

Potensi Pelanggan Tunggal PDAM Menggunakan Metode *K-Medoids* dengan Optimasi *Ant Colony Optimization* (ACO)

Hardiyusa¹, Made Sudarma², Nyoman Pramaita³

Abstract—PDAM in carrying out operational activities is greatly influenced by the receivables or arrears of customer water bills. Some factors that influence customer patterns in delinquent water bills are customer class and consumption of water usage, which affects the water bill paid by the customer. This study will apply the *K-Medoids* clustering method to find out customers who are delinquent in the PDAM by optimizing the selection of cluster centers using the *Ant Colony Optimization* (ACO) algorithm. In this study the combination of ACO and *K-Medoids* methods is called *ACOMedoids*. The results with the *ACOMedoids* method can produce a high level of accuracy from the comparison of clustering data with actual bill data. This can be seen from the results of accuracy which is always better than the *K-Medoids* method, which is the highest achieves 97.65% accuracy for *ACOMedoids* while *K-Medoids* is 88.29%. Accuracy results show that the ACO algorithm can produce optimal cluster center points in the clustering process of the *K-Medoids* method.

Intisari—PDAM dalam menjalankan kegiatan operasional sangat dipengaruhi oleh piutang atau tunggakan tagihan air pelanggan. Beberapa faktor yang mempengaruhi pola pelanggan dalam menunggak tagihan air yaitu golongan pelanggan dan konsumsi pemakaian air sehingga mempengaruhi tagihan air yang dibayar oleh pelanggan. Penelitian ini akan diterapkan metode *clustering K-Medoids* untuk mengetahui pelanggan yang menunggak di PDAM dengan melakukan optimasi pada pemilihan titik pusat *cluster* menggunakan algoritma *Ant Colony Optimization* (ACO). Dalam penelitian ini penggabungan metode ACO dan *K-Medoids* disebut *ACOMedoids*. Hasil dengan metode *ACOMedoids* dapat menghasilkan tingkat akurasi yang tinggi dari perbandingan data hasil *clustering* dengan data tagihan aktual. Hal ini terlihat pada hasil akurasi yang selalu lebih baik dibandingkan dengan metode *K-Medoids* yaitu tertinggi mencapai akurasi 97,65% untuk *ACOMedoids* sedangkan *K-Medoids* 88,29%. Hasil akurasi menunjukkan algoritma ACO dapat menghasilkan titik pusat *cluster* yang optimal pada proses *clustering* metode *K-Medoids*.

Kata Kunci—PDAM, Tunggakan, *Clustering*, *K-Medoids*, ACO, *ACOMedoids*.

I. PENDAHULUAN

Perusahaan Daerah Air Minum (PDAM) merupakan badan usaha milik daerah yang dibentuk oleh Pemerintah Daerah

Provinsi, Kabupaten dan atau Kota yang bergerak di bidang pelayanan air minum [1]. Dalam pengelolaannya, PDAM mengalami kendala dan hambatan dalam melakukan kegiatan operasional di antaranya, seperti keterbatasan jumlah SDM dan kompetensi, efektivitas penagihan, efektivitas produksi, kehilangan air, jumlah tunggakan pelanggan yang besar dan lain-lain yang menimbulkan kerugian bagi PDAM [1]. Besarnya piutang yang tidak dapat di tagihkan akan mempengaruhi kelancaran operasional suatu PDAM, sehingga berdampak kepada pelanggan PDAM yang tidak mendapatkan pelayanan secara maksimal. Beberapa faktor yang mempengaruhi pola pelanggan dalam menunggak tagihan air adalah golongan dari pelanggan dan konsumsi pemakaian air sehingga mempengaruhi tagihan air yang dibayar oleh pelanggan. Pemanfaatan faktor tersebut dapat digunakan untuk proses pengelompokan pelanggan sesuai dengan polanya masing-masing, sehingga dapat diketahui potensi pelanggan yang akan menunggak.

