

Perbandingan Algoritma *K-Means* dan *K-Medoids* untuk Pengelompokan Data Transaksi Bongkar Muat di Provinsi Riau

¹Insanul Kamila, ²Ulya Khairunnisa, ³Mustakim

^{1,2,3}Program Studi Sistem Informasi, Fakultas Sains dan Teknologi UIN Sultan Syarif Kasim Riau

^{1,2,3}Puzzle Research Data Technology, Fakultas Sains dan Teknologi UIN Sultan Syarif Kasim Riau
Jl. HR Soebrantas KM.18 No. 155 Panam Pekanbaru – Riau 28293

Email: ¹11653201531@student.uin-suska.ac.id,

²11653201346@student.uin-suska.ac.id, ³mustakim@uin-suska.ac.id

ABSTRAK

Data mining merupakan pemrosesan sebuah informasi dari suatu *database* yang dapat digunakan untuk berbagai kebutuhan sektor swasta. Salah satu metode dalam *data mining*, yaitu *Clustering* yang bertujuan untuk menemukan pengelompokan dari serangkaian pola, titik, objek maupun dokumen. Algoritma *K-Means clustering* merupakan algoritma yang berperan penting dalam bidang data mining serta sederhana untuk diimplementasikan dan dijalankan. Selain itu, terdapat pengembangan varian dari metode *K-Means Clustering* yaitu *K-Medoids* yang muncul sebagai penanggulangan kelemahan Algoritma *K-Means* yang sensitif terhadap *outlier* dikarenakan sebuah objek dengan sebuah nilai yang besar mungkin secara substansial menyimpang dari distribusi data. Penelitian ini membandingkan kedua algoritma pada suatu *dataset*. Adapun data yang digunakan merupakan data transaksi bongkar muat selama tahun 2017 pada PT Pelabuhan Indonesia I Cabang Dumai berdasarkan atribut agen, keterangan barang, jenis, dan jumlah ton. Dari percobaan yang dilakukan, diperoleh hasil pengolahan *K-Means* hanya membutuhkan waktu rata-rata 1 detik sedangkan pengolahan data pada *K-Medoids* membutuhkan waktu rata-rata 1 menit 38 detik pada RapidMiner. Nilai DBI pada *K-Means* lebih rendah dibandingkan *K-Medoids* yaitu masing-masing 0.112 dan 0.119. Perolehan *cluster* dominan, menunjukkan bahwa agen Buana Listya Tama TBK, PT mendominasi diikuti agen Samudera Sarana Karunia, PT.

Kata kunci: Bongkar Muat Barang, *Clustering*, *Data Mining*, *Davies-Bouldin Index*, *K-Means*, *K-Medoids*.

A. PENDAHULUAN

Data mining merupakan pemrosesan sebuah informasi dari suatu *database* yang dapat dilakukan meski tidak terlihat secara langsung. *Data mining* dapat digunakan untuk berbagai kebutuhan sektor swasta. Bidang industri misalnya pada perbankan, asuransi, dan obat-obatan dimana biasanya menggunakan teknik data mining untuk mengurangi biaya, meningkatkan pengetahuan, dan meningkatkan penjualan. Teknik analisis data pada umumnya berkisar seperti analisis regresi, analisis *cluster*, taksonomi numerik, analisis multidimensi, metode statistik multivariat, model stokastik, analisis deret waktu, teknik estimasi nonlinear, dan lainnya [1]. Kuantitas data yang diperoleh akan semakin bertumbuh lebih banyak lagi, dimana dibutuhkan alat dan teknik yang tepat dalam memproses data tersebut sehingga ekstraksi informasi yang dilakukan menjadi berguna untuk mengambil keputusan pada suatu perusahaan maupun organisasi kedepannya. Maka, dengan adanya *data mining*, mulai berkembang pesat menjadi bidang penelitian penting serta dibutuhkan untuk meningkatkan kajian ilmiah di bidang kecerdasan buatan dan statistik [1][2].

