

Analisis Sentimen Dompot elektronik Pada Twitter Menggunakan Metode K-Means

Nova Ferawati Purba¹, Sri Agustina Rumapea², Arina Prima Silalahi³

^{1,2,3} Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Methodist Indonesia, Medan

e-mail: ¹ novaferawatipurba@gmail.com, ² sri_rumapea78@gmail.com, ³ primaarinasilalahi@gmail.com

Abstrak

Saat ini penggunaan dompet elektronik untuk transaksi non-tunai semakin berkembang pesat. Hal ini dikarenakan kemudahan dan kepraktisan transaksi yang ditawarkan oleh penyedia layanan dompet elektronik. Oleh karena itu, dalam penelitian ini akan dilakukan analisis tentang sentimen masyarakat terhadap dompet elektronik untuk mendapatkan tolak ukur seberapa jauh masyarakat mengetahui dompet elektronik. K-Means merupakan metode pengklasteran secara partitioning yang memisahkan data ke dalam kelompok yang berbeda. Dengan partitioning secara iteratif, K-Means mampu meminimalkan rata-rata jarak setiap data ke klasternya. Metode K-Means digunakan sebagai pendekatan analisis untuk mengelompokkan dan mengevaluasi pola sentimen yang muncul dalam percakapan online. Davies bouldin index (DBI) adalah matriks untuk mengevaluasi atau mempertimbangkan hasil algoritma clustering. DBI suatu cluster akan dianggap memiliki skema clustering yang optimal adalah yang memiliki DBI minimal. Hasil dari penelitian ini berupa jumlah sentimen positif 203 data, jumlah sentimen negatif 50 data dan jumlah sentimen netral 224 data kemudian metode K-means menghasilkan sentimen positif dompet elektronik Gopay sebanyak 63.801%, Ovo 56.951% dan untuk Dana sebanyak 27.907%. Penggunaan DBI untuk mengukur sejauh mana pengelompokan data yang dihasilkan oleh algoritma clustering tertentu adalah pengelompokan yang baik atau valid. Semakin kecil nilai DBI, semakin baik kualitas pengelompokan yang diperoleh.

Kata kunci: Twitter, Sentimen, K-Means, Dompot digital.

Abstract

Currently, the use of electronic wallets for non-cash transactions is growing rapidly. This is due to the ease and practicality of transactions offered by electronic wallet service providers. Therefore, in this research an analysis of public sentiment towards electronic wallets will be carried out to obtain a benchmark for how much the public knows about electronic wallets. K-Means is a partitioning clustering method that separates data into different groups. By iterative partitioning, K-Means is able to minimize the average distance of each data to its cluster. The K-Means method is used as an analytical approach to group and evaluate sentiment patterns that emerge in online conversations. Davies Bouldin Index (DBI) is a matrix for evaluating or considering the results of clustering algorithms. The DBI of a cluster will be considered to have an optimal clustering scheme which has a minimum DBI. The results of this research are 203 positive sentiment data, 50 negative sentiment data and 224 neutral sentiment data, then the K-means method produces 63,801% positive sentiment for Gopay electronic wallets, 56,951% for Ovo and 27,907% for Dana. The use of DBI is to measure the extent to which the data grouping produced by a particular clustering algorithm is a good or valid grouping. The smaller the DBI value, the better the grouping quality obtained.

Keywords: Twitter, Sentiment, K-Means, Digital wallet.

1. PENDAHULUAN

Dompot elektronik adalah penggunaan teknologi dalam sistem keuangan yang menghasilkan produk, layanan, teknologi, atau model bisnis baru serta dapat berdampak pada stabilitas moneter, stabilitas sistem keuangan, atau efisiensi, kelancaran, keamanan, dan keandalan sistem pembayaran [1]. Pengguna dompet elektronik saat ini telah diatur oleh Bank Indonesia dalam peraturan Bank Indonesia Nomor 20/6/PBI/2018 tentang dompet elektronik. Saat ini penggunaan dompet elektronik untuk transaksi non-tunai semakin berkembang pesat [2]. Hal ini dikarenakan kemudahan dan kepraktisan transaksi yang ditawarkan oleh penyedia layanan dompet elektronik. Kemunculan dompet elektronik membantu mempermudah masyarakat mengakses produk atau layanan keuangan. Masyarakat biasanya menggunakan dompet elektronik untuk keperluan pembayaran dengan efisien karena tidak perlu dilakukan secara tunai. Konsumen juga berkesempatan menikmati jenis-jenis promosi[3] .