Segmentasi atau pengelompokan pelanggan pada suatu perusahaan dapat membantu perusahaan tersebut dalam membuat keputusan terhadap kelompok pelanggan tertentu sesuai kebijakan suatu perusahaan. Beberapa penelitian mengenai analisis pelanggan antara lain penelitian mengenai pengelompokan pelanggan berdasarkan pola konsumsi pemakaian air terhadap kebutuhan air perkotaan menggunakan metode *clustering Kohonen* atau *Self Organized Maps* (SOM) [2]. Penelitian tersebut menghasilkan kelompok data pelanggan terhadap pola konsumsi air untuk kategori rumah tangga dan non-perusahaan. Hasil *cluster* digunakan dalam optimalisasi pengelompokan pada pelanggan baru berdasarkan pada kriteria atau faktor yang telah ditentukan. Penelitian berikutnya mengenai segmentasi dan klasifikasi perilaku pembayaran nasabah pada perusahaan penyedia jasa multimedia dengan algoritma *K-Means* dan C4.5 [3]. Dalam penelitian tersebut tingkat potensi nasabah juga bisa menjadi acuan dalam promosi, retensi, dan pencegahan pelanggan menunggak. Penelitian selanjutnya mengenai analisa dalam memprediksi pelanggan apakah mereka akan meninggalkan perusahaan atau tidak, penelitian tersebut di lakukan pada industri telekomunikasi [4] dengan membandingkan metode *decision tree* J-48 dengan teknik regresi logistik. Penelitian tersebut digunakan untuk mencari potensi pelanggan yang berhenti berlangganan (*churn*), sehingga dapat menghindari kerugian yang lebih besar di alami oleh perusahaan. Penelitian berikutnya yaitu mengembangkan model prediksi risiko pinjaman pada bank dengan menggunakan data *mining* [5]. Tiga algoritma klasifikasi yang digunakan yaitu *J-48*, *bayesNet* dan *naiveBayes*. Hasil dari penelitian tersebut adalah prediksi nasabah bank itu baik atau tidak dalam melakukan pinjaman.

p-ISSN:1693 – 2951; e-ISSN: 2503-2372

¹Mahasiswa Magister Teknik Elektro Fakultas Teknik Universitas Udayana Jalan Kampus Bukit Jimbaran 80361 INDONESIA tel: 0361-703315; fax: 0361-703315 e-mail: hardiyusa.bss@gmail.com

^{2,3}Dosen Jurusan Teknik Elektror Fakultas Teknik Universitas Udayana, Jalan Kampus Bukit Jimbaran 80361 INDONESIA tel: 0361-703315; fax: 0361-703315; e-mail: msudarma@unud.ac.id, n_pramaita@yahoo.com

Hardiyusa: Potensi Pelanggan Tunggal PDAM ...



Berdasarkan beberapa penelitian mengenai potensi pelanggan, serta memperhitungkan aspek keuangan dalam hal efektivitas penagihan PDAM, maka dikembangkan sebuah ide untuk melakukan penelitian mengenai potensi pelanggan tunggakan pada PDAM dengan menggunakan metode *clustering K-Medoids*. Dalam implementasinya *K-Medoids* memang lebih baik dibandingkan *K-Means*, tetapi metode ini memiliki kelemahan yaitu kompleksitas komputasinya yang tinggi sehingga berdampak pada performa proses *clustering* secara keseluruhan. Selain itu, proses pemilihan pusat *cluster* awal secara acak membuat hasil dari proses *clustering* menjadi tidak stabil, sehingga *K-Medoids* juga belum dapat mengatasi masalah lokal optimal. Oleh karena itu diperlukan suatu algoritma optimasi yang dapat meningkatkan performa serta mengatasi kelemahan yang ada pada metode *K-Medoids* tradisional.

Beberapa penelitian mengenai optimasi metode *clustering* antara lain penelitian mengenai pengelompokan (*clustering*) data dengan teknik *k-medoids* yang digabungkan dengan algoritma *Bat* (kelelawar) [6]. Algoritma *Bat* digunakan untuk efisiensi dalam menentukan titik pusat *cluster* yang lebih baik. Penggabungan kedua metode tersebut dapat menghasilkan pengelompokan analisis data yang lebih baik. Penelitian berikut mengenai penentuan kompetensi mahasiswa dengan algoritma genetika dan metode *Fuzzy C-Means* [7]. Hasil percobaan yang dilakukan, penggabungan Algoritma Genetik dan *Fuzzy C-Means* memberikan hasil yang lebih baik dibandingkan *Fuzzy C-Means* saja dengan persentase rata-rata kesesuaian pada pengujian yang dilakukan adalah sebesar 88.89%. Penelitian selanjutnya mengenai optimasi pusat *cluster K-Prototype* dengan algoritma genetika [8]. Dari beberapa hasil percobaan yang dilakukan metode *K-Prototype* dengan Algoritma Genetika menghasilkan hasil yang terbaik dari metode *K-Prototype* tanpa Algoritma Genetika dan metode *K-Means*.