Salah satu metode pada *data mining*, yaitu *Clustering* yang merupakan metode pengelompokan data. Tujuan dari pengelompokan *cluster* ini adalah untuk menemukan pengelompokan dari serangkaian pola, titik, objek maupun dokumen. Objek yang

berada didalam pengelompokan *cluster* yang sama memiliki kemiripan antar satu kelompok dan memiliki perbedaan dengan objek oleh kelompok *cluster* lain [4]. Algoritma *K-Means clustering* merupakan algoritma yang berperan penting dalam bidang *data mining* serta sederhana untuk diimplementasikan dan dijalankan. Dalam prakteknya, algoritma ini umum digunakan karena relatif cepat dan mudah beradaptasi [13]. Selain itu, terdapat pengembangan varian dari metode *K-Means Clustering* yaitu *K-Medoids*. Bertujuan mengurangi sensitivitas dari partisi yang dihasilkan sehubungan dengan nilai-nilai ekstrim yang terdapat dalam *dataset*, penggunaan *medoids* tidak berdasarkan dari pengamatan *mean* yang dimiliki oleh setiap *cluster*. Algoritma *K-Medoids* muncul sebagai penanggulangan kelemahan Algoritma *K-Means* yang sensitif terhadap *outlier* dikarenakan sebuah objek dengan sebuah nilai yang besar mungkin secara substansial menyimpang dari distribusi data [14].

Berdasarkan pada penelitian sebelumnya oleh Arbin, dkk (2015) terkait penggunaan metode *K-Medoids* dan *K-Means* sebagai perbandingan pengelompokan data statistik, disimpulkan bahwa analisis dari kedua pendekatan metode menghasilkan hasil yang baik dengan *mean square error* kurang dari 3%. Berdasarkan *dataset*, *K-Medoids* menjadi pendekatan atau metode terbaik untuk mengelompokkan data. Untuk analisis lebih

lanjut, *dataset* yang berbeda dari berbagai kategori data harus dilakukan pengujian. Jenis *dataset* yang berbeda dapat menghasilkan hasil pilihan pendekatan atau metode yang berbeda [3].

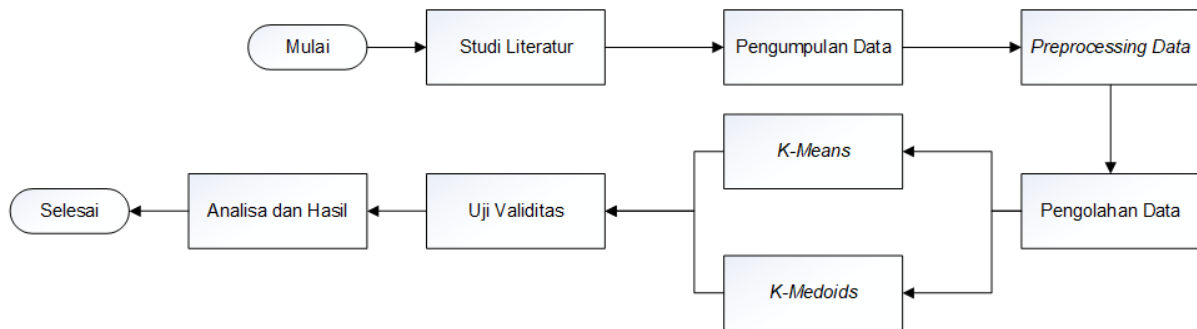
Kemudian pada penelitian oleh Aryuni, dkk (2018) mengenai Klasterisasi data Internet Banking di Bank XYZ menggunakan metode *K-Means* dan *K-Medoids Clustering* untuk melakukan segmentasi customer, diperoleh nilai *k* yang paling optimal pada kedua algoritma adalah 3 dengan nilai *average within centroid distance (W)* sebesar 35.241 untuk *K-Means* dan sebesar 88.849 untuk *K-Medoids*. Selanjutnya total waktu eksekusi dari proses *elbow method* berturut-turut untuk *K-Means* dan *K-Medoids* adalah 3 detik dan 13 menit 44 detik. Maka dapat disimpulkan dalam pengolahan data, algoritma *K-Means* memiliki performa yang lebih baik daripada *K-Medoids* baik dari sisi nilai *average within centroid distance* dan kompleksitas waktu [15].

