

KOMPARASI *PAYMENT DIGITAL* UNTUK ANALISIS SENTIMEN BERDASARKAN ULASAN DI *GOOGLE PLAYSTORE* MENGGUNAKAN METODE *SUPPORT VECTOR MACHINE*

¹⁾Renaldi AL Anshari, ²⁾Syariful Alam, ³⁾Moch Hafid T

^{1,2,3)}Sekolah Tinggi Teknologi Wastukencana (Program Studi, Teknik Informatika, Sekolah Tinggi Teknologi Wastukencana Purwakarta)

¹⁾renaldial75@wastukencana.ac.id, ²⁾syarifulalam@wastukencana.ac.id,

³⁾mhafid@wastukencana.ac.id

INFO ARTIKEL

Riwayat Artikel :

Diterima : 27 Juli 2023

Disetujui : 29 Juli 2023

Kata Kunci :

Google play store, Support Vector Machine, Analisis Sentimen, Dana, Ovo

ABSTRAK

Pada era modern saat ini semuanya menjadi sangat mudah dan praktis tentunya dengan bantuan perkembangan teknologi aktivitas manusia menjadi sangat terbantu sekali. salah satunya dalam bidang keuangan atau biasa disebut *finance*. seperti yang sudah di ketahui saat ini banyak produk keuangan yang di gabungkan dengan kemampuan teknologi atau sering juga di sebut dengan *fintech* (*finansial technology*). salah satu produk dari fintech adalah payment digital. Sederhananya digital payment ini merupakan metode pembayaran yang memanfaatkan teknologi jaringan internet dan sejumlah perangkat untuk mendapatkan benefit kepraktisan, efesiensi dan kemudahan. Salah satu aplikasi *digital payment* yang digunakan dalam penelitian ini adalah Dana dan Ovo, data yang diambil ini adalah data ulasan pengguna dari google playstore dengan cara *scraping* menggunakan *API google play scraper*. Tujuan dari penelitian ini utnuk mengetahui ulasan pengguna dan membandingkan mana aplikasi yang mendapatkan banyak komentar positif dan negatif menggunakan algoritma *Support Vector Machine*, dari masing-masing aplikasi diambil 6000 data, dan untuk hasil akhirnya aplikasi dana yang mendapat nilai tertinggi berdasarkan nilai akurasi 0,92, presisi 0,92 dan nilai *recall* 100%. Dengan hal itu aplikasi dana cenderung mendapatkan sentiment positif dan negatif lebih banyak di masyarakat berdasarkan ulasan pengguna.

ARTICLE INFO

Article History :

Received : July 27, 2023

Accepted : July 29, 2023

Keywords:

Google play store, Support Vector Machine, Sentiment Analysis, Dana, Ovo

ABSTRACT

In today's modern era, everything has become very easy and practical, of course, with the development of technological assistance, human activities have become very helpful. One of them is in the field of finance or commonly called finance. As is well known, many financial products are combined with technological capabilities or often referred to as fintech (financial technology). one of the products from fintech is digital payments. Simply put, this digital payment is a payment method that utilizes internet network technology and a number of devices to gain practicality, efficiency and convenience. One of the digital payment applications used in this study is Dana and Ovo, the data taken is a review of user data from Google Playstore by scraping using the Google Play Scraper API. The purpose of this study is to find out user reviews and compare which applications get a lot of positive and negative comments using the Support Vector Machine algorithm, from each application as much as

6000 data is taken, and for the final result the fund application gets the highest score based on an accuracy value of 0.92, precision of 0.92 and a recall value of 100%. With that, fund applications tend to get more positive and negative sentiments in the community based on user reviews.

1. PENDAHULUAN

Pada era modern saat ini semuanya menjadi sangat mudah dan praktis tentunya dengan bantuan perkembangan teknologi aktivitas manusia menjadi sangat terbantu sekali. salah satunya dalam bidang keuangan atau biasa disebut *finance*. seperti yang sudah di ketahui saat ini banyak produk keuangan yang di gabungkan dengan kemampuan teknologi atau sering juga di sebut dengan *fintech (finansial technology)*. *fintech* adalah sebuah perusahaan yang menggabungkan jasa keuangan dengan teknologi, *fintech* juga merupakan istilah untuk menyebut sebuah inovasi teknologi dan digitalisasi pada layanan finansial. salah satu produk dari *fintech* adalah *digital payment*.(Technology et al., 2020).

Sederhananya *digital payment* ini merupakan metode pembayaran yang memanfaatkan teknologi jaringan internet dan sejumlah perangkat untuk mendapatkan benefit kepraktisan, efisiensi dan kemudahan. (Mubarak and Akhmadi, no date)

untuk aplikasi dari digital payment ini sangat banyak diantaranya adalah dana dan ovo yang sedang trending belakangan ini. Dana adalah sistem dan layanan pembayaran bergerak berupa uang elektronik, *e-wallet*, transfer uang, dan layanan pendukung lainnya yang dapat diakses melalui perangkat telekomunikasi. Sedangkan ovo adalah aplikasi untuk penyedia sistem pembayaran yang memungkinkan pembayaran tanpa batas dan menyediakan akses ke produk dan layanan keuangan digital lainnya yang diluncurkan bekerja sama dengan mitra terpilih. (Hidayat, Hanggara and Prakoso, 2021).

Dalam hal ini baik dana dan ovo memiliki kendalanya masing-masing di masyarakat seperti kebijakan dan biaya, ketidak sesuaian fitur atau layanan, gangguan dan kegagalan sistem yang menjadi sentimen negatif di masyarakat. Dari masalah tersebut penelitian ini bertujuan untuk bagaimana mengetahui perbedaan sentimen masyarakat terhadap

kelayakan aplikasi, cara bertransaksi pembayaran, fitur yang ada dalam aplikasi dan juga keamanan dan privasi aplikasi tersebut lalu masyarakat bisa memberikan opini negatif ataupun positif untuk aplikasi dana dan ovo. *Analisis sentimen* dan *opinion mining* adalah cabang bidang ilmu study yang menganalisa opini publik, sentimen, evaluasi, sikap dan emosi yang diterapkan hampir oleh semua pelaku bisnis dan sosial. Alasan sentimen ini digunakan karena opini merupakan hal penting di segala kegiatan manusia dan merupakan salah satu hal yang mempengaruhi kebiasaan masyarakat. kedepannya analisis sentimen tersebut akan menadi acuan untuk perusahaan meningkatkan pelayanannya.