Berdasarkan latar belakang tersebut, maka dalam penelitian ini diperlukan suatu algoritma optimasi yang dapat meningkatkan performa serta mengatasi kelemahan yang ada pada metode *K-Medoids* tradisional. *Ant Colony Optimization* (ACO) merupakan suatu algoritma yang didasarkan oleh pola semut di dalam membentuk suatu koloni dan mencari makanan [9]. Algoritma ACO akan digunakan untuk optimasi pada pemilihan titik pusat *cluster* sebelum dilakukan proses *clustering K-Medoids*. Penggabungan metode *clustering* dan algoritma ACO dalam penelitian ini disebut *ACOMedoids*. Berdasarkan hasil *clustering* penelitian ini dihitung nilai akurasi dengan menggunakan metode *confusion matrix*.

II. ALGORITMA DAN METODE PENGUJIAN

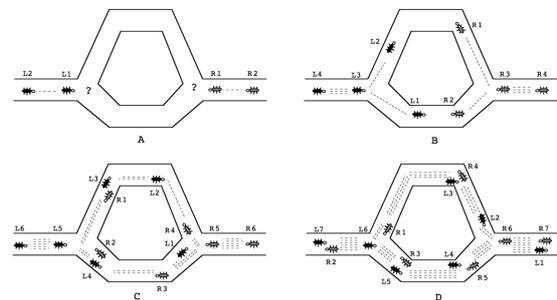
A. *K-Medoids*

Strategi dasar dari algoritma *clustering K-Medoids* adalah untuk menemukan k *cluster* dalam n objek dengan pertama kali secara *arbitrarily* menemukan wakil dari objek (*medoid*) untuk tiap-tiap *cluster* [10]. Masing-masing sisa objek di *cluster* dengan *medoids* ke yang paling mirip. Strategi ini kemudian secara iteratif menggantikan satu *medoids* dari yang non *medoids* sepanjang kualitas dari hasil *clustering*

ditingkatkan. PAM (*partition around medoids*) adalah jenis *K-Medoids* algoritma *clustering*. Pada PAM ditemukan k *cluster* di objek n dengan terlebih dahulu mencari benda perwakilan (*medoid*) untuk setiap *cluster*. Set awal *medoids* dapat dipilih sesuai keinginan. Kemudian proses iteratif menggantikan salah satu *medoids* oleh salah satu non-*medoids* lainnya selama total jarak pengelompokan yang dihasilkan mengalami peningkatan.

B. *Ant Colony Optimization* (ACO)

Ant Colony Optimization (ACO) termasuk dalam kelompok *swarm intelligence*, yang merupakan salah satu jenis pengembangan paradigma yang digunakan untuk menyelesaikan masalah optimasi dimana inspirasi yang digunakan untuk memecahkan masalah tersebut berasal dari perilaku kumpulan atau kawanan (*swarm*) serangga [11]. *Ant-based techniques* pertama kali digunakan oleh Marco Dorigo dan Luca Maria Gambardella dengan menggunakan ACO untuk menyelesaikan *Traveling Salesman Problem* (TSP) [9].



Gambar 1: Koloni semut mencari makan [9]

Algoritma *Ant Colony Optimization* berdasarkan pada tingkah laku semut dalam mencari makanan. Semut, dalam perjalanan mencari makanan, akan mengeluarkan aroma kimia yang dikenal sebagai feromon. Feromon yang dikeluarkan berfungsi sebagai jejak untuk semut lainnya agar dapat menemukan sumber makanan yang sudah ditemukan oleh semut sebelumnya dan setiap semut akan meninggalkan feromon sebagai jejak juga untuk dirinya dan semut lainnya. Semut yang melewati jalur terpendek dan terdekat akan meninggalkan feromon yang lebih kuat untuk dicium oleh semut lainnya. Semut lainnya, secara alami akan mengikuti jejak feromon yang terkuat atau terpendek. Berikut adalah langkah-langkah semut dalam mencari makanan:

1. Semut-semut akan mencari sumber makanan secara acak melalui jalur yang memungkinkan.
2. Saat menemukan makanan mereka kembali ke sarangnya dengan memberikan tanda dengan jejak feromon.
3. Semut-semut yang lain tidak akan mencari makanan secara acak kembali melainkan akan mengikuti atau mengambil jalur dengan feromon terkuat.
4. Akhirnya semut akan mengambil jalur terpendek, di mana jalur lainnya akan semakin kehilangan feromon yang telah menguap.

C. *Pengujian Confusion Matrix*

Proses pengujian dilakukan dengan metode *confusion matrix*. *Confusion matrix* adalah suatu metode yang digunakan

untuk melakukan perhitungan akurasi pada konsep *data mining* [12]. Proses pengujian dengan *confusion matrix* dilakukan pada hasil metode *clustering K-Medoids* tanpa menggunakan algoritma optimasi dan hasil metode *clustering K-Medoids* dengan menggunakan algoritma optimasi ACO.