Agar tidak terjadi perbedaan pendapat lebih luas, penelitian ini akan langsung membandingkan kedua algoritma pada *dataset* bertujuan untuk menganalisa pola pengelompokan dan menentukan algoritma yang baik dalam pengolahan data. Adapun data yang digunakan merupakan data transaksi bongkar muat oleh agen didermaga B selama tahun 2017 pada PT Pelabuhan Indonesia I Cabang Dumai. Dermaga B pada perusahaan tersebut merupakan

transaksi muatan curah cair dan menghasilkan data terbanyak dibandingkan dermaga lainnya serta berpotensi menjadikan perusahaan PT Pelabuhan Indonesia I Cabang Dumai *First Class Port* dengan catatan ekspor minyak terbesar se-Sumatera[16]. Tanpa adanya agen yang melakukan transaksi dan mengandalkan fasilitas pelabuhan tersebut, tentu akan mempengaruhi objek pendapatan perusahaan. Untuk itu, berdasarkan pada atribut yang didapatkan yaitu agen, keterangan barang, jenis, dan jumlah ton akan dilakukan perbandingan algoritma *K-Means* dan *K-Medoids*. Percobaan yang dilakukan menggunakan aplikasi RapidMiner dengan nilai *Davies-Bouldin Index* sebagai acuan pengelompokan *cluster*. Pola pengelompokan *cluster* yang didapatkan kemudian dianalisa dan menjadi sebuah informasi baru bagi perusahaan sebagai penentu tindakan lanjut sesuai analisa. Selain itu, diperoleh hasil algoritma terbaik dari uji validitas yang dilakukan pada *K-Means* dan *K-Medoids*.

B. METODOLOGI PENELITIAN

Berdasarkan pendahuluan yang telah dijabarkan, penelitian ini melalui beberapa tahap, berikut tahapan kerangka kerja dalam penyusunan penelitian:



Gambar 1. Metodologi Penelitian

C. TINJAUAN PUSTAKA

C.1. Data Mining

Data mining merupakan proses mencari pola atau informasi menarik dalam data terpilih atau disebut juga suatu istilah yang digunakan untuk menguraikan penemuan pengetahuan di dalam *database* dengan menggunakan metode atau teknik tertentu. *Data mining* adalah proses menggunakan teknik statistik, matematika, kecerdasan buatan, dan *machine learning* dalam mengidentifikasi dan mengekstraksikan informasi yang bermanfaat bagi pengetahuan yang terkait dengan berbagai basis data berkapasitas besar [5]. Salah satu metode yang terdapat pada *data mining* yaitu *Clustering* yang berperan penting dalam mengeksplorasi data, menciptakan prediksi dan mengatasi anomali data. Pengelompokan dalam suatu *cluster*,

dikelompokkan menggunakan teknik reiteratif/perulangan dengan karakteristik identik suatu *dataset* [6].

C.2. Clustering

Clustering atau pengelompokan data merupakan kegiatan mempertimbangkan sebuah pendekatan penting untuk mencari kesamaan dalam data dan menempatkan data yang sama ke dalam kelompok-kelompok. *Clustering* dianggap sebagai *unsupervised learning* yang paling penting dimana setiap masalah pada jenis ini berkaitan untuk menemukan struktur dalam kumpulan data yang tidak berlabel. Sebuah pengelompokan *cluster* membagi kumpulan data kedalam sejumlah kelompok dimana kesamaan pada sebuah kelompok

tertentu lebih besar daripada kelompok lainnya [7][18].

Penggunaan algoritma *clustering* tergantung pada jenis data yang ada pada tujuan dan aplikasi tertentu. Jika analisis *cluster* digunakan sebagai alat deskriptif atau eksplorasi, maka dimungkinkan untuk mencoba beberapa algoritma pada data yang sama untuk memperoleh apa yang diungkap oleh data tersebut. Secara umum, metode *clustering* dapat diklasifikasikan ke dalam beberapa kategori, salah satunya kategori *partitioning method*. Metode partisi ini didasarkan pada penentuan awal jumlah kelompok, kemudian realokasi objek secara iteratif untuk menemukan kembali kelompok-kelompok yang berada dalam satu titik. Salah satu algoritma yang populer dalam pengaplikasian metode partisi ini yaitu algoritma *K-Means* dan algoritma *K-Medoids* [18].

C.3. K-Means

K-Means clustering ialah metode analisis *cluster* yang bertujuan untuk memecah objek menjadi *k cluster* kemudian diamati di mana setiap objek *cluster* diperoleh melalui rata-rata terdekat. Algoritma ini merupakan salah satu pembelajaran yang terkenal sederhana dan mudah dipelajari sebagai pemecahan masalah pengelompokan dari sebuah *dataset*. Algoritma *K-Means* adalah algoritma evolusi yang mana metode operasinya memiliki artian yang serupa dengan nama algoritmanya. Algoritma ini mengelompokkan pengamatan ke dalam kelompok *k*, di mana *k* merupakan parameter *input*. Tiap data kemudian ditetapkan pada setiap pengamatan *cluster* berdasarkan kedekatan pengamatan nilai rata-rata *cluster*. Nilai rata-rata pada *cluster* kemudian dihitung secara berulang pada proses awal. Adapun tahap-tahap melakukan *K-Means Clustering* sebagai berikut [8]:

1. Pilih nilai *k* sebagai pusat *cluster* awal secara acak.
2. Setiap data dalam *dataset* dibagi kedalam beberapa kelompok *k cluster* antara setiap titik dan pusat *cluster* yang didapatkan berdasarkan pada jarak *Euclidean*.

Adapun rumus perhitungan *Euclidean Distance* ditunjukkan pada persamaan (3.1) berikut [9]:

$$d(x,y) = \|x-y\|^2 = \dots\dots\dots(1)$$

dengan :

- d= jarak data ke pusat *cluster*
- x= data ke pada atribut ke
- y= titik pusat ke, pada atribut

3. Setiap pusat *cluster* dihitung ulang berdasarkan dari nilai rata-rata dalam *cluster* yang didapatkan.
4. Langkah 2 dan 3 ulangi sampai kelompok tersebut sesuai. Perulangan dapat didefinisikan

secara berbeda tergantung pada implementasi, tetapi biasanya pada proses langkah 2 dan 3 dapat diulangi apabila kelompok *cluster* masih mengalami perubahan dan proses akan terhenti apabila tidak ada perubahan antar material dalam *cluster*.

C.4. K-Medoids

Algoritma PAM (*Partitioning Around Medoids*) atau biasa juga disebut dengan algoritma *K-Medoids*, merupakan algoritma yang diwakili oleh *cluster* yaitu *medoid*. Perbedaan antara algoritma *K-Medoids* dengan algoritma *K-Means* yaitu algoritma *K-Medoids* menggunakan objek sebagai perwakilan (*medoid*) pusat *cluster* untuk tiap *cluster*, sementara algoritma *K-Means* membutuhkan nilai rata-rata (*mean*) sebagai pusat *cluster* [17][10].

Langkah-langkah penyelesaian *K-Medoids* adalah sebagai berikut[11]:

1. Inisialisasikan pusat *cluster* sebanyak jumlah *cluster* (*k*).
2. Setiap data atau objek dialokasikan ke *cluster* terdekat menggunakan persamaan ukuran jarak *Euclidian Distance* dengan rumus persamaan (3.1).
3. Pilih objek pada masing-masing *cluster* secara acak sebagai kandidat *medoid* baru.
4. Hitung jarak setiap objek yang terdapat pada masing-masing *cluster* dengan calon *medoid* baru.
5. Hitung total simpangan (*S*) dengan menghitung nilai total jarak baru – total jarak lama. Jika didapatkan $S < 0$, tukarlah objek dengan data *cluster* untuk membuat sekumpulan *k* objek baru sebagai *medoid*.
6. Ulangi langkah 3 sampai dengan 5 hingga tidak terjadi perubahan *medoid*, sehingga diperoleh *cluster* serta anggota *cluster* masing-masing.

C.5. Davies-Bouldin Index (DBI)

Davies-Bouldin Index (DBI) adalah metrik untuk mempertimbangkan/mengevaluasi hasil algoritma *clustering* yang diperkenalkan oleh David L. Davies dan Donald W. Bouldin bertepatan pada tahun 1979 [12]. Dengan menggunakan DBI suatu *cluster* akan dianggap memiliki skema *clustering* yang optimal adalah yang memiliki DBI minimal.

D. HASIL DAN ANALISIS

D.1. Preprocessing Data

Pengumpulan data dilakukan dengan melakukan wawancara dan observasi langsung di PT. Pelabuhan Indonesia I Cabang Dumai. Sumber data utama yaitu berasal dari data Dermaga B pada tahun 2017 sebanyak 3103 *record data*. Data tersebut akan diproses untuk mengasilkan pola rata-rata transaksi bongkar muat curah cair oleh agen yang melakukan transaksi pada Dermaga B. Hal ini dapat memberikan informasi baru kepada

perusahaan sehingga bermanfaat untuk membuat sebuah kebijakan bisnis.

Untuk melakukan proses klusterisasi dibutuhkan beberapa atribut. Dari 11 atribut yang ada pada data dermaga b, diambil 4 atribut yang digunakan dalam proses perhitungan yaitu agen, keterangan barang, jenis bongkar muat, dan jumlah ton. Atribut yang dipilih untuk dijadikan acuan dalam proses pengelompokan yaitu atribut

keterangan barang, jenis barang, dan jumlah ton, sedangkan atribut agen berfungsi untuk kebutuhan analisa pengelompokan.

Cleaning data dilakukan untuk mengurangi *noise* yang dapat mempengaruhi perhitungan. Dalam proses *cleaning data*, data yang tidak lengkap dibuang, sehingga didapatkan sebanyak 3089 *record data*. Data yang telah melakukan proses *cleaning data* dapat dilihat pada tabel 1.

Tabel 1. *Cleaning Data*

No.	Agen	Keterangan Barang	Jenis BM	Jumlah (Ton)
1	Buana Listya Tama Tbk, PT	<i>Crude Palm Oil</i>	Muat	1000
2	Rafina Tirta Samudra, PT	<i>Crude Palm Oil</i>	Muat	939
3	Usda Seroja Jaya, PT	<i>Crude Palm Oil</i>	Muat	5175
4	Bahari Sandi Pratama, PT	<i>Crude Palm Oil</i>	Bongkar	3501
5	Snepac Shipping, PT	<i>Crude Palm Oil</i>	Bongkar	3501
....
362	Snepac Shipping, PT	<i>RBD Palm Olien</i>	Muat	10000
363	Snepac Shipping, PT	<i>RBD Palm Stearin</i>	Muat	1500
....
3089	Rich Sino, MT	<i>Crude Palm Oil</i>	Muat	1000

Setelah dilakukannya *Cleaning data*, selanjutnya melakukan transformasi data. Data yang ada ditransformasi kedalam bentuk simbol untuk memudahkan dalam melakukan perhitungan nantinya. Penetapan perubahan data kedalam bentuk simbol dapat dilihat pada tabel 2 sampai dengan tabel 4.

Tabel 2. Transformasi Agen

No.	Agen	Transfor masi
1	Bahari Eka Nusantara, PT	A1
2	Bahari Sandi Pratama, PT	A2
3	Barakomindo Shipping, PT	A3
4	Bintang Samudera Utama, PT	A4
5	Buana Listya Tama Tbk, PT	A5
6	Cahaya Lautan Niaga, PT	A6
7	Dahlia Bina Utama, PT	A7
8	Gemalindo Shipping, PT	A8
9	Jatarim Binau Line, PT	A9
10	Kanaka Dwimitra Manunggal, PT	A10
11	Pacific Samudera Sentosa, PT	A11
12	Pelayaran Cahaya Papua, PT	A12
13	PT. Mega Tama Mestika	A13
14	Putra Segara Abadi, PT	A14
15	Rafina Tirta Samudra, PT	A15
16	Rizkia Armada Nusantara, PT	A16
17	Samudera Anambas, PT	A17
18	Samudera Sarana Karunia, PT	A18
19	Sejahtera Mandiri Lines, PT	A19
20	Snepac Shipping, PT	A20
21	Spectra Tirtasegara Line, PT	A21
22	Tri Daya Laju, PT	A22
23	Usda Seroja Jaya, PT	A23
24	Winde Ship Centralindo, PT	A24

Tabel 3. Transformasi Keterangan Barang

No	Keterangan Barang	Transfor masi
1	Bahan Bakar	1
2	<i>Crude Glycerine</i>	2
3	<i>Crude Palm Kernel Oil</i>	3
4	<i>Crude Palm Oil</i>	4
5	<i>Crude Palm Olien</i>	5
6	<i>Crude Palm Stearin</i>	6
7	<i>Fatty Acid Methyl Ester</i>	7
8	<i>Palm Acid Oil</i>	8
9	<i>Palm Fatty Acid Distillate</i>	9
10	<i>Palmata</i>	10
11	<i>Palmitic Acid</i>	11
12	<i>RBD Coconut Oil</i>	12
13	<i>RBD Olien</i>	13
14	<i>RBD Palm Kernel Oil</i>	14
15	<i>RBD Palm Oil</i>	15
16	<i>RBD Palm Olien</i>	16
17	<i>RBD Palm Stearin</i>	17
18	<i>Soft Stearin</i>	18
19	<i>Split Palm Kernel Fatty Acid</i>	19

Tabel 4. Transformasi Jenis Bongkar Muat

No	Jenis Bongkar Muat	Transfor masi
1	Bongkar	1
2	Muat	2

Berdasarkan atribut pada tabel 2 sampai dengan 4 maka dapat diperoleh transformasi secara keseluruhan, yang ditunjukkan pada tabel 5:

Tabel 5. Normalisasi Data

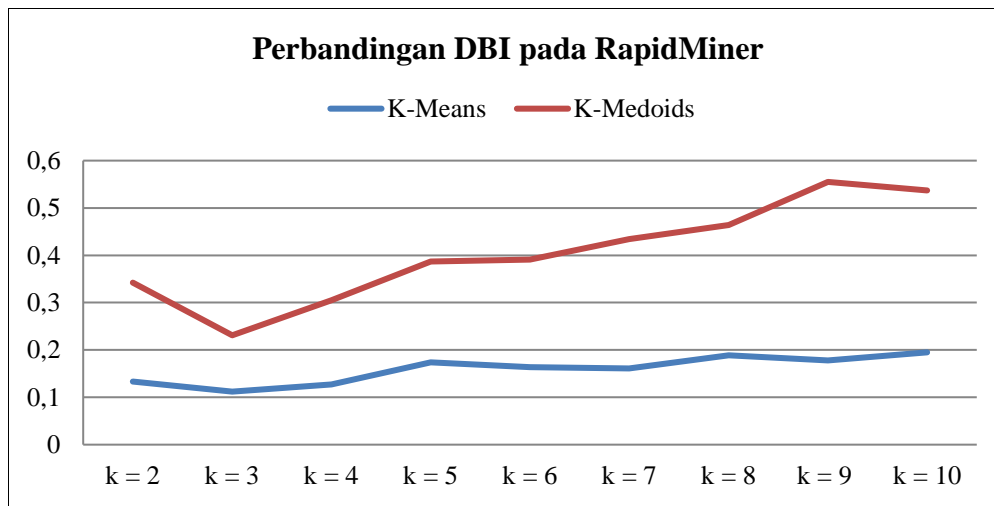
No.	Agen	Keterangan Barang	Jenis BM	Jumlah (Ton)
-----	------	-------------------	----------	--------------

1	A5	0,167	1,000	0,055
2	A15	0,167	1,000	0,052
3	A23	0,167	1,000	0,287
4	A2	0,167	0,000	0,194
5	A20	0,167	0,000	0,194
....
362	A20	0,833	1,000	0,555
363	A20	0,889	1,000	0,083
....
3089	A5	0,167	1,000	0,055

D.2. Proses Data

Setelah melakukan tahapan *preprocessing data*, data kemudian diolah menggunakan RapidMiner untuk menunjukkan nilai *Davies-Bouldin Index* (DBI) sebagai acuan

pengelompokan *cluster* terbaik. Pengujian dilakukan pada *cluster* k =2 sampai dengan k=10. Berikut hasil pengolahan data ditunjukkan pada gambar 2:



Gambar 2. Perbandingan DBI pada RapidMiner

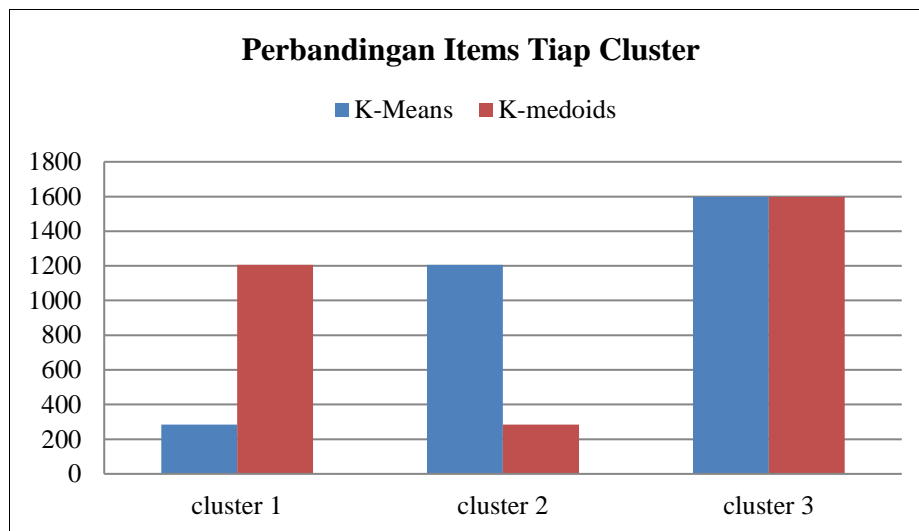
Untuk melihat secara rinci pengujian tiap *cluster* pada masing-masing algoritma, ditunjukkan pada tabel 6:

Tabel 6. Hasil Perbandingan DBI

Cluster	K-Means	K-Medoids
k = 2	0.133	0.209
k = 3	0.112	0.119
k = 4	0.127	0.178
k = 5	0.174	0.213
k = 6	0.164	0.227
k = 7	0.161	0.273

Cluster	K-Means	K-Medoids
k = 8	0.189	0.275
k = 9	0.178	0.377
k = 10	0.195	0.342

Berdasarkan perbandingan nilai *Davies-Bouldin Index* diatas, menunjukkan bahwa k=3 menjadi *cluster* terbaik. Untuk melihat pembagian tiap *items* pada masing-masing pengelompokan k=3 dapat ditunjukkan pada gambar 3.



Gambar 3. Perbandingan *Items* Tiap *Cluster*

Diperoleh pola pengelompokan tiap *cluster* hanya pada satu jenis atribut bongkar barang atau muat barang, dimana pada *cluster* dominan didapatkan atribut berjenis muat barang dengan transaksi CPO (*Crude Palm Oil*) yang paling dominan. Pada *cluster 3* ini juga menunjukkan agen Buana Listya Tama TBK, PT mendominasi sebanyak 18,77% dengan jumlah rata-rata transaksi 2847,22 ton. Selanjutnya oleh agen Samudera Sarana Karunia, PT sebanyak 17,02% dengan jumlah rata-rata transaksi 2827,32 ton.

E. KESIMPULAN

Perbandingan antar kedua algoritma *clustering*, *K-Medoids* dan *K-Means* tidak menunjukkan perbedaan yang signifikan mengenai pengelompokan data. Adapun dalam pengolahannya, *K-Means* hanya membutuhkan waktu rata-rata 1 detik sedangkan pengolahan data pada *K-Medoids* membutuhkan waktu rata-rata 1 menit 38 detik yang artinya apabila makin tinggi iterasi dan pengelompokan yang ditentukan, maka pengolahan data akan semakin lama. Selanjutnya pada nilai *Davies-Bouldin Index* diperoleh nilai $k=3$ pada *K-Means* lebih rendah dibandingkan *K-Medoids*. Selain itu, pengolahan data berlanjut pada $k=10$ juga menunjukkan nilai DBI pada *K-Means* lebih rendah. Pola pengelompokan masing-masing *cluster* diperoleh berdasarkan pembagian jenis barang sesuai jenis bongkar atau muat barang. Perolehan *cluster* dominan, menunjukkan bahwa agen Buana Listya Tama TBK, PT mendominasi diikuti dengan agen Samudera Sarana Karunia, PT. sebagai agen yang melakukan transaksi terbanyak.

F. UCAPAN TERIMA KASIH

Terima kasih yang sebesar-besarnya kepada Tim Puzzle Research Data Technology (PREDATECH) Fakultas Sains dan Teknologi Universitas Islam Negeri Sultan Syarif Kasim Riau

atas segala masukan, dukungan moral dan materi, koreksi dan bantuan dalam kegiatan penelitian ini sehingga penelitian ini dapat terselesaikan dengan baik

REFERENSI

- [1] Okfalisa, Mustakim, Gazalba I & Reza NGI . *Comparative Analysis Of K-Nearest Neighbor And Modified K-Nearest Neighbor Algorithm For Data Classification*. Information Technology, Information Systems and Electrical Engineering (ICITISEE), 2017 2nd International conferences on IEEE. 2017 : 294-298.
- [2] Uddin H & Nafis MT. *Students Academic Performance Using Partitioning Clustering Algorithms*. International Journal of Advanced Research in Computer Science. 2017 ; 8(5) : 640-644.
- [3] Arbin N, Suhaimi, NS, Mokhtar, NZ & Othman Z. *Comparative Analysis between K-Means and K-Medoids for Statistical Clustering*. 2015 3rd International Conference on Artificial Intelligence, Modelling and Simulation. IEEE. 2015: 117-121.
- [4] Balabantaray RC, Sarma C & Jha M. *Document clustering using K-Means and K-Medoids*. International Journal of Knowledge Based Computer System. 2013 ; 1(1) : 7-13.
- [5] Han J, Kamber M, dan Jian P. *Data Mining : Concepts and techniques*. USA : Elsevier Inc. 2012.
- [6] Arora P, Dr Deepali, Varshney S. *Analysis of K-Means and K-Medoids Algorithm For Big Data*. International Conference on Information Security & Privacy (ICISP2015). Nagpur. 2016; 507-512.

- [7] RuiXu & Donald, C.W. 2009. *Clustering*. John Wiley & Sons, Inc.
- [8] Singh K, Malik D & Naveen Sharma. *Evolving Limitations in K-Means Algorithm in Data Mining and Their Removal*. IJCEM International Journal of Computational Engineering & Management. 2011; 12: 105-109.
- [9] Ramadhan A, Efendi Z, & Mustakim. Perbandingan K-Means dan Fuzzy C-Means untuk Pengelompokan Data User Knowledge Modeling. *Seminar Nasional Teknologi Informasi Komunikasi dan Industri*. 2017; 219-226.
- [10] Kaur, Noor K et al. "K-Medoids Clustering Algorithm". Vol. 1, Issue 1. April 2014.
- [11] Han J, Kamber M. *Data Mining: Concept and Techniques*. Waltham: Morgan Kauffman Publisher; 2006.
- [12] Irhamni F, Damayanti F, Khusnul K B & A Miftachul. *Optimalisasi Pengelompokan Kecamatan Berdasarkan Indikator Pendidikan Menggunakan Metode Clustering Dan Davies Bouldin Index*. Seminar Nasional Sains dan Teknologi 2014. Jakarta. 2014 : 1-5.
- [13] Wu, Xindong & Kumar, Vipin. 2009. *The Top Ten Algorithms in Data Mining*. London: CRC Press.
- [14] Triyanto WA. *Algoritma K-Medoids Untuk Penentuan Strategi Pemasaran Produk*. Jurnal Teknik Industri, Mesin, Elektro, dan Ilmu Komputer. 2015; 6(1):183-188
- [15] Aryuni M, Madyatmadja ED, Miranda E. *Penerapan K-Means Dan K-Medoids Clustering Pada Data Internet Banking Di Bank Xyz*. Jurnal Teknik dan Ilmu Komputer. 2018 ; 7(27) : 349-356.
- [16] Diakses oleh <http://ekonomi.metrotvnews.com> pada Minggu, 22 Oct 2017.
- [17] IBRAHIM, L. Fattouh. *Using Of Clustering And Ant-Colony Algorithms CWSP-PAM-ANT In Network Planning*. In: *Digital Telecommunications,, 2006. ICDT'06. International Conference on*. IEEE, 2006. p. 63-63.
- [18] Madhulatha TS. *Comparison Between K-Means and K-Medoids Clustering Algorithms*. International Journal of Advanced Computing (IJAC). 2011; 3(2) : 65-69.