Pada penelitian ini, data yang diambil dari *google playstore* dengan menggunakan *API google play scraper*. Untuk langkah awal dalam menganalisis data yaitu memisahkan sentimen negatif dan positif berdasarkan ulasan komentar dari pengguna tersebut, selanjutnya membandingkan kedua aplikasi tersebut untuk mengetahui tingkatan dari aplikasi tersebut berdasarkan nilai akurasi.

Penelitian sebelumnya dengan judul Perbandingan *analisis sentimen* terhadap *digital payment* Go-pay dan Ovo di media sosial *twitter* menggunakan algoritma *naive bayes* dan *wordcloud*. Pada aplikasi Gopay Penelitian tersebut mendapatkan akurasi model 84,82% untuk *tweet* positif hingga 214 *tweet*. Namun, pembelajaran mesin sampai pada kesimpulan yang berbeda: 119 *tweet* dianggap benar positif dan 95 *tweet* negatif, dengan representasi sebenarnya dari akurasi masing-masing 55,61 dan 50,21 persen, sedangkan jumlah *tweet* yang dianggap negatif adalah 1189 *tweet*. Namun, pembelajaran mesin menyimpulkan bahwa *tweet* negatif pun diberi peringkat 1171. Dengan kinerja mengingat 90,04% recall dan akurasi 91,85%. Pada aplikasi Egg, penelitian ini mencapai akurasi model sebesar 82,81%. Jumlah *tweet* yang dinilai positif sebanyak 198 *tweet*. Namun, pembelajaran mesin menyimpulkan

bahwa 99 tweet benar-benar positif dan 99 tweet dianggap negatif. Dengan performa recall 50,00, presisinya 24,63%. Sementara itu, jumlah tweet yang tergolong negatif sebanyak 2141 tweet. Namun, pembelajaran mesin menyimpulkan bahwa hingga 1838 tweet sebenarnya negatif dan 303 tweet dianggap positif. Performa mengingat dengan akurasi 85,85 dan precision sebesar 94,89%. (Agrani, 2020).

Penelitian lainnya dengan judul *Analisis sentimen* pengguna untuk aplikasi Google Maps, Maps.Me dan Waze menggunakan metode *Support Vector Machine*. Data yang berhasil diambil oleh penelitian ini adalah 750 data review menggunakan Google Playscraper, setelah itu data tersebut diberi label sentimen pada masing-masing aplikasi sebanyak 300 data untuk dijadikan data training. Hasil dari aplikasi google maps mendapatkan akurasi model sebesar 89,73%, presisi 100% dan untuk nilai recall sebesar 24,45%. Hasil dari aplikasi Maps.me mendapatkan nilai akurasi model sebesar 86,40%, presisi 86,55% dan untuk nilai recal sebesar 99,69%. Dan yang terakhir hasil dari aplikasi waze mendapatkan nilai akurasi sebesar 79,47%, presisi 81,69 dan untuk nilai recall 80,07%. (Fariz *et al.*, 2023)

Penelitian selanjutnya ini dilakukan oleh Anggi Puji Astuti, Syariful Alam dan Irsan Jaelani dengan judul *Komparasi algoritma Support Vector Machine* dengan Naïve Bayes untuk Analisis Sentimen Pada Aplikasi BRImo. Hasil dari penelitian ini untuk mengetahui perbandingan akurasi dari dua algoritma yaitu *Support Vector Machine* dan Naïve Bayes. Data yang digunakan yaitu 5000 data ulasan hasil scraping. Data tersebut di klasifikasikan dengan 5 kali percobaan menggunakan algoritma *Support Vector Machine* yang menghasilkan nilai akurasi sebesar 97,56%, sedangkan untuk algoritma Naïve bayes menghasilkan nilai akurasi sebesar 96,52% serta evaluasi data ini menggunakan confusion matriks. Sehingga dapat disimpulkan algoritma *Support Vector Machine* ini memiliki nilai lebih baik dari pada algoritma Naïve bayes untuk klasifikasi data ulasan aplikasi mobile banking BRImo. (Astuti, Alam and Jaelani, 2023)

Berdasarkan informasi diatas tujuan untuk penelitian ini yaitu membandingkan aplikasi payment digital Dana dan Ovo berdasarkan

ulasan di *google playstore* menggunakan metode *suport vector machine*.

klasifikasi *Support Vector Machine*, tahap ini adalah tahap paling penting yaitu menentukan suatu akurasi model dari data tersebut menggunakan *python google collaboratory*. Lalu setelah itu melakukan tahap evaluasi yang menggunakan *Confussion matriks* jika tahap evaluasi sudah maka yang terakhir adalah tahap visualisasi yang menggunakan *worldcloud* untuk memunculkan kata yang sering digunakan baik negatif maupun positif.