TABEL I
MODEL *CONFUSION MATRIX* POTENSI PELANGGAN TUNGGAKAN

Kelas		Hasil Klasifikasi	
		Menunggak	Lancar
Hasil Aktual	Menunggak	<i>True Positive</i> (TP)	<i>False Negatives</i> (FN)
	Lancar	<i>False Positives</i> (FP)	<i>True Negatives</i> (TN)

TP = Jumlah data jika hasil klasifikasi dari sistem menunggak (positif) dan jika hasil aktual juga menunggak (positif).

FG = Jumlah data jika hasil klasifikasi dari sistem lancar (negatif) dan jika hasil aktual menunggak (positif).

FP = Jumlah data jika hasil klasifikasi dari sistem menunggak (positif) dan jika hasil aktual juga lancar (negatif).

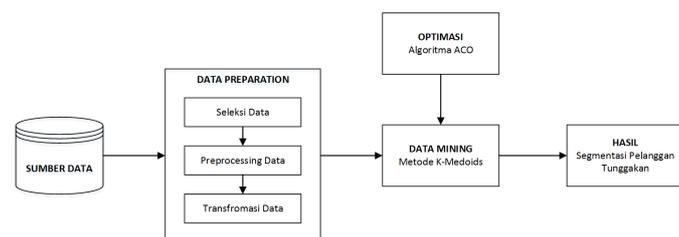
TN = Jumlah data jika hasil klasifikasi dari sistem lancar (negatif) dan jika hasil aktual juga lancar (negatif).

Pengujian dengan *confusion matrix* berdasarkan tabel di atas, nilai dari TP, FN, FP dan TN dapat menghasilkan nilai akurasi. Nilai akurasi adalah persentase ketepatan *record* data yang diklasifikasikan secara benar setelah dilakukan pengujian pada hasil klasifikasi. Dengan kata lain, nilai akurasi merupakan perbandingan antara data yang terklasifikasi benar dengan keseluruhan data.

$$Akurasi = \frac{TP+TN}{TP+TN+FP+FN} \times 100\% \quad (1)$$

III. PROSES DATA MINING

Sistem yang akan dikembangkan dalam penelitian ini adalah penerapan *data mining* untuk mengetahui potensi pelanggan tunggakan PDAM dengan metode *clustering* yaitu *K-Medoids* yang akan di optimasi dalam pencarian titik pusat *cluster* dengan algoritma ACO. Sistem yang akan dikembangkan terdiri dari dua proses utama yaitu proses *data preparation* dan proses *data mining*. Gambar 2 menunjukkan gambaran umum dari sistem yang akan dikembangkan.



Gambar 2: Gambaran umum sistem

A. Sumber Data

Sumber data dalam penelitian ini merupakan data tagihan pelanggan PDAM yang berasal dari *database* sistem *billing* PDAM milik PT. Bima Sakti Sanjaya. PT. Bima Sakti Sanjaya adalah perusahaan yang menjadi mitra PDAM dalam hal sistem informasi manajemen yang bertempat di A. Yani Utara Peguyangan, Denpasar, Bali. PDAM yang digunakan dalam penelitian ini sebanyak 6 PDAM dengan periode tagihan adalah 3 tahun yaitu dari periode Agustus 2014 sampai dengan periode Agustus 2017.

B. Data preparation

Sumber data yang telah didapatkan selanjutnya akan diolah menjadi dataset akhir atau data yang akan digunakan dalam proses pada tahap *data mining*. Pada tahapan ini mencakup pemilihan tabel, *record*, serta atribut-atribut data, termasuk di dalamnya proses pembersihan dan transformasi data untuk kemudian dijadikan masukan dalam tahap *data mining*. Tahap ini dapat diulang beberapa kali untuk mendapatkan hasil yang lebih baik sesuai dengan kebutuhan. Proses *data preparation* terdiri dari 3 tahap, yaitu tahap seleksi data, tahap *preprocessing* data dan transformasi data.

1. Proses seleksi data dilakukan dengan cara memilih atribut-atribut pada tabel tersebut yang diperlukan saat proses *data mining*, yaitu atribut periode, nosamb, nama, tarif, pakai, total, dan flaglunas.
2. *Preprocessing* dilakukan proses pembersihan (*data cleansing*) sehingga mendapatkan data tanpa adanya *data noise* dan *missing value*, yaitu atribut pakai dan total tidak diperbolehkan bernilai 0.
3. Transformasi data, dalam penelitian ini akan di gunakan tipe titik 3 dimensi sehingga atribut-atribut yang akan digunakan akan disesuaikan kembali menjadi atribut pakai, tarif dan total.

Jumlah hasil *data preparation* untuk masing-masing PDAM dapat dilihat pada Tabel II.

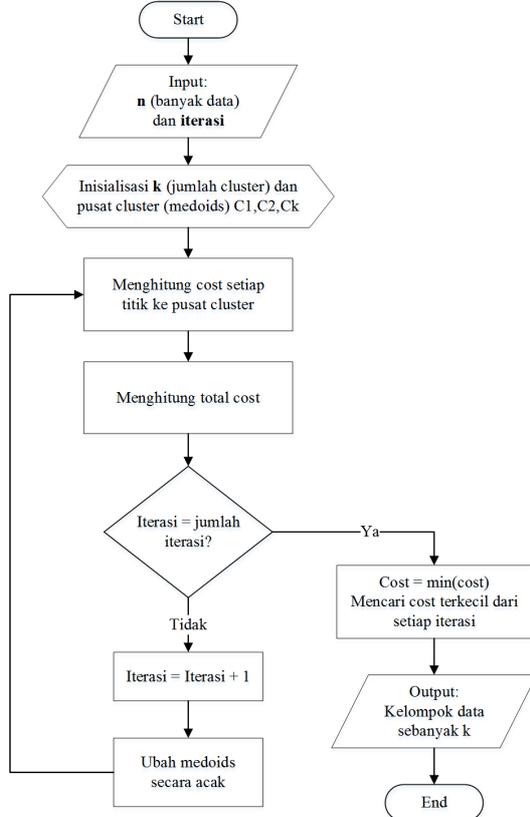
TABEL II
JUMLAH HASIL *DATA PREPARATION* MASING-MASING PDAM

No	Nama PDAM	Pelanggan	Data Rekening	Data Preparation
1	PDAM Kab. Timor Tengah Utara	4.025	70.781	44.698
2	PDAM Kab. Buru	4.099	46.025	34.581
3	PDAM Kota Pekanbaru	12.309	193.846	88.340
4	PDAM Kab. Polewali Mandar	15.912	471.353	391.342
5	PDAM Kab. Klungkung	29.919	807.442	763.100
6	PDAM Kota Tangerang	44.389	1.083.695	896.368



C. Metode K-Medoids

Teknik *data mining* yang digunakan dalam penelitian ini adalah *clustering* dengan metode *K-Medoids*. Pada tahap ini data yang telah disiapkan pada tahap sebelumnya *data preparation* dilakukan pengolahan untuk mendapatkan beberapa kelompok data (*cluster*). Gambar 3 menunjukkan bagan alir dari proses *data mining* dengan metode *K-Medoids*

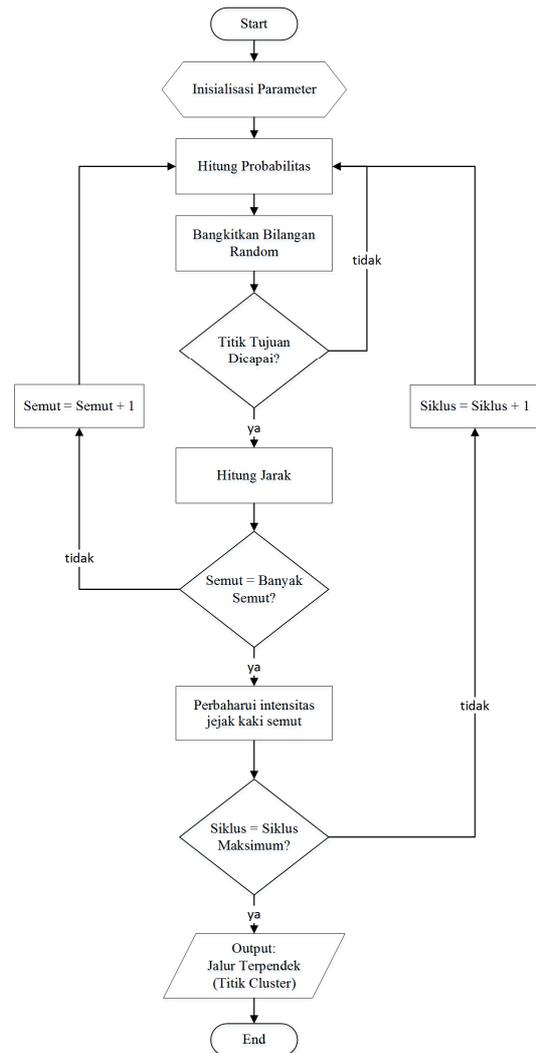


Gambar 3: Flowchart algoritma *K-Medoids*

Pembentukan kelompok data pelanggan sesuai dengan tahapan seperti pada *flowchart* metode *K-Medoids* diatas dapat dijelaskan sebagai berikut:

1. Masukkan jumlah data n yang akan di proses dan jumlah iterasi yang akan dilakukan.
2. Menentukan jumlah *cluster* k dan pusat *cluster* (*medoids*) untuk masing-masing *cluster* dari inputan data n .
3. Menghitung *cost* setiap titik data ke pusat *cluster* (*medoids*) menggunakan rumus *Minkowski*.
4. Menghitung total *cost* dari *cluster* yang sudah terbentuk
5. Mengecek iterasi yang telah dilakukan sama dengan jumlah iterasi yang telah ditentukan.
6. Jika ya pilih nilai total *cost* yang terkecil dari total *cost* yang telah dilakukan selama iterasi. Data hasil *cluster* dengan total *cost* yang terkecil tersebut akan menjadi kelompok data yang akan digunakan untuk proses selanjutnya.
7. Jika tidak tambahkan iterasi dan ubah *medoids* secara acak.
8. Ulangi langkah 3 sampai 8 sampai tidak ada perubahan dalam *medoids* tersebut.

D. Optimasi Titik Cluster K-Medoids dengan ACO



Gambar 4: Flowchart algoritma ACO

Pada tahap ini data hasil *data preparation* dilakukan pencarian titik cluster sebelum tahap *clustering* *K-Medoids* dilakukan. Gambar 4 menunjukkan bagan alir dari proses algoritma ACO.

IV. HASIL DAN PEMBAHASAN

A. Proses Data mining

Pembentukan *cluster* data pelanggan dengan menggunakan metode *K-Medoids* pada penelitian ini terdiri dari beberapa parameter yaitu jumlah data yang digunakan sesuai dengan jumlah data hasil *data preparation*, jumlah iterasi yaitu 10 kali, jumlah *cluster* yaitu 2 *cluster* (menunggak dan lancar), perhitungan jarak antar titik dengan rumus *Minkowski*. Hasil proses *data mining* dengan menggunakan metode *K-Medoids* dapat dilihat pada Tabel IV.

Metode *ACOMedoids* adalah metode penggabungan antara algoritma ACO yang digunakan untuk membentuk atau menentukan pusat *cluster* dengan metode *K-Medoids* dalam pembentukan *clustering* berdasarkan pusat *cluster* tersebut.

Tahap pertama menentukan titik pusat *cluster* dengan algoritma ACO digunakan beberapa parameter yaitu jumlah iterasi sebanyak 3 kali, nilai α digunakan 1, nilai β adalah 1, nilai ρ (ρ) 0.5 dan jumlah semut yang digunakan adalah 3.

Tahap kedua yaitu hasil pusat *cluster* dari hasil algoritma ACO digunakan sebagai pusat *cluster* pada metode *clustering K-Medoids* untuk mendapatkan *cluster* sebagai data latih.

Proses pembentukan *cluster* data pelanggan pada metode *K-Medoids* digunakan beberapa parameter yaitu jumlah data yang digunakan sesuai dengan jumlah data hasil *data preparation*, jumlah iterasi yaitu 3 kali, jumlah *cluster* yaitu 2 *cluster* (menunggak dan lancar), perhitungan jarak antar titik dengan rumus *Minkowski*. Hasil proses *data mining* dengan menggunakan metode *ACOMedoids* terlihat pada Tabel IV.

TABEL IV
 HASIL PROSES METODE *K-MEDOIDS* DAN METODE *ACOMEDOIDS*

No	Nama PDAM	<i>K-Medoids</i>			<i>ACOMedoids</i>		
		Waktu (detik)	Pusat Cluster 1	Pusat Cluster 2	Waktu (detik)	Pusat Cluster 1	Pusat Cluster 2
1	PDAM Kab. Timor Tengah Utara	1,88	3400,1,37000	2600,62,86200	12,05	3400,20,30000	1000,95,76000
2	PDAM Kab. Buru	1,43	1580,37,122218	2080,79,135500	13,00	1580,10,29400	2080,59,77960
3	PDAM Kota Pekanbaru	3,70	7700,26,101100	2800,70,99900	19,70	2300,76,310600	3500,57,74000
4	PDAM Kab. Polewali Mandar	16,38	15000,13,30300	15500,54,77500	21,21	15000,4,22500	15500,66,88100
5	PDAM Kab. Klungkung	34,66	1400,29,82600	1900,74,115600	47,90	1400,29,82600	1900,85,153000
6	PDAM Kota Tangerang	39,83	2775,27,126375	3275,92,264125	57,19	2775,5,17375	3275,71,136875

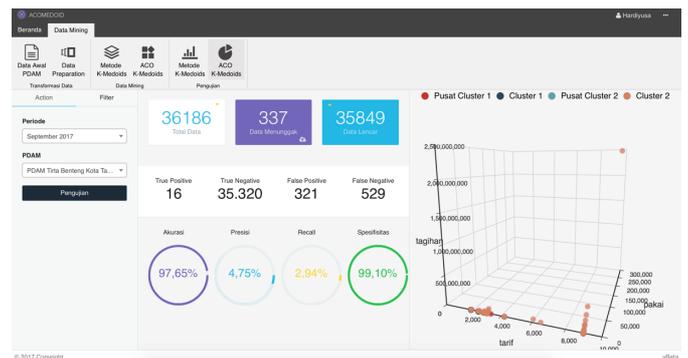
B. Hasil Pengujian

Pengujian pada metode *K-Medoids* menggunakan datalatih hasil *cluster* pada proses *data mining* dengan metode *K-Medoids* sebelumnya. Data aktual (data transaksi pembayaran) yang digunakan dalam pengujian adalah periode tagihan September 2017 dengan transaksi pembayaran bulan Oktober 2017. Pengujian langsung menggunakan aplikasi dan terlihat pada Gambar 5 dan Tabel V menunjukkan hasil pengujian untuk masing-masing PDAM.

Pengujian pada metode *ACOMedoids* digunakan datalatih dan data aktual sama seperti dengan pengujian pada metode *K-Medoids*. Pengujian langsung menggunakan aplikasi dan terlihat pada Gambar 6 dan Tabel V menunjukkan hasil pengujian untuk masing-masing PDAM.



Gambar 5: Hasil pengujian metode *K-Medoids*



Gambar 6: Hasil pengujian metode *ACOMedoids*

Tabel V terlihat nilai akurasi terus naik sesuai jumlah data latih yang digunakan semakin besar data latih, maka semakin besar tingkat akurasi yang didapatkan.

TABEL V
 HASIL PENGUJIAN AKURASI METODE *K-MEDOIDS* DAN METODE *ACOMEDOIDS*

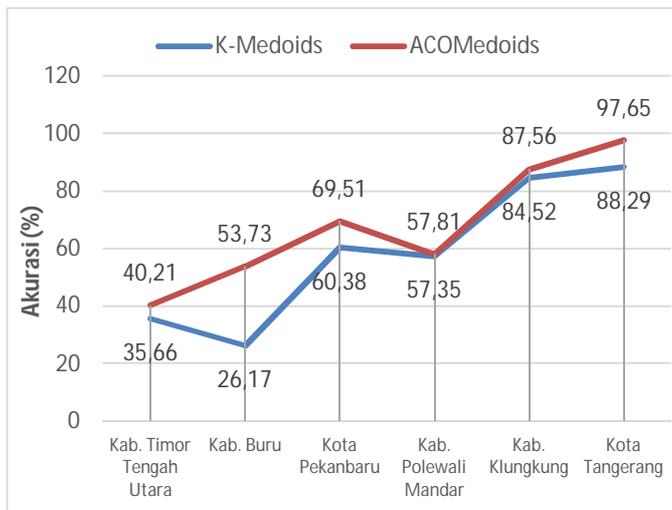
No	Nama PDAM	<i>K-Medoids</i>					<i>ACOMedoids</i>				
		<i>confusion matrix</i>				Akurasi (%)	<i>confusion matrix</i>				Akurasi (%)
		TP	TP	TP	TN		TP	TP	TP	TN	
1	PDAM Kab. Timor Tengah Utara	497	897	324	2191	35,66	728	844	377	1960	40,21
2	PDAM Kab. Buru	540	516	166	2813	26,17	1759	409	273	1594	53,73
3	PDAM Kota Pekanbaru	886	6565	2516	2373	60,38	113	8465	616	3146	69,51
4	PDAM Kab. Polewali Mandar	1161	7727	3038	3571	57,35	1121	7838	2927	3611	57,81
5	PDAM Kab. Klungkung	212	24995	3305	1309	84,52	155	25959	2344	1366	87,56
6	PDAM Kota Tangerang	71	31877	3764	474	88,29	16	35320	321	529	97,65



C. Perbandingan Metode *K-Medoids* dengan *ACOMedoids*

Perbandingan dilakukan terhadap metode *K-Medoids* dengan metode *ACOMedoids* terhadap akurasi dari hasil pengujian dengan metode *confusion matrix*.

Akurasi merupakan persentase ketepatan *record* data yang diklasifikasikan secara benar setelah dilakukan pengujian pada hasil klasifikasi. Perbandingan akurasi hasil pengujian antara metode *K-Medoids* dengan metode *ACOMedoids* dapat dilihat pada Gambar 7. Pada Gambar 7 menunjukkan akurasi metode *ACOMedoids* terlihat lebih tinggi dan lebih baik dibandingkan metode *K-Medoids*, hal ini menunjukkan titik pusat *cluster* yang didapat dari hasil algoritma ACO terbukti optimal dibandingkan dengan penentuan titik pusat *cluster* dengan menggunakan metode *K-Medoids*.



Gambar 7: Grafik perbandingan akurasi hasil pengujian

Nilai akurasi metode *ACOMedoids* selalu lebih tinggi dan lebih baik dibandingkan dengan metode *K-Medoids*, hal ini disebabkan oleh dalam penentuan titik *cluster* *K-Medoids* melakukan pemilihan secara acak dari data latih dan bergantung jumlah iterasi yang dilakukan semakin banyak iterasi memungkinkan mendapat titik *cluster* yang lebih baik tetapi itu akan membutuhkan waktu yang sangat lama. Sedangkan metode *ACOMedoids* dalam menentukan titik *cluster* menggunakan algoritma ACO yang mencari jarak terpendek dari setiap data latih, sehingga didapatkan fungsi minimal sebagai titik *cluster*. Metode *ACOMedoids* memerlukan waktu yang lebih lama pada saat proses algoritma ACO karena melibatkan semua data latih, tetapi tidak memerlukan jumlah iterasi yang banyak pada metode *clustering K-Medoids* berikutnya. Sehingga jika dalam efisiensi dan efektifitas dapat dikatanya metode *ACOMedoids* lebih baik dibandingkan *K-Medoids*. Nilai akurasi masing-masing PDAM terlihat bervariasi dan cenderung naik mengikuti jumlah data latih yang digunakan, semakin besar data latih, maka semakin besar tingkat akurasi yang

didapatkan. Tetapi hal ini tidak dapat menjadi acuan karena data tagihan rekening PDAM Kab. Klungkung dan PDAM Kota Tangerang memiliki data rekening yang lebih bersih dibandingkan PDAM yang lain, jadi tingkat akurasi juga dipengaruhi oleh beberapa faktor selain jumlah data yaitu data rekening yang bersih tidak memiliki lebih dari 3 tunggakan dan kebijakan masing-masing PDAM dalam hal pelanggan aktif dan penerbitan rekening tagihan pelanggan.

V. KESIMPULAN

Metode *K-Medoids* dan *ACOMedoids* dapat menghasilkan kelompok serta jumlah data pelanggan yang memiliki potensi untuk menunggak tagihan air (pelanggan tunggakan) maupun tidak menunggak (pelanggan lancar) pada periode tagihan berjalan. Algoritma ACO dapat menghasilkan titik pusat *cluster* yang optimal pada proses *clustering* metode *K-Medoids*. Metode *ACOMedoids* dapat menghasilkan tingkat akurasi yang tinggi dari perbandingan data hasil *clustering* dengan data tagihan aktual dari PDAM. Hal ini terlihat pada hasil akurasi yang selalu lebih baik dibandingkan dengan metode *K-Medoids* yaitu tertinggi mencapai akurasi 97,65% pada PDAM Tirta Benteng Kota Tangerang.

REFERENSI

- [1] BPPSPAM. *Buku Kinerja PDAM 2016*. Jakarta: BPPSPAM, 2016.
- [2] C. Laspidou, et al. "Exploring Patterns In Water Consumption By Clustering" *Elsevier: Procedia Engineering 13th Computer Control for Water Industry Conference, CCWI*, 2015, 119:1439-1446.
- [3] S. Moedjiono, F. Fransisca and A. Kusdaryono. "Segmentation and Classification Customer Payment Behavior at Multimedia Service Provider Company with K-Means and C4.5 Algorithm" *International Journal of Computer Networks and Communications Security*, Vol. 4, No. 9:265-275, 2015.
- [4] K. Dahiya, S. Bhatia. "Customer Churn Analysis in Telecom Industry" *IEEE 4th International Conference on Reliability, