming* juga memiliki tujuan untuk merapikan suatu kata dari pengejaan yang kurang tepat.

3.5.6. Labeling

Setelah data yang telah di proses sebelumnya maka selanjutnya akan masuk ke tahap pelabelan yaitu tahap untuk memberikan identitas dan akan dibagi menjadi dua kategori yaitu kata dengan ulasan positif atau kata dengan ulasan negatif.

3.6. Analisa Data

Analisis data pada penelitian ini menggunakan algoritma SVM (*support vector machine*). Pada tahap analisa data akan dibagi menjadi dua bagian yaitu data *testing* dan data *training*. Dalam tahap analisa data akan di lakukan proses pengklasifikasian kata dengan label positif dan kata dengan label negatif pada ulasan aplikasi DANA di *google play store* menggunakan algoritma SVM (*support vector machine*) untuk mendapatkan hasil tingkatan akurasi teratas.

3.7. Pengujian Akurasi

Penentuan akurasi dan menguji seberapa efektif algoritma yang di gunakan untuk klasifikasi ini di uji akurasinya dengan membandingkan kebenaran data yang telah di Analisa dengan menggunakan perhitungan Confusion Matrix.

3.8. Hasil dan Pembahasan

Hasil yang didapatkan dilakukan proses validasi dengan data yang sebelumnya sudah dilakukan pelabelan secara manual untuk mengetahui akurasi dari proses analisa yang dilakukan. Hasil dari prediksi analisa dibandingkan kebenaran datanya menggunakan perhitungan *Confusion Matrix* sehingga nilai akurasi dari proses klasifikasi yang sudah dilakukan dapat diketahui.

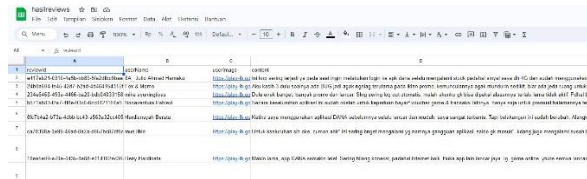
3.9. Kesimpulan dan Saran

Ditahap ini memberikan kesimpulan berdasarkan hasil penelitian yang dilakukan, hasil sentimen masyarakat ini dapat diketahui hasil dari proses analisa yang dilakukan, juga dapat diketahui bagaimana sentimen masyarakat terhadap dompet digital DANA

dan untuk kedepannya dapat dilakukan penelitian lanjutan berdasarkan penelitian yang sudah dilakukan untuk dapat lebih mendalami dan memberikan manfaat terhadap pembaca maupun pihak yang bersangkutan.

4. HASIL DAN PEMBAHASAN

Dalam pengumpulan data pada komentar *Google play store*, peneliti memerlukan pembuatan akun *google drive* untuk menyimpan data hasil dari pengumpulan data menggunakan *web scraping* dengan *google colab*. Berikut adalah hasil *scraping* yang telah tersimpan di *google drive*:



The screenshot shows a Google Drive interface. At the top, there's a search bar and navigation icons. Below, a folder named 'hasil' is expanded, showing a file named 'hasil scraping'. The file icon is a document with a green checkmark. The file size is 1.1 MB. The last modified date is 14 April 2023, 10:00 WIB. The file is shared with 'me' and '1 user'.

Gambar 3. Isi hasil *web scraping*

4.1. Text Cleaning

Pada tahap ini, data yang telah dikumpulkan akan dihapus karakter-karakter selain yang ditentukan seperti huruf atau karakter di luar dari alfabet a-z (termasuk tanda baca). Hasil *cleaning* menggunakan *Rapidminer* sebagai berikut.

Tabel 2. Hasil dari *text cleaning*

No	Ulasan	Text Cleaning
1.	Ini knp sering terjadi ya pada saat ingin melakukan login ke apk dana selalu mengalami stuck padahal sinyal saya dh 4g dan sudah menggunakan wifi pun ttp tidak bisa jadi saia tiap ingin login harus apus dan buka kembali untuk membuka aplikasi dananya mohon untuk segera diperbaiki bug bug seperti ini agar pelayanan lebih lancar 🙏🙏. Terimakasih.	Ini knp sering terjadi ya pada saat ingin melakukan login ke apk dana selalu mengalami stuck padahal sinyal saya dh 4g dan sudah menggunakan wifi pun ttp tidak bisa jadi saia tiap ingin login harus apus dan buka kembali untuk membuka aplikasi dananya mohon untuk segera diperbaiki bug bug seperti ini agar pelayanan lebih lancar Terimakasih

4.2. Tokenizing

Pada tahap ini, data yang telah melalui *cleaning* diubah dari deretan kalimat menjadi kata per kata. Contoh penerapan *tokenizing* pada *Rapidminer* seperti berikut.

Tabel 3. Hasil dari *tokenizing*

No	Ulasan	Tokenizing
1.	Ini knp sering terjadi ya pada saat ingin melakukan login ke apk dana selalu mengalami	“Ini”, “knp”, “sering”, “terjadi”, “ya”, “pada”, “saat”, “ingin”, “melakukan”, “login”, “ke”, “apk”, “dana”,

No	Ulasan	Tokenizing
	stuck padahal sinyal saya dh 4g dan sudah menggunakan wifi pun ttp tidak bisa jadi saia tiap ingin login harus apus dan buka kembali untuk membuka aplikasi dananya mohon untuk segera diperbaiki bug bug seperti ini agar pelayanan lebih lancar 🙏🙏. Terimakasih.	“selalu”, “mengalami”, “stuck”, “padahal”, “sinyal”, “saya”, “dh”, “4g”, “dan”, “sudah”, “menggunakan”, “wifi”, “pun”, “ttp”, “tidak”, “bisa”, “jadi”, “saia”, “tiap”, “ingin”, “login”, “harus”, “apus”, “dan”, “buka”, “kembali”, “untuk”, “membuka”, “aplikasi”, “dananya”, “mohon”, “untuk”, “segera”, “diperbaiki”, “bug”, “bug”, “seperti”, “ini”, “agar”, “pelayanan”, “lebih”, “lancar”, “terimakasih”

4.3. Case Folding

Data yang sudah melalui proses *cleaning* sebelumnya selanjutnya masuk ke tahap *case folding*, di mana data melalui proses penyamaan huruf pada dokumen ulasan dengan mengubah huruf kapital menjadi huruf kecil. Tabel di bawah ini merupakan contoh data yang telah melalui tahap *case folding* di *Rapidminer*.

Tabel 4. Hasil dari *case folding*

No	Ulasan	Case Folding
1.	“Ini”, “knp”, “sering”, “terjadi”, “ya”, “pada”, “saat”, “ingin”, “melakukan”, “login”, “ke”, “apk”, “dana”, “selalu”, “mengalami”, “stuck”, “padahal”, “sinyal”, “saya”, “dh”, “4g”, “dan”, “sudah”, “menggunakan”, “wifi”, “pun”, “ttp”, “tidak”, “bisa”, “jadi”, “saia”, “tiap”, “ingin”, “login”, “harus”, “apus”, “dan”, “buka”, “kembali”, “untuk”, “membuka”, “aplikasi”, “dananya”, “mohon”, “untuk”, “segera”, “diperbaiki”, “bug”, “bug”, “seperti”, “ini”, “agar”, “pelayanan”, “lebih”, “lancar”, “terimakasih”	“ini”, “knp”, “sering”, “terjadi”, “ya”, “pada”, “saat”, “ingin”, “melakukan”, “login”, “ke”, “apk”, “dana”, “selalu”, “mengalami”, “stuck”, “padahal”, “sinyal”, “saya”, “dh”, “4g”, “dan”, “sudah”, “menggunakan”, “wifi”, “pun”, “ttp”, “tidak”, “bisa”, “jadi”, “saia”, “tiap”, “ingin”, “login”, “harus”, “apus”, “dan”, “buka”, “kembali”, “untuk”, “membuka”, “aplikasi”, “dananya”, “mohon”, “untuk”, “segera”, “diperbaiki”, “bug”, “bug”, “seperti”, “ini”, “agar”, “pelayanan”, “lebih”, “lancar”, “terimakasih”

4.4. Filter Token

Pada tahap ini, data yang telah melalui tahap *case folding* dipilih dengan jumlah kata minimal 4 huruf

pada setiap kata. Hasil filter token menggunakan *Rapidminer* sebagai berikut.

Tabel 5. Hasil dari *filter token*

No	Ulasan	Filter Token
1.	"ini", "knp", "sering", "terjadi", "ya", "pada", "saat", "ingin", "melakukan", "login", "ke", "apk", "dana", "selalu", "mengalami", "stuck", "padahal", "sinyal", "saya", "dh", "4g", "dan", "sudah", "menggunakan", "wifi", "pun", "ttp", "tidak", "bisa", "jadi", "saia", "tiap", "ingin", "login", "harus", "apus", "dan", "buka", "kembali", "untuk", "membuka", "aplikasi", "dananya", "mohon", "untuk", "segera", "diperbaiki", "bug", "bug", "seperti", "ini", "agar", "pelayanan", "lebih", "lancar", "terimakasih"	"sering", "terjadi", "pada", "saat", "ingin", "melakukan", "login", "dana", "selalu", "mengalami", "stuck", "padahal", "sinyal", "saya", "sudah", "menggunakan", "wifi", "tidak", "bisa", "jadi", "saia", "tiap", "ingin", "login", "harus", "apus", "buka", "kembali", "untuk", "membuka", "aplikasi", "dananya", "mohon", "untuk", "segera", "diperbaiki", "seperti", "agar", "pelayanan", "lebih", "lancar", "terimakasih"

4.5. Stopword Removal

Tahap selanjutnya pada *preprocessing* yaitu menghapus kata-kata yang tidak berkaitan dengan isi ulasan Proses *stopword removal* pada tahap ini, perlu di tambahkan sebuah *library* kata Bahasa indonesia (sastrawi), karena di *rapidminer* belum terdapat *library* Bahasa indonesia, Contoh penerapan *stopword removal* dari hasil *stemming* sebagai berikut.

Tabel 6. Hasil dari *stopword removal*

No	Content	Stopword Removal
1.	"sering", "terjadi", "pada", "saat", "ingin", "melakukan", "login", "dana", "selalu", "mengalami", "stuck", "padahal", "sinyal", "saya", "sudah", "menggunakan", "wifi", "tidak", "bisa", "jadi", "saia", "tiap", "ingin", "login", "harus", "apus", "buka", "kembali", "untuk", "membuka", "aplikasi", "dananya", "mohon", "untuk",	"login", "dana", "mengalami", "stuck", "sinyal", "wifi", "login", "apus", "buka", "membuka", "aplikasi", "dananya", "mohon", "segera", "diperbaiki", "pelayanan", "lancar", "terimakasih"

	"segera", "diperbaiki", "seperti", "agar", "pelayanan", "lebih", "lancar", "terimakasih"	
--	---	--

4.6. Stemming

Data hasil proses *cleaning* selanjutnya diolah menggunakan proses *stemming* untuk mengetahui kata dasar yang terdapat di dalam sebuah ulasan. Proses *stemming* ini perlu di tambahkan sebuah *library* kata Bahasa indonesia (sastrawi), karena di *rapidminer* belum terdapat *library* Bahasa indonesia, Contoh penerapan *stemming* pada data hasil *cleaning* sebagai berikut.

Tabel 6. Hasil dari *stemming*

No	Ulasan	Stemming
1.	"login", "dana", "mengalami", "stuck", "sinyal", "wifi", "login", "apus", "buka", "membuka", "aplikasi", "dananya", "mohon", "segera", "diperbaiki", "pelayanan", "lancar", "terimakasih"	"login", "dana", "alami", "stuck", "sinyal", "wifi", "login", "apus", "buka", "buka", "aplikasi", "dana", "mohon", "segera", "baik", "layan", "lancar", "terimakasih"

Setelah melewati tahap – tahap seperti *text cleaning*, *tokenizing*, *case folding*, *filter token*, *stopword removal*, *stemming* maka data awal yang berjumlah 1200 kata menjadi 1089 kata.

4.7. Proses Ekstraksi Fitur Menggunakan TF-IDF

Pada tahap ini, data yang telah dimasukkan kedalam proses training dan testing akan diekstraksi menggunakan TF-IDF. Proses ini digunakan untuk mengetahui bobot suatu term (kata) dalam sebuah dokumen (ulasan). Gambar di bawah ini menunjukkan hasil TF-IDF sebagai berikut.

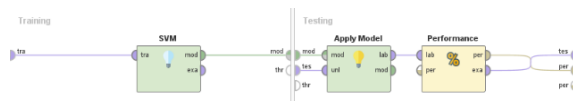
Tabel 7. Hasil proses TF-IDF

indeks_ulasan	indeks_token	tf-idf
0	35	0,9139
0	41	0,3582
0	125	0,1771
0	164	0,5177
0	165	0,3667
0	209	0,1771
0	441	0,1837

4.8. Pemodelan Data Menggunakan Support Vector Machine (SVM)

Dalam tahap pemodelan ini juga dilakukan *k-fold cross validation* dengan *k=10* dan *kernel rbf*. Nilai *k* tersebut menunjukkan bahwa data akan dibagi menjadi 10 subset dengan ukuran yang sama dan data yang berbeda. Pada setiap iterasi, satu bagian digunakan untuk pengujian dan sisanya untuk pelatihan. Berikut

bentuk *k-fold cross validation* dari *Support Vector Machine*.



Gambar 4. *Cross validation svm*

4.9. Evaluasi Klasifikasi SVM

Setelah data melalui tahap *modeling*, proses data selanjutnya akan masuk pada tahap evaluasi atas hasil dari tahap *modeling* dengan menggunakan *confusion matrix*. Disini penulis akan membandingkan tingkat akurasi dari hasil *modeling*. Dalam menganalisis akurasi, penulis juga menampilkan indikator *True Positive* (TP), *True Negative* (TN), *False Positive* (FP), dan *False Negative* (FN). Indikator *True Positive* (TP) berada di kelas kata positif yang dikelompokkan sebagai kelas positif, *True Negative* (TN) berada di kelas negatif yang dikelompokkan sebagai kelas negatif, *False Positive* (FP) berada di kelas positif yang dikelompokkan ke dalam kelas negatif, *False Negative* (FN) berada di kelas negatif yang dikelompokkan ke dalam kelas positif.

Dalam pengolahan data menggunakan *Support Vector Machine* (SVM) menghasilkan tingkat *accuracy* dari model dengan menggunakan *tool rapidminer* seperti gambar di bawah ini.

accuracy: 80.81% +/- 3.00% (micro average: 80.81%)

	true negatif	true positif	class precision
pred. negatif	617	117	84.06%
pred. positif	92	263	74.08%
class recall	87.02%	69.21%	

Gambar 5. Tingkat *accuracy svm*

Dapat disimpulkan dari gambar di atas bahwa tingkat akurasi yang didapatkan dari proses analisis pengklasifikasian menggunakan metode *Support Vector Machine* (SVM) adalah 80,81% dengan jumlah *True Positive* 263 record data yang diklasifikasikan sebagai positif, *False Positive* 117 record data dikelompokkan sebagai negatif, *True Negative* 617 record data dikelompokkan sebagai negatif, dan *False Negative* 92 record data diklasifikasikan sebagai positif.

5. KESIMPULAN DAN SARAN

Kesimpulan dari penelitian ini yaitu berdasarkan data yang didapat algoritma SVM cukup akurat untuk mengklasifikasikan komentar positif dan negatif dari aplikasi DANA di *google play store* dengan skor akurasi positif sebesar 80.81% dan skor akurasi negatif sebesar 84.06%, saran yang dapat peneliti sampaikan untuk penelitian selanjutnya adalah pengembangan pada proses *stemming* dan *stopword removal* dengan menambahkan *keyword* yang cukup bervariasi dari

KBBI sehingga mendapatkan hasil yang lebih valid dan akurat.

DAFTAR PUSTAKA

- [1] Panday, R., 2020. Dampak COVID19 Pada Kesiapan Teknologi dan Penerimaan Teknologi di Kampus. *Jurnal Kajian Ilmiah*, 1(1).
- [2] Isnain, A.R., Sakti, A.I., Alita, D. and Marga, N.S., 2021. Sentimen Analisis Publik Terhadap Kebijakan Lockdown Pemerintah Jakarta Menggunakan Algoritma Svm. *Jurnal Data Mining Dan Sistem Informasi*, 2(1), pp.31-37.
- [3] Siagian, R.Y., 2011. Klasifikasi Parket Kayu Jati Menggunakan Metode Support Vector Machine (SVM). Skripsi, Jurusan Teknik Informatika, Fakultas Teknologi Industri, Universitas, Jawa Barat.
- [4] Pratiwi, D.S. and Nuryana, I.K.D., 2021. Analisis Tingkat Penerimaan dan Kepercayaan Pengguna Teknologi Terhadap Penggunaan Dompot Digital DANA. *Journal of Emerging Information System and Business Intelligence (JEISBI)*, 2(4), pp.76-84.
- [5] Arini, A.C., Sulistyarini, S. and Harjanti, D.T., PENGGUNAAN LAHAN GAMBUT UNTUK PERTANIAN BERBASIS KEARIFAN LOKAL DI DESA RASAU JAYA DUA. *Jurnal Pendidikan dan Pembelajaran Khatulistiwa (JPPK)*, 10(11).
- [6] Normawati, D. and Prayogi, S.A., 2021. Implementasi Naïve Bayes Classifier Dan Confusion Matrix Pada Analisis Sentimen Berbasis Teks Pada Twitter. *J-SAKTI (Jurnal Sains Komputer Dan Informatika)*, 5(2), pp.697-711.
- [7] Sutanto, E.M., 2020. Analisis pemilihan aplikasi pembayaran DANA (Doctoral dissertation, Petra Christian University).
- [8] Andika, L. A., Azizah, P. A. N. and Respatiwan, R. (2019) "Analisis Sentimen Masyarakat terhadap Hasil Quick Count Pemilihan Presiden Indonesia 2019 pada Media Sosial Twitter Menggunakan Metode Naive Bayes Classifier," *Indonesian Journal of Applied Statistics*, 2(1), p. 34. doi: 10.13057/ijas.v2i1.29998.
- [9] View of identifikasi konten kasar Pada tweet Bahasa Indonesia (no date) Inacl.id. Available at: <http://inacl.id/journal/index.php/jlk/article/view/15/16> (Accessed: July 19, 2023).
- [10] Isnain, A.R., Sakti, A.I., Alita, D. and Marga, N.S., 2021. Sentimen Analisis Publik Terhadap Kebijakan Lockdown Pemerintah Jakarta Menggunakan Algoritma Svm. *Jurnal Data Mining Dan Sistem Informasi*, 2(1), pp.31-37.