2. METODE

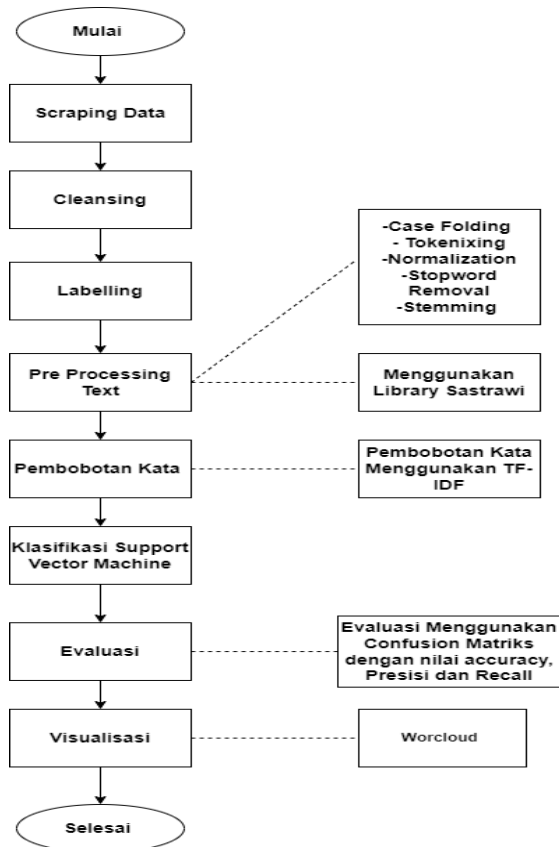
1. Teknik Pengumpulan Data

Metode pengambilan data menggunakan API google play scraper, tahap ini adalah tahap awal dimana analisis sentimen ini dilakukan. Data yang diambil pada penelitian ini berdasarkan ulasan komentar di google playstore dengan bantuan tools google colaboratory yang menggunakan Bahasa pemrograman python sebagai media pengambilan data tersebut. Cukup memasukan link dari aplikasi yang akan kita ambil data nya dan berapa jumlah data yang akan di ambil maka google colaboratory akan merespon dan mengambil data tersebut. Selanjutnya melakukan tahap pre processing data dimana data diolah melalui tahap-tahap tertentu agar menjadi data yang dibutuhkan lalu setelah sudah di olah data akan memasuki tahap pembobotan kata TF-IDF. Pada dasarnya TF-IDF ini bekerja dalam menentukan frekuensi relative suatu kata kemudian kata tersebut di bandingkan dengan proporsi kata lainnya pada seluruh dokumen. Selanjutnya yaitu tahap klasifikasi *Support Vector Machine*, tahap ini adalah tahap paling penting yaitu menentukan suatu akurasi model dari data tersebut menggunakan *python google collaboratory*. Lalu setelah itu melakukan tahap evaluasi yang menggunakan *Confussion matriks* jika tahap evaluasi sudah maka yang terakhir adalah tahap visualisasi yang menggunakan *worldcloud* untuk memunculkan kata yang sering digunakan baik negatif maupun positif.

2. Kerangka Pemikiran

Kerangka pemikiran ini adalah jalur penelitian yang di rancang berdasarkan kegiatan peneliti

lakukan. Dari kerangka pikir ini di bentuk suatu alur penelitian yang terjadi berisikan hubungan *variable* bebas dan *variable* terikat dalam rangka memberikan jawaban sementara untuk penelitian.(Penggunaan *et al.*, 2017)
 Rancangan penelitian ini disajikan dalam bentuk kerangka pemikiran pada gambar 1 dibawah ini.



Gambar 1. Kerangka Pemikiran

3.Scraping Data

Pada tahap *scraping* ini data diambil dari *google playstore* dengan menggunakan *API google play scraper* dengan Bahasa pemrograman *pyhton* dengan bantuan tools *google colaboratory*. Data yang diambil hanya beberapa data yang di butuhkan saja seperti *Username, score/rating, waktu* dan komentar dari pengguna lalu data yang sudah diambil tersebut akan disimpan ke dalam bentuk *file excel* dengan format *csv*.

4.Cleansing

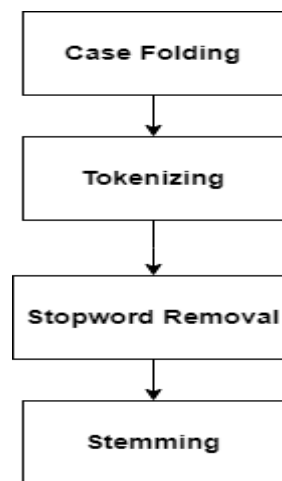
Tahap *Cleansing* ini merupakan tahap penghapusan data duplikat dan data yang tidak sesuai dengan komentar aplikasi Dana dan Ovo.

5.Labelling

Pada tahap *labelling* ini menggunakan aplikasi *Microsoft excel* dengan cara manual yaitu melihat isi komentar/content tersebut satu persatu atau terkadang menyesuaikan dengan *rating* yang ada dalam aplikasi tersebut.

5.Preprocessing Text

Tahap berikutnya ini adalah *Preprocessing text*. Pada langkah ini data diolah menjadi informasi yang dapat digunakan pada langkah selanjutnya sesuai dengan format yang ada. *Text preprocessing* dibagi menjadi 5 langkah, yang dapat dilihat pada Gambar 2 berikut.



Gambar 2. Preprocessing Text

1. Case Folding

Tahap ini adalah tahap penyamarataan huruf didalam sebuah dokumen. Untuk diketahui lebih lanjutnya tidak semua dokumen yang ada dalam teks tersebut selalu selaras menggunakan huruf kapital jadi diperlukan case folding ini untuk penyamarataan sebuah huruf.

2. Tokenizing

Tahapan *tokenizing* ini berfungsi untuk memisahkan kata per kata sebuah huruf dalam sebuah dokumen.

3. Stopword Removal

Stopword removal ini adalah tahapan untuk menghilangkan kata yang tidak di perlukan.

4. Stemming

Tahap terakhir yaitu *Stemming*, tahap ini bertujuan untuk mengubah kata yang menerima sufiks/imbunan dan pindahkan pilihan ke kata dasar. (Sentimen *et al.*, 2021).

6. Pembobotan Kata (TF-IDF)

Setelah melewati tahap *preprocessing text*, selanjutnya adalah pembobotan kata TF-IDF. Pada dasarnya, TF-IDF bekerja dengan menentukan *frekuensi relatif* suatu kata dan kemudian membandingkannya dengan proporsi kata tersebut di seluruh dokumen. (Khatib *et al.*, 2022) Semakin banyak kata yang dikandung dokumen, semakin besar nilai atau bobot kata tersebut.

7. Klasifikasi Support Vector Machine

Tahap selanjutnya setelah pembobotan kata TF-IDF yaitu Klasifikasi *Support Vector Machine*. *Support Vector Machine* ini adalah algoritma pembelajaran mesin yang diawasi dan dapat digunakan untuk klasifikasi maupun regresi. *Support Vector Machine* ini digunakan untuk mencari hyperlane terbaik dengan memaksimalkan jarak antar kelas. (Noviana *et al.*, 2023) Hyperlanes itu sendiri adalah fungsi dari pemisahan kelas. Fungsi yang digunakan untuk mengklasifikasikan antar kelas dalam 2D disebut garis sedangkan dalam 3D fungsi yang digunakan untuk mengklasifikasikan antar kelas juga disebut bidang sedangkan fungsi yang digunakan untuk mengklasifikasikan dalam kelas berdimensi lebih tinggi disebut hyperlane (Ditami, Ripanti and Sujaini, 2022).

8. Confusion matriks

Tahap berikutnya adalah evaluasi, tahap evaluasi ini menggunakan *confusion matriks* Matriks konfusi adalah tabel yang digunakan untuk mengukur performa model pembelajaran mesin. Tabel ini menjelaskan informasi tentang data yang diklasifikasikan sebagai benar atau salah. Adapun pengertian lain dalam *confusion matriks* ini adalah Salah satu tools/alat untuk menunjukkan nilai sebenarnya atau *true value* dan membandingkannya dengan nilai prediksi model yang dapat digunakan untuk menghasilkan matriks evaluasi seperti *Accuracy* (Akurasi), *Precision*, *Recal* dan *F measure*. (Himawan, 2021). Ada 4 nilai yang dihasilkan dalam tabel confusion matriks ini diantaranya *True Positif (TP)*, *False Positif (FP)*, *False Negatif (FN)*, dan *True Negatiive (TN)*. Untuk ilustrasi *Confusion matriks* ini dapat dilihat pada tabel 1 dibawah ini.

Tabel 1. *Confusion Matriks*

		Nilai Aktual	
		Positif	Negatif
Nilai Prediksi	Positif	TP	FP
	Negatif	FN	TN

True Positif (TP) : Jumlah data yang positif tetapi diharapkan positif

False Positif (FP) : Himpunan data yang negatif tetapi diharapkan positif

False Negatif (FN) : Jumlah data yang positif tetapi diharapkan positif

True Negatif (TN) : Jumlah data yang bernilai negative dan diprediksi benar sebagai negative.

Visualisasi

Tahap terakhir ini adalah Visualisasi, dalam visualisasi ini peneliti menggunakan wordcloud untuk menampilkan gambar seberapa banyak nya kata positif dan kata negatif dari sentiment analisis aplikasi Dana dan Ovo tersebut.

3. HASIL DAN PEMBAHASAN

Pengambilan data diambil dari *Google play store* dengan bantuan *tools google colab* yang menggunakan bahasa *python*. Lalu setelah itu data disimpan kedalam bentuk file *csv* dalam *Microsoft excel*, dilanjut dengan pelabelan secara manual dan terkadang sesuai rating atau content dalam data tersebut. Untuk data dengan sentiment negatif diberi label 2 dan untuk data dengan sentiment positif diberi label 1. Bersamaan dengan pelabelan peneliti juga membuang data yang tidak sesuai dengan komentar pengguna aplikasi tersebut.

3.1 Scraping Data

Dari hasil *scraping* ini diperoleh data dari masing-masing aplikasi dana dan ovo 6000 data dan jika dijumlahkan peneliti mengambil 12.000 data secara keseluruhan. Berikut hasil dari scraper bisa dilihat pada gambar 2 dibawah ini.

```

from google_play_scraper import app

import pandas as pd

import numpy as np
    
```

Gambar 3. *Library Phyton*

Langkah pertama *import library* yang dibutuhkan terlebih dahulu seperti *pandas* dan *numpy*. Lalu setelah itu peneliti mengambil data dengan menyalin *link* dari aplikasi *dana* dan *ovo* tersebut dengan jumlah masing-masing 6000 data. Data yang diambil berupa *user name*, *komentar/content*, *waktu* dan juga *rating* seperti pada gambar 4 dibawah ini.

```

from google_play_scraper import Sort, reviews

result, continuation_token = reviews(
    'id.dana',
    lang='id',
    country='id',
    sort=Sort.MOST_RELEVANT,
    count=6000,
    filter_score_with=None
)
    
```

Gambar 4. Hasil *Scraping*

3.2. *Cleansing*

Setelah itu adalah tahap *cleansing*, tahap ini bertujuan menghapus tabel kosong atau kata yang tidak penting seperti pada tabel 4 dibawah ini.

Tabel 2. Hasil *Cleansing*

Content	<i>Cleaning Teks</i>
Apk ga bner.. Baru top up 1jt masuk terus ga bisa login, niat nyanyari bantuan ke cs malah di lempar sana sini,	apk top up jt masuk login niat nyari bantu cs lempar email balesyang apk awas ojk kaya apk abal abal
Jelas transaksi pending 24jam. Saat transaksi sudah fix Gagal, saldo tidak lekas di kembalikan Padahaluang sendiri, tpi udha kaya lagi ngutang	transaksi pending jam transaksi fix gagal saldo lekas kembali uang udha kaya hutang

3.3. *Preprocessing Text*

Tahap *preprocessing* data ini adalah tahap dimana data yang berbentuk teks yang tidak terstruktur atau ter acak diubah menjadi data yang terstruktur. Data yang dimasukkan dalam tahap ini berupa data yang asli dan belum melalui proses apapun. Maka dari itu hasil dari proses ini berupa data yang sudah bersih sesuai

dengan format yang ada dan untuk mempermudah kedalam tahap berikutnya yaitu klasifikasi. Pada *preprocessing* ini dilakukan 4 tahap yaitu *case folding*, *tokenizing*, *stopword removal* dan juga *stemming*.

Tabel 3. Hasil *Case Folding*

Content	Case Folding
Hai pihak dana mau menyarankan untuk kirim uang dengan system klaim mungkin waktunya bisa ditambah jadi 24 jam batas klaim	hai pihak dana mau menyarankan untuk kirim uang dengan system klaim mungkin waktunya bisa ditambah jadi jam batas klaim
Ovo sekarang jelek gak kaya dulu, mau masuk susah banget . bukan malah bagus jadi makin jelek	Ovo sekarang jelek gak kayak dulu mau masuk susah banget bukan malah bagus jadi makin jelek

Tabel diatas menunjukkan bahwa data yang sebelum di *case folding* terdapat huruf kapital dan juga nomer dan setelah melalui tahap *case folding* semua kata di samaratakan yang dimana huruf kapital menjadi huruf biasa.

Tabel 4. Hasil *Tokenizing*

Hai	Pihak	Dana	Mau
Menyarankan	Kirim	Uang	system
Ovo	Sekarang	Jelek	Gak
Kaya	Dulu	Mau	masuk

Selanjutnya yaitu tahap *tokenizing*, tahap ini adalah tahap memisahkan suatu kata satu persatu dalam sebuah kalimat seperti yang terlihat pada tabel diatas.

Tabel 5. Hasil *Stopword Removal*

Content	Stopword Removal
Hai pihak dana mau menyarankan untuk kirim uang dengan system klaim mungkin waktunya bisa ditambah jadi 24 jam batas klaim krn ini cocok untuk	hai dana menyarankan kirim uang sistem klaim ditambah jam batas klaim krn cocok pengiriman uang tujuan gift jdi pengirim pemberi

Content	Stopword Removal
pengiriman uang dengan tujuan gift jdi pengirim atau pemberi hadiah	hadiah
Ovo sekarang jelek gak kayak dulu, mau masuk susah banget.. bukan malah bagus jadi makin jelek	ovo jelek gak kayak masuk susah banget bagus jelek

Selanjutnya yaitu tahap *Stopword Removal*, dalam tahapan ini kata yang tidak penting akan dihilangkan sesuai dengan format yang ada seperti pada tabel diatas.

Tabel 6. Hasil *Stemming*

Content	Stemming
Hai pihak dana mau menyarankan untuk kirim uang dengan system klaim mungkin waktunya bisa ditambah jadi 24 jam batas klaim krn ini cocok untuk pengiriman uang dengan tujuan gift jdi pengirim atau pemberi hadiah	hai dana saran kirim uang sistem klaim tambah jam batas klaim krn cocok kirim uang tuju gift jdi kirim beri hadiah
Ovo sekarang jelek gak kayak dulu, mau masuk susah banget.. bukan malah bagus jadi makin jelek	ovo jelek gak kayak masuk susah banget bagus jelek

Selanjutnya tahap terakhir dari *preprocessing* ini adalah *stemming*. *Stemming* ini adalah tahap dimana kata yang ber imbuhan akan diubah menjadi kata dasar sesuai dengan format yang ada. Dalam stemming ini diperlukan suatu *library* yaitu sastrawi untuk menjalankannya. Berikut contoh kata dalam tahap stemming dapat dilihat pada tabel diatas.

3.4. Pembobotan kata TF-IDF

Setelah *preprocessing* selesai maka tahap berikutnya adalah Pembobotan kata TF-IDF. Pada tahap ini perhitungan bobot kata tiap *term* dicari pada setiap data, hal ini bertujuan untuk dapat mengetahui ketersediaan dan kemiripan suatu term didalam data, semakin banyak kata yang muncul pada data maka akan semakin banyak tinggi bobot pada atau nilai term tersebut.

Tabel 7. Sampel Aplikasi Dana

No	TEKS	SENTIMEN
1	mudah banget transaksi apa tarik uang mudah banget alfa gratis biaya tarik uang rekomend deh pokok	POSITIF
2	proses mudah fitur nya banyakin aman tingkat tampil aplikasi simpel ya biar berat apk nya lengkap sukses	POSITIF
3	aplikasi baik serba serbaguna naik rate nya tolong berantas hacker cari tempat dana error masuk curi uang guna dana tolong berantas ya minsegera tangkap	POSITIF
4	mantap bosku tingkatkan kecepatan aplikasinya dalam hal bertransaksi pembayaran dan lain lainnya jgn sering error terus ya	POSITIF
5	bagus aplikasi nya mudah kemaren errornya sulit bagus	POSITIF
6	dana aplikasi tipu transaksi proses uang transaksi kembali dompet utama live chat tanggap robot aplikasinya rugi	NEGATIF
7	upgred premium susah ampun top up rb transfer maksudnya cobaa	NEGATIF
8	kecewa aplikasi dana kali transaksi buka menu error update performa turun kecewa	NEGATIF
9	kecewa dana top up saldo tambah tanggal mey solusi bingung adu pusat bantu dana terima lapor tindak	NEGATIF
10	pakai aplikasi suka apa jalan masalah habis top up beli token masuk saldo masuk kesini beres aplikasi suruh masukin nomor hp pakai masukin nomor	NEGATIF

Tabel 8. Sampel Aplikasi Ovo

No	TEKS	SENTIMEN
1	transaksi mudah banget pakai ovo ribet set set selesai	POSITIF
2	bantu mudah transaksi ovo	POSITIF
3	Aplikasi ovo emang hebat dan mantap mudah sekali lagi buat transaksi atau tf	POSITIF
4	Selama menggunakan ovo, belum pernah ada kendala. dimohon untuk lebih bagusnya lagi jika ada bonus pemakaiannya. terima kasih ovo	POSITIF
5	salut si ovo respon admin cepet banget akun log in gara log o	POSITIF
6	Aplikasi ribet ada masalah bukan nya cepat di bantu malah diem aja kecewa dgn apk ini jelek	NEGATIF
7	ovo aman instal rugi transfer update ktp pakai pinjol jelek	NEGATIF
8	Eenggak bisa digunakan transfer walau sudah update sama sudah upgrade jadi ovo premier	NEGATIF
9	aplikasi tarik uang atm gara gara upgeade premium kacau uang tarik butuh gara ovo woi du tahan uang ya transfer hah	NEGATIF
10	Aplikasi ovo tidak bisa di transfer ke bank,, jadi jelek banget aplikasinya, padahal saya pengguna ovo premium	NEGATIF

Setelah mengambil sampel data dari masing-masing aplikasi maka akan dilanjutkan dengan mencari nilai *term frequency* dengan hasil 1 untuk kata yang muncul dan 0 untuk kata yg tidak muncul. Berikut hasil dari *term frequency* dapat dilihat pada tabel dibawah ini.

Tabel 9. Term Frequency Aplikasi Dana

Term	TF									
	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10
uang	1	0	1	0	0	1	0	0	0	0
dana	0	0	1	0	0	1	0	1	1	0
aplikasi	0	1	1	1	1	1	0	1	0	1
error	0	0	1	1	1	0	0	1	0	0
top up	0	0	0	0	0	0	1	0	1	1
mudah	1	1	0	0	0	0	0	0	0	0
kecewa	0	0	0	0	0	0	0	1	1	0
transaksi	1	0	0	1	0	1	0	1	0	0
saldo	0	0	0	0	0	0	0	0	1	1
masuk	0	0	1	0	0	0	0	0	0	1

Tabel 10. Term Frequency Aplikasi Ovo

Term	TF									
	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10
ovo	1	1	1	1	1	0	1	1	1	1
mudah	1	0	1	1	0	0	0	0	0	0
aplikasi	0	0	1	0	0	1	0	0	1	1
transaksi	1	1	1	0	0	0	0	0	0	0
transfer	0	0	0	0	0	0	1	1	1	1
update	0	0	0	0	0	0	1	1	0	0
cepat	0	0	0	0	1	1	0	0	0	0
premium	0	0	0	0	0	0	0	0	1	1
bantu	0	1	0	0	0	1	0	0	0	0
jelek	0	0	0	0	0	1	1	0	0	1

Setelah melewati tahap *term frequency* dan mendapatkan hasilnya maka tahap selanjutnya adalah *inverse document*. Setelah semua kata di beri bobot selanjutnya menghitung *deocument frequency (DF)*. Kemudian total dari seluruh dokumen dibagi dengan *dokumen frequency (DF)* untuk menghasilkan nilai dokumen. Hasil dari *inverse dokumen* dapat dilihat pada gambar di bawah ini.

Tabel 11. Inverse Document Aplikasi Dana

Term	TF										df	D/DF	f (log D/c)	1+Df
	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10				
ovo	1	1	1	1	1	0	1	1	1	1	9	1,11111	0,04576	1,04576
mudah	1	0	1	1	0	0	0	0	0	0	3	3,33333	0,52288	1,52288
aplikasi	0	0	1	0	0	1	0	0	1	1	4	2,50000	0,39794	1,39794
transaksi	1	1	1	0	0	0	0	0	0	0	3	3,33333	0,52288	1,52288
transfer	0	0	0	0	0	0	1	1	1	1	4	2,50000	0,39794	1,39794
update	0	0	0	0	0	0	1	1	0	0	2	5,00000	0,69897	1,69897
cepat	0	0	0	0	1	1	0	0	0	0	2	5,00000	0,69897	1,69897
premium	0	0	0	0	0	0	0	0	1	1	2	5,00000	0,69897	1,69897
bantu	0	1	0	0	0	1	0	0	0	0	2	5,00000	0,69897	1,69897
jelek	0	0	0	0	0	1	1	0	0	1	3	3,33333	0,52288	1,52288

Tabel 12. Inverse Document Aplikasi Ovo

Term	TF										df	D/DF	f (log D/c)	1+Df
	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10				
uang	1	0	1	0	0	1	0	0	0	0	3	3,33333	0,52288	1,52288
dana	0	0	1	0	0	1	0	1	1	0	4	2,50000	0,39794	1,39794
aplikasi	0	1	1	1	1	1	0	1	0	1	8	1,25000	0,09691	1,09691
error	0	0	1	1	1	0	0	1	0	0	4	2,50000	0,39794	1,39794
top up	0	0	0	0	0	0	1	0	1	1	3	3,33333	0,52288	1,52288
mudah	1	1	0	0	0	0	0	0	0	0	2	5,00000	0,69897	1,69897
kecewa	0	0	0	0	0	0	0	1	1	0	2	5,00000	0,69897	1,69897
transaksi	1	0	0	1	0	1	0	1	0	0	4	2,50000	0,39794	1,39794
saldo	0	0	0	0	0	0	0	0	1	1	2	5,00000	0,69897	1,69897
masuk	0	0	1	0	0	0	0	0	0	1	2	5,00000	0,69897	1,69897

Tahap berikutnya adalah tahap *term frequency inverse document frequency*. Untuk mendapatkan nilai *tf-idf* dilakukan cara dengan mengalikan bobot tiap kata dari seluruh dokumen dengan nilai *tf-idf* kata tersebut seperti terlihat pada gambar dibawah berikut.

Tabel 13. TF-IDF Aplikasi Dana

W=tf*(idf+1)										
D1	D2	D3	D4	D5	D6	D7	D8	D9	D10	
1,52288	0,00000	1,52288	0,00000	0,00000	1,52288	0,00000	0,00000	0,00000	0,00000	0,00000
0,00000	0,00000	1,39794	0,00000	0,00000	1,39794	0,00000	1,39794	1,39794	0,00000	0,00000
0,00000	1,09691	1,09691	1,09691	1,09691	1,09691	0,00000	1,09691	0,00000	1,09691	0,00000
0,00000	0,00000	1,39794	1,39794	1,39794	1,39794	0,00000	1,39794	0,00000	1,39794	0,00000
0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	1,52288	0,00000	1,52288	1,52288
1,69897	1,69897	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000
0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	1,69897	1,69897	0,00000
1,39794	0,00000	0,00000	1,39794	0,00000	1,39794	0,00000	1,39794	0,00000	1,39794	0,00000
0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	1,69897	1,69897
0,00000	0,00000	1,69897	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	1,69897
4,61979	2,79588	7,1464	3,89279	2,49485	5,41567	1,52288	6,98970	6,31876	6,01773	

Tabel 14. TF-IDF Aplikasi Ovo

W=tf*(idf+1)										
D1	D2	D3	D4	D5	D6	D7	D8	D9	D10	
1,04576	1,04576	1,04576	1,04576	1,04576	0,00000	1,04576	1,04576	1,04576	1,04576	1,04576
1,52288	0,00000	1,52288	1,52288	1,52288	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000
0,00000	0,00000	1,39794	0,00000	0,00000	1,39794	0,00000	0,00000	1,39794	1,39794	0,00000
1,52288	1,52288	1,52288	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000
0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	1,39794	1,39794	1,39794	1,39794	0,00000
0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	1,69897	1,69897	0,00000	0,00000
0,00000	0,00000	0,00000	0,00000	0,00000	1,69897	1,69897	0,00000	0,00000	0,00000	0,00000
0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	1,69897	1,69897
0,00000	1,69897	0,00000	0,00000	0,00000	1,69897	0,00000	0,00000	0,00000	0,00000	0,00000
0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	1,52288	1,52288	0,00000	0,00000	1,52288
4,09151	4,26761	5,48945	2,56864	2,74473	6,31876	5,66555	4,14267	5,54061	7,06349	

3.4. Klasifikasi Support Vector Machine

Selanjutnya jika sudah melewati tahap pembobotan kata maka akan langsung ke dalam proses klasifikasi, pada klasifikasi disini peneliti menggunakan algoritma *Support Vector Machine*. Algoritma ini sangat berguna untuk mengambil keputusan dengan memprediksi kasus berdasarkan hasil klasifikasi yang diperoleh. Dalam kasus ini peneliti menggunakan Bahasa *pyhton* dan *tools google colaboratory* untuk klasifikasi *Support Vector Machine*. Hal pertama dilakukan ialah memisahkan data *testing* dan *training* menggunakan *pyhton* seperti dilihat pada gambar dibawah ini.

```

x = selected x
y = data.label
#Memisahkan data training dan data testing dengan perbandingan 80:20
x_train, x_test, y_train, y_test = train_test_split(x,y,test_size=0.2, random_state=0)

[28] #print
print("Banyak data x_train :",len(x_train))
print("Banyak data x_test :",len(x_test))
print("Banyak data y_train :",len(y_train))
print("Banyak data y_test :",len(y_test))

Banyak data x_train : 4749
Banyak data x_test : 1188
Banyak data y_train : 4749
Banyak data y_test : 1188
    
```

Gambar 5. Data Training dan Testing Dana


```
[28] x = selected_x
y = data.label
#Memisahkan data training dan data testing dengan perbandingan 80:20
x_train, x_test, y_train, y_test = train_test_split(x,y,test_size=0.2, random_state=0)

#print
print('Banyak data x_train :',len(x_train))
print('Banyak data x_test :',len(x_test))
print('Banyak data y_train :',len(y_train))
print('Banyak data y_test :',len(y_test))

Banyak data x_train : 4757
Banyak data x_test : 1190
Banyak data y_train : 4757
Banyak data y_test : 1190
```

Gambar 6. Data Training dan Testing Ovo

Dari gambar diatas dapat terlihat banyak nya data *training* dan *testing* aplikasi dana ini adalah 4749 dan data testing nya 1188. Sedangkan untuk aplikasi ovo ini menghasilkan data training sebanyak 4757 dan data testing nya 1190.

Setelah tahap pemisahan data trainig dan testing selesai maka lanjut ke dalam tahap uji model dengan evaluasi data *training* yang nantinya akan menentukan prediksi benar, prediksi salah, akurasi, presisi dan juga recall seperti yang ditunjukkan pada gambar di bawah ini.

```
[29] #Training Model
from datetime import datetime
start_time = datetime.now()
from joblib import dump
#algoritme fitting

text_algorithm = svm.SVC(kernel="rbf", C=1.0)

model = text_algorithm.fit(x_train, y_train)

# save the model to disk
dump(model, filename="model_sentiment_naive.joblib")

end_time = datetime.now()
result_time = end_time-start_time
print("Duration:",result_time)

Duration: 0:00:26.366993

[30] #Prediksi
predicted = model.predict(x_test)

#Hasil ini akan kita bandingkan dengan nilai y_test (labeling dari manusia)
predicted

array([2, 2, 2, ..., 2, 2, 2])
```

Gambar 7. Uji Model

Setelah melewati tahap uji model maka dihasilkan suatu tingkat akurasi dari masing-masing aplikasi menurut data training dan data testing Berikut hasil dari perbandingan akurasi, presisi dan recal dari aplikasi dana dan ovo dapat terlihat pada tabel di bawah ini.

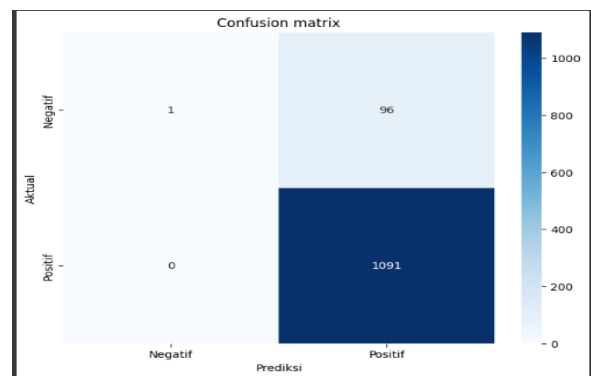
Tabel 15. Hasil Perbandingan

Aplikasi	Akurasi	Precision	Recall
Dana	0,92%	0,92%	100%
Ovo	0,90%	0,88%	0,98%

Dari perbandingan diatas aplikasi dana mendapatkan nilai akurasi, presisi dan recal menggunakan algoritma *Support Vector Machine* adalah 0,92% akurasi(Accuracy), Pressisi (Precision) 0,92%, dan nilai recall 100%. Sedangkan untuk aplikasi ovo mendapatkan nilai akurasi (*Accuracy*) 0,90%, Presisi(*Precision*) 0,88%, dan nilai recal 0,98%. Maka diperoleh hasil aplikasi dana memiliki nilai lebih baik karena tingkat akurasi , presisi dan recall nya lebih tinggi dari pada aplikasi ovo .

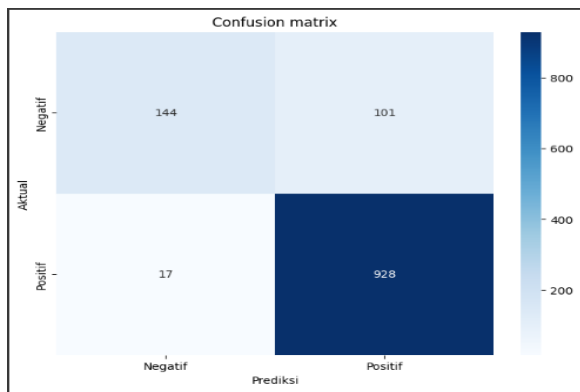
3.5 Evaluasi

Tahap evaluasi ini bertujuan untk mengecek kebenaran hasil klasifikasi data *testing* dengan menggunakan confusion matriks pada tools *google colab* yang menggunakan *python* dan dapat dilihat pada gambar dibawah ini.



Gambar 8. Confusion Matriks Aplikasi Dana

Pada gambar diatas menunjukkan hasil dari *confusion matriks* aplikasi dana dengan perhitungan *True negatif* 0, *false negatif* 0, *true positif* 1105 dan *false positif* 95. Sedangkan untuk *confusion matriks* aplikasi ovo dapat dilihat pada gambar di bawah ini.



Gambar 9. Confusion Matriks Aplikasi Ovo

Pada gambar diatas ini menunjukkan hasil dari *confusion matriks* aplikasi ovo dengan perhitungan *True negatif* 21, *false negatif* 154, *true positif* 924 dan *false positif* 100.

3.6 Visualisasi

Tahap terakhir adalah visualisasi, tahap ini di gambarkan dengan *wordcloud* pada seluruh sentiment ulasan positif dan ulasan negatif dari aplikasi dana dan ovo pada *google playstore*. Berikut hasil visualisasi dapat dilihat pada gambar dibawah ini.



Gambar 10. Visualisasi Aplikasi Dana

Pada gambar *wordcloud* diatas terdapat kata positif yang sering muncul pada keseluruhan komentar pengguna dari Aplikasi dana adalah : dana, aplikasi, *update*, saldo, login, transaksi, masuk dan lain lain.



Gambar 11. Visualisasi Aplikasi Ovo

Pada gambar *wordcloud* diatas terdapat kata – kata yang sering muncul pada keseluruhan

ulasan Aplkasi Ovo diatas adalah : ovo, aplkasi, transaksi, mudah, cepat, bantu bangt, top, saldo dan lain lain.

4. PENUTUP

4.1. Kesimpulan

Dari hasil penelitian diatas yakni mengenai Komparasi *Payment Digital* untuk analisis sentimen berdasarkan ulasan di *google playstore* berhasil mengumpulkan data yang berjumlah 6000 data dari masing-masing aplikasi yaitu Dana dan Ovo dan juga sudah melewati tahap *preprocess text* yang terdiri dari *case folding*, *tokenizing*, *stopword removal* dan *stemming* data tersebut menjadi 5936 untuk dana dan 5946 untuk Ovo. Penelitian ini menggunakan algoritma *Support Vector Machine* serta evaluasi data menggunakan *Confusion Matrix* menggunakan *Google Colaboratory* dapat disimpulkan bahwa tanggapan / ulasan dari Aplikasi Dana terbilang lebih bagus dari pada Aplikasi Ovo dengan tingkat persentase akurasi pada aplikasi dana 92% (*accuracy*), untuk nilai *precision* 92% dan 100% untuk tingkat keberhasilan *recall* nya. Sedangkan pada aplikasi ovo mempunyai nilai akurasi 90% (*accuracy*), untuk nilai presisi nya 88%, dan tingkat keberhasilan *recall* nya 98%. Dari nilai tersebut menunjukkan bahwa klasifikasi algoritma *Support Vector Machine* ternyata cukup baik dalam pemrosesan perbandingan data ulasan di *Google play store*. Berdasarkan nilai tersebut membuktikan aplikasi dana mendapat nilai akurasi paling besar yaitu 92% dengan ini aplikasi dana lebih cenderung mendapat sentimen positif oleh masyarakat dari pada aplikasi ovo.

4.2 . Saran

Saran yang dapat diberikan peneliti yaitu adalah sebagai berikut :

- Untuk Peneliti selanjutnya bisa mencoba atau menggunakan algoritma yang lainnya dalam menganalisis sentiment seperti algoritma K-Nearest Neighbor, Naive bayes, ataupun algoritma lainnya.
- Dapat memperbanyak aplikasi *payment digital* untuk komparasi/perbandingan menggunakan metode *support vector machine* ini.

- c. Dapat memperluas dataset yang lebih besar dan lebih beragam agar dapat membantu meningkatkan kinerja model analisis sentiment. Melibatkan data dari berbagai sumber, industry, atau Bahasa agar dapat memberikan wawasan yang lebih luas tentang sentimen

5. DAFTAR PUSTAKA

Agrani, A. (2020) 'PERBANDINGAN ANALISIS SENTIMEN TERHADAP DIGITAL PAYMENT " GO- PAY " DAN " OVO " DI MEDIA SOSIAL TWITTER MENGGUNAKAN ALGORITMA NAÏVE BAYES DAN WORD CLOUD COMPARISON OF SENTIMENT ANALYSIS AGAINST DIGITAL PAYMENT " GO- PAY " AND " OVO " IN SOCIAL MEDIA TWITTER USING NAÏVE BAYES ALGORITHM AND WORD CLOUD', 7(2), pp. 2534–2542.

Astuti, A. P., Alam, S. and Jaelani, I. (2023) 'Komparasi Algoritma Support Vector Machine Dengan Naive Bayes Untuk Analisis Sentimen Pada Aplikasi BRImo', XI(02), pp. 1–7.

Ditami, G. R., Ripanti, E. F. and Sujaini, H. (2022) 'Implementasi Support Vector Machine untuk Analisis Sentimen Terhadap Pengaruh Program Promosi Event Belanja pada Marketplace', 8(3), pp. 508–516.

Fariz, I. *et al.* (2023) 'Ilham Fariz Asya Mubarak, 2) Baenil Huda, 3) Agustia Hananto, 4) Tukino, 5) Huban Kabir', 8(1), pp. 69–74.

Hidayat, A. N., Hanggara, B. T. and Prakoso, B. S. (2021) 'Analisis Perbandingan Pengalaman Pengguna Pada Aplikasi Dompot Digital (Studi Kasus Pada OVO dan DANA)', 5(6), pp. 2181–2187.

Himawan, R. D. (2021) 'Perbandingan Akurasi Analisis Sentimen Tweet terhadap Pemerintah Provinsi DKI Jakarta di Masa', 7(1), pp. 58–63.

Khatib, J. *et al.* (2022) 'Indonesian Journal of Computer Science', 11(1), pp. 1036–1052.

Mubarak, D. A. and Akhmadi, M. H. (no date) 'IMPLEMENTASI SISTEM APLIKASI PEMBAYARAN DIGITAL PAYMENT

DALAM PELAKSANAAN PEMBAYARAN BERBASIS CASHLESS DI MASA PANDEMI COVID-19'.

Noviana, R. *et al.* (2023) 'PENERAPAN ALGORITMA NAIVE BAYES DAN SVM UNTUK ANALISIS SENTIMEN BOY', 2(2), pp. 51–60.

Penggunaan, P. *et al.* (2017) 'p 2337-4721', 5(1), pp. 145–151.

Sentimen, A. *et al.* (2021) 'Analisis Sentimen Pada Maskapai Penerbangan di Platform Twitter Menggunakan Algoritma Support Vector Machine (SVM) Sentiment Analysis of Airline on Twitter Platform Using Support Vector Machine (SVM) Algorithm', 10(1), pp. 18–26. doi: 10.34148/teknika.v10i1.311.

Technology, F. *et al.* (2020) 'financial technology', 2(2), pp. 155–170.