Penggunaan dompet elektronik tidak jarang mengalami beberapa permasalahan diantaranya seringkali terjadi delay pembayaran atau bisa disebut juga keterlambatan masuknya sejumlah uang, dimana proses transaksi bayar secara non-tunai tersebut nominal uang yang dibayarkan bisa sampai pada penyedia jasa dalam kurun waktu yang lama. Kendala lainnya terjadi akibat adanya gangguan pada kualitas internet yang mengakibatkan lambatnya kinerja aplikasi dalam pembayaran. Bahkan penipuan yang kerap terjadi. Dari kendala di atas banyak opini masyarakat pada media sosial twitter mengenai kinerja dompet elektronik. Twitter merupakan media online hasil dari teknologi yang banyak digunakan masyarakat dalam penyaluran aspirasi, ide, informasi, minat dan lain – lain[4][5]. Banyak isu yang diangkat secara bebas di twitter bahkan dengan adanya hastag (#), suatu topik dapat menjadi bahan pembicaraan masyarakat dunia. Pada twitter terdapat official account dari penyedia layanan dompet elektronik dengan akun @gopayindonesia, @ovo_id, @danawallet yang menjadi pen jembatan antara pengguna dengan pemberi layanan dalam memberikan review, masukan, kritik serta saran.

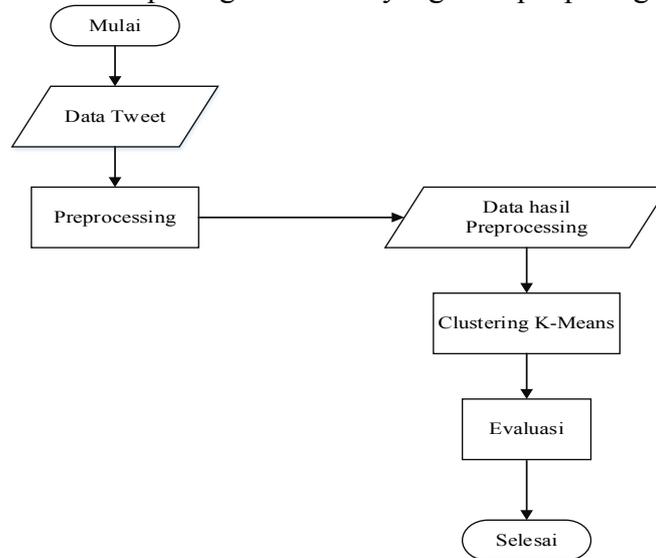
Terdapat beberapa penelitian terkait analisis sentimen masyarakat terhadap dompet elektronik [6][7]. M . Wino dalam penelitiannya menggunakan metode Naïve Bayes Classifier terhadap 3 dompet elektronik Dana, Ovo dan Gopay. Hasil dari penelitian ini menunjukkan Gopay merupakan dompet elektronik yang lebih banyak dinilai positif oleh pengguna twitter dengan nilai 46,67%.[6] Sedangkan Siti Saidah dalam penelitiannya menggunakan metode lexion based dan k-nearest terhadap 3 dompet elektronik Ovo, Gopay dan LinkAja. Dengan hasil penelitian bahwa dompet elektronik Gopay memiliki nilai sentimen positif sebesar 94,05% [7]. Pada kedua penelitian tersebut belum dilakukan perhitungan akurasi terhadap hasil cluster dari analisis sentimen yang dilakukan.

Oleh karena itu maka dilakukan analisis sentimen terhadap 3 dompet elektronik oleh Ovo, Gopay dan Dana. Dengan menggunakan metode K-Means untuk mengelompokkan dan mengevaluasi atau mempertimbangkan hasil algoritma clustering menggunakan metode DBI.

2. METODE PENELITIAN

Metode yang digunakan dalam penelitian ini dimulai dengan *preprocessing* pada data yang diperoleh dari platform media sosial twitter, kemudian akan dilanjutkan dengan menggunakan perangkat lunak rapid miner yang bertujuan untuk menerapkan

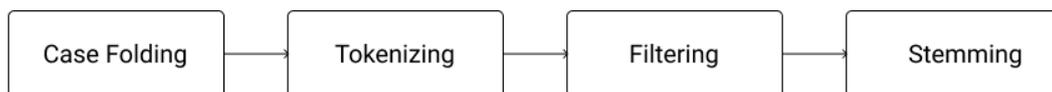
algoritma k-means dalam menentukan dompet digital terbaik. Berikut ini diagram proses dalam menentukan dompet digital terbaik yang terdapat pada gambar 1



Gambar 1. Arsitektur penelitian

2.1 Preprocessing teks

Setelah mendapatkan hasil *crawling* data dan dijadikan dataset tahapan selanjutnya adalah tahapan *preprocessing*. Tahapan ini dengan cara mengubah data mentah yang tidak terstruktur menjadi data tweet yang baik dan siap untuk diolah dengan jumlah cluster yang digunakan sebanyak 3 *cluster* yaitu positif, negatif dan netral. Dokumen-dokumen yang ada kebanyakan tidak memiliki struktur yang pasti sehingga informasi di dalamnya tidak bisa diekstrak secara langsung. *Preprocessing* diperlukan untuk memilih kata yang digunakan sebagai indeks. Kata kata yang mewakili dokumen yang nantinya digunakan untuk membuat pemodelan untuk *Information Retrieval* maupun aplikasi teks mining lain. *Preprocessing* pada penelitian ini terdiri dari beberapa tahapan dapat dilihat pada Gambar 2 dibawah ini



Gambar 2. Preprocessing teks

Adapun tahapan dari preprocessing teks sebagai berikut:

1. Case Folding

Pada proses case folding merupakan mengubah huruf kapital menjadi huruf kecil. Proses ini dilakukan agar semua data yang diproses memiliki penyeragaman karakter, sehingga ketika diproses akan mempermudah melakukan tahap selanjutnya. Contoh proses pelaksanaan Case Folding pada tahap preprocessing dapat dilihat pada Tabel 1.

Tabel 1. Contoh Proses Pelaksanaan Case Folding

Sebelum	Sesudah
berapa kali transaksi dengan dana seperti ini payah live chat lewat mana ya adanya digital assistant	berapa kali transaksi dengan dana seperti ini payah live chat lewat mana ya adanya digital assistant

2. *Tokenizing*

Pada proses tokenizing merupakan memisahkan kalimat menjadi data tunggal. Pada proses tokenizing ini dilakukan untuk mempermudah saat masuk tahap transformasi dengan tidak memproses kalimatnya tapi memproses kata demi kata dari kalimat tersebut. Contoh proses pelaksanaan Tokenizing pada tahap preprocessing dapat dilihat pada Tabel 2.

Tabel 2. Contoh Proses Pelaksanaan Tokenizing

Sebelum	Sesudah
berapa kali transaksi dengan dana seperti ini payah live chat lewat mana ya adanya digital assistant	'berapa', 'kali', 'transaksi', 'dengan', 'dana', 'seperti', 'ini', 'payah', 'live', 'chat', 'lewat', 'mana', 'ya', 'adanya', 'digital', 'assistant'

3. *Filtering*

Data yang sudah melakukan tahap *Tokenizing*, selanjutnya melakukan proses *filtering*. Pada proses *filtering* merupakan membuang kata yang tidak berpengaruh. Kata umum yang biasanya muncul dan tidak memiliki makna disebut dengan *stopword*. Pada proses ini menggunakan dokumen yang berisikan kata-kata *stopword*, yang digunakan pada proses klasifikasi. Contoh kata yang tidak diperlukan yaitu kata : yang, di, dan, dari, nya, dan lain lain. Contoh proses pelaksanaan Filtering pada tahap preprocessing dapat dilihat pada Tabel 3.

Tabel 3. Contoh Proses Pelaksanaan Filtering

Sebelum	Sesudah
berapa kali transaksi dengan dana seperti ini payah live chat lewat mana ya adanya digital assistant	berapa kali transaksi dengan dana seperti ini live chat lewat mana ada digital assistant

4. *Steaming*

Pada proses *steaming* merupakan membersihkan imbuhan dalam sebuah kata yang terdapat pada awal, akhir ataupun kombinasi dari keduanya. Dilakukan proses tahapan ini untuk mengurangi frekuensi dari sebuah kata turunan. Contoh proses pelaksanaan Filtering pada tahap preprocessing dapat dilihat pada Tabel 4.

Tabel 4. Contoh Proses Pelaksanaan Filtering

Sebelum	Sesudah
berapa kali transaksi dengan dana seperti ini live chat lewat mana ada digital assistant	berapa kali transaksi dengan dana seperti ini live chat lewat mana ada digital assistant

2.2 Metode K-Means

K-means clustering merupakan salah satu algoritma *machine learning* yang sederhana dan populer digunakan untuk memecahkan masalah pengelompokan dataset yang belum di label ke dalam kluster yang berbeda. Dalam *machine learning*, *k-means clustering* termasuk ke dalam jenis algoritma *unsupervised learning*. Algoritma *K-means* mengambil dataset yang tidak berlabel sebagai input, kemudian membagi dataset menjadi sejumlah *k cluster*, dan mengulangi proses tersebut sampai tidak menemukan *cluster* terbaik. Nilai *k* harus ditentukan sebelumnya dalam algoritma ini [8].

Langkah-langkah algoritma *k-Means* adalah sebagai berikut:

1. Menentukan beberapa *klaster k*
2. Menentukan nilai *centroid*. Nilai awal *centroid* dicari dengan secara acak. Sedangkan untuk nilai *centroid* saat iterasi, menggunakan rumus sebagai berikut:

$$V_{ij} = \frac{1}{N_i} \sum_{k=0}^{N_i} x_{kj} \dots\dots\dots (1)$$

Dimana :

V_{ij} adalah *centroid* atau rata rata kluster ke-*i* untuk variabel ke-*j* ;

N_i adalah jumlah data yang menjadi anggota dari *kluster* ke-*i* ;

i, k adalah indeks dari *kluster* ; j adalah indeks dari variabel ;

X_{kj} adalah nilai data ke- k yang terdapat pada kluster tersebut untuk variabel ke- j ;

3. Menghitung jarak antara titik *centroid* dengan setiap titik objek. Menghitung dengan menggunakan *euclidean distance*, dengan rumus sebagai berikut:

$$De = \sqrt{x_i - S_i + (y_1 - y_2)^2} \dots\dots\dots (2)$$

Dimana ;

De adalah euclidean distance ;

i adalah banyaknya objek ; x, y adalah koordinator objek ; s, t adalah koordinator *centroid* ;

4. Mengelompokkan objek. Yaitu untuk menentukan anggota kluster dengan menghitung jarak minimum objek. Nilai yang diperoleh dalam keanggotaan data pada jarak matriks adalah 0 atau 1, dimana nilai 1 untuk data yang dialokasikan ke kluster dan nilai 0 untuk data yang dialokasikan ke kluster lain.
5. Jika ada perubahan susunan kelompok, maka kembali lagi ke langkah kedua. Melakukan iterasi hingga hasil centroid tetap dan anggota kluster tidak berpindah ke kluster lain. Jika tidak, maka didapatkan kelompok data.

Setelah fitur-fitur numerik dihasilkan, algoritma *K-means* diterapkan untuk mengelompokkan tweet menjadi kelompok-kelompok berdasarkan kesamaan fitur. Jumlah kelompok atau kluster harus ditentukan sebelumnya. Algoritma *K-means* akan menghitung jarak antara setiap tweet dan pusat kluster, serta mengupdate posisi pusat kluster secara iteratif untuk mencapai konvergensi. Berikut ini perhitungan manual algoritma kemans pada data gopay, ovo dan dana. Nilai *centroid* sebagai titik data dapat dilihat pada Tabel 1.

Tabel 1. Jumlah cluster

DOC	Tweet	Label
Doc 1	Positif	C1
Doc 2	Netral	C2
Doc 3	Negatif	C3

Dari tabel diatas menjadi titik *centroid* yang dapat dilakukan pengelompokkan objek pada iterasi pertama dengan nilai *centroid* yang dapat dilihat pada Tabel 2

Tabel 2 Nilai Centroid Iterasi ke-1

KODE	Asisten	digital	Dana	Siap	Bantuan	keluhan	Transaksi	top up	sudah	transfer
C1	0,60	0,60	0,38	1,08	1,08	0,78	0	0	0	0
C2	0	0	0,38	0	0	0	0	1,08	0,60	0,60
C3	0	0	0	0	0	0	0	0	0	0

Perhitungan Jarak Data Dengan *Centroid* tersebut Kemudian akan dihitung jarak dari setiap data ke setiap pusat cluster yang ada dengan rumus *Euclidean distance* dapat dilihat pada Tabel 3.

Tabel 3. Tabel Perhitungan Jarak Iterasi ke-1

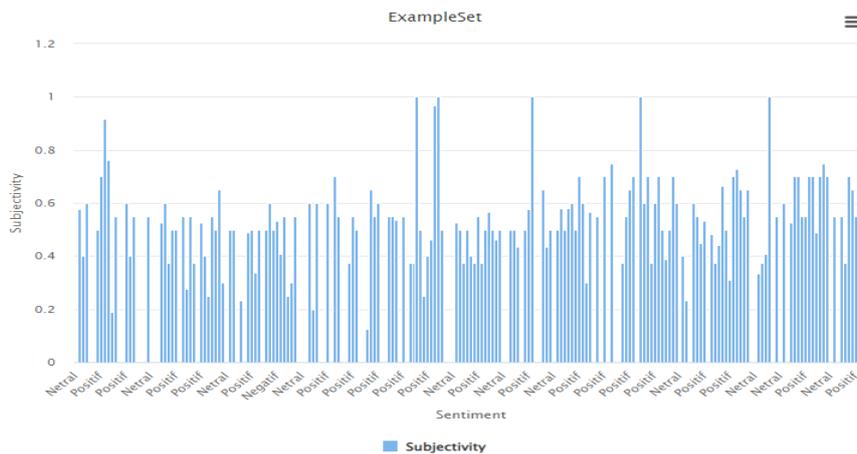
NO	DOKUMEN	C1	C2	C3	MIN	CLUSTER
1	D1	0	3,07	3,29	0	C1
2	D2	3,29	2,87	0	0	C3
3	D3	3,07	0	2,87	0	C2
4	D4	4,06	3,73	3,92	3,73	C2
5	D5	4,62	3,83	4,66	3,83	C2
6	D6	4,33	4,34	4,5	4,33	C1
7	D7	6,27	6,06	5,95	5,95	C3
8	D8	4,15	3,83	3,97	3,83	C2
9	D9	4,01	3,67	3,82	3,67	C2
10	D10	3,12	2,61	2,89	2,61	C2
11	D11	3,13	2,68	2,89	2,68	C2

3. HASIL DAN PEMBAHASAN

Data yang telah selesai di analisis maka akan diperoleh hasil berupa kuantitatif yang berbentuk data nominal yang menunjukkan opini masyarakat selaku konsumen mengenai layanan jasa dompet elektronik. Penjelasan hasil implementasi pada perangkat lunak yang telah dibuat pada penelitian ini. kemudian akan divisualisasikan data masing masing dari dompet digital.

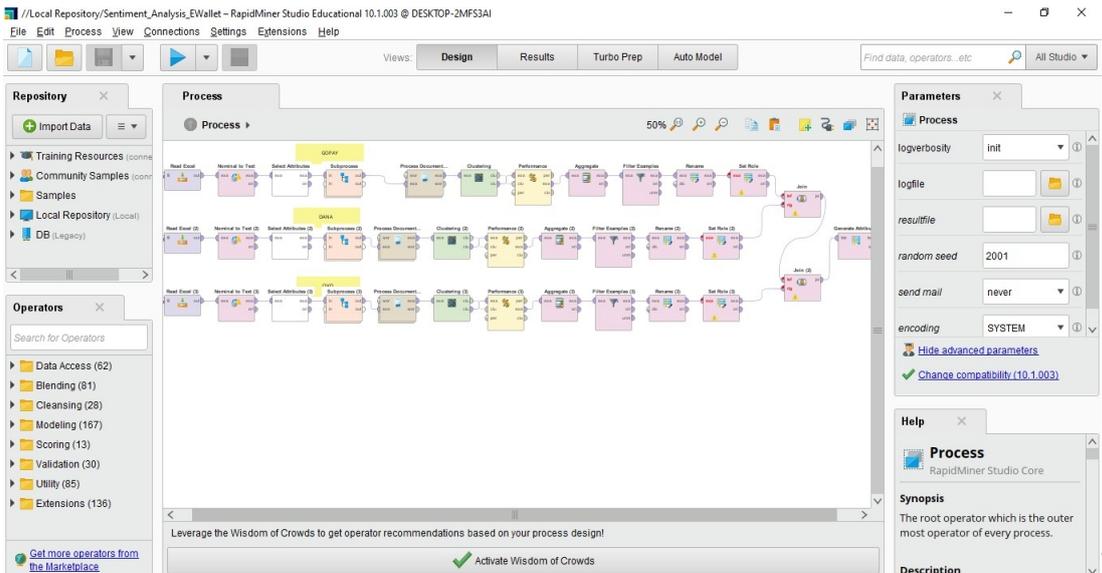
3.1 Visualisasi cluster

Visualisasi *cluster* dalam analisis klastering adalah cara yang efektif untuk memahami dan memvisualisasikan grup-grup yang terbentuk dari data, pada penelitian ini cluster 0 adalah sentimen positif, cluster 1 adalah sentimen negatif dan cluster 2 adalah sentimen netral. Berikut ini visualisasi klaster yang dihasilkan. Tampilan hasil visualisasi data dompet digital dapat dilihat pada gambar 3. berikut ini



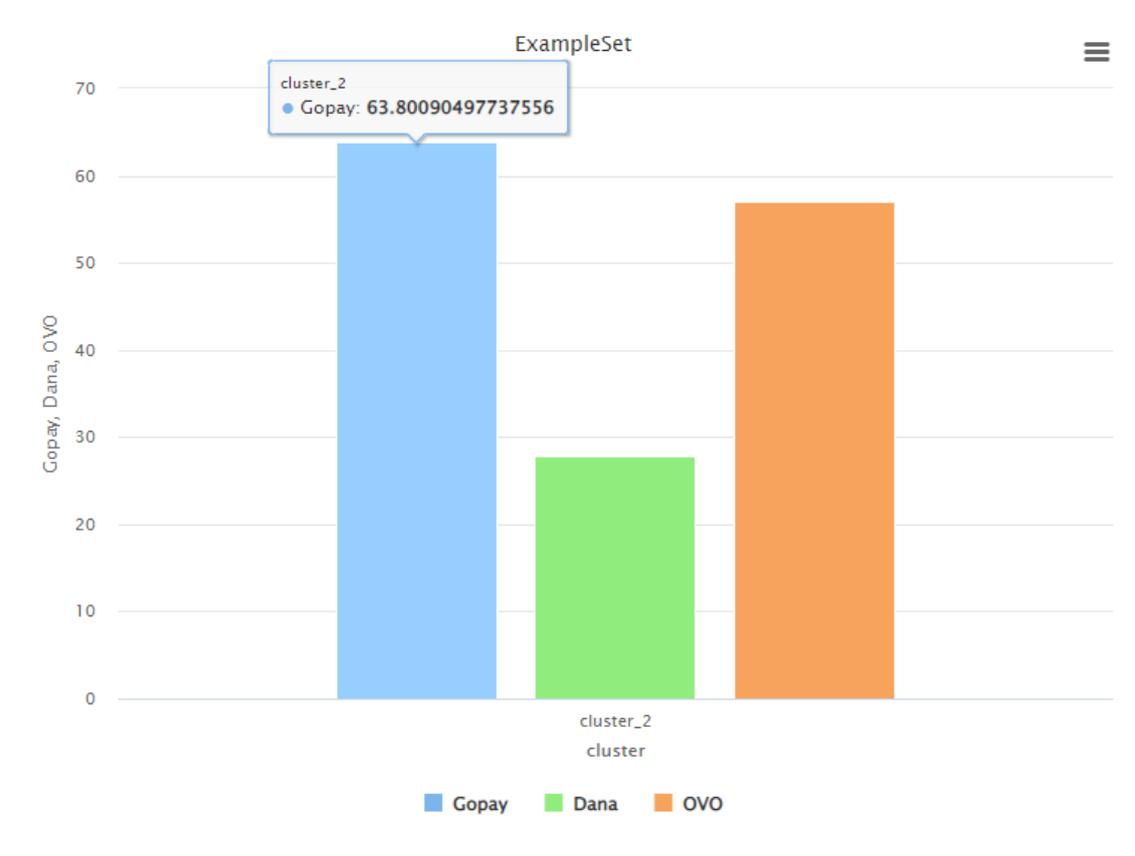
Gambar 3. Visualisasi data dompet digital

Setelah dilakukan penerapan algoritma K-means kemudian akan melakukan konfigurasi dalam sentiment analisis untuk melihat perbandingan data antara dompet digital yang akan diproses dengan algoritma kmeans yang dapat dilihat pada Gambar 4 berikut ini.



Gambar 4. Perbandingan data

Kemudian hasil dari proses k-means akan melakukan clustering secara otomatis dengan memanfaatkan algoritma k-means menggunakan rumus jarak mahattan distance, maka berikut ini hasil dari proses k-means sehingga dapat disimpulkan **Sentiment dompet elektronik GoPay** dari data tweet bernilai **POSITIF**. Tampilan hasil grafik dari analisis ini dapat dilihat pada Gambar 5 berikut ini.



Gambar 5. Hasil algoritma K-means

Setelah didapatkan hasil dari penerapan algoritma K-means kemudian akan dilakukan evaluasi dengan Indeks Davies-Bouldin (DBI) adalah metrik evaluasi yang digunakan untuk mengukur kualitas dan kebaikan kluster dalam analisis clustering. Tujuan dari DBI adalah untuk mengukur sejauh mana setiap kluster berada dalam kaitannya dengan kluster lainnya, berdasarkan jarak antara pusat kluster dan seberapa baik data dalam kluster tersebar. Berikut ini hasil dari Indeks Davies-Bouldin (DBI) yang terdapat pada Gambar 6.



Gambar 6. Evaluasi kinerja algoritma

4. KESIMPULAN DAN SARAN

Analisis sentimen terhadap dompet elektronik dengan menggunakan metode K-Means mendapatkan hasil sentimen positif dompet elektronik Gopay sebanyak 63.801%, Ovo 56.951% dan untuk Dana sebanyak 27.907%. Dari hasil proses diatas maka dompet elektronik gopay dengan 63.801% sentiment positif dapat ditentukan bahwa gopay merupakan dompet elektronik yang layak digunakan masyarakat. Saran untuk peneliti selanjutnya Data yang digunakan pada penelitian selanjutnya dapat ditambahkan sehingga akurasi dalam penggunaan Davies-Bouldin-Indeks (DBI) dapat menghasilkan nilai yang terbaik.

DAFTAR PUSTAKA

- [1] "Pengukuran Penerimaan Pengguna Dompet Elektronik Berbasis Server D," no. [1].
- [2] L. Angelica and U. Soebiantoro, "Analisa menggunakan dompet digital," *J. Manaj.*, vol. 14, no. 2, pp. 232–238, 2022, doi: 10.30872/jmmn.v14i2.11209.
- [3] E. S. Basryah, A. Erfina, and C. Warman, "Analisis Sentimen Aplikasi Dompet Digital Di Era 4.0 Pada Masa Pandemi Covid-19 Di Play Store Menggunakan Algoritma Naive Bayes Classifier," *SISMATIK (Seminar Nas. Sist. Inf. dan Manaj. Inform. Univ.*, vol. 1, no. 1, pp. 189–196, 2021.
- [4] "BAB I," pp. 1–11.
- [5] M. M. Rizki, "Analisis Sentimen Terhadap Produk Otomotif Dari Twitter Menggunakan Kombinasi Algoritma K-Nearest Neighbor dan Pendekatan Lexicon (Studi Kasus: Mobil Toyota)," *Repository.Uinjkt.Ac.Id*, pp. 1–127, 2019, [Online]. Available:

- <http://repository.uinjkt.ac.id/dspace/handle/123456789/48643>
- [6] M. W. A. Putra, Susanti, Erlin, and Herwin, “Analisis Sentimen Dompok Elektronik Pada Twitter Menggunakan Metode Naïve Bayes Classifier,” *IT J. Res. Dev.*, vol. 5, no. 1, pp. 72–86, 2020, doi: 10.25299/itjrd.2020.vol5(1).5159.
- [7] S. S. Salim and J. Mayary, “Analisis Sentimen Pengguna Twitter Terhadap Dompok Elektronik Dengan Metode Lexicon Based Dan K – Nearest Neighbor,” *J. Ilm. Inform. Komput.*, vol. 25, no. 1, pp. 1–17, 2020, doi: 10.35760/ik.2020.v25i1.2411.
- [8] M. A. Z. Larasati, N. A. S. Winarsih, M. S. Rohman, and G. W. Saraswati, “Penerapan Metode K-Means Clustering Dalam Menganalisis Sentimen Masyarakat Terhadap K-Popers Pada Twitter,” *Progresif J. Ilm. Komput.*, vol. 18, no. 2, p. 201, 2022, doi: 10.35889/progresif.v18i2.877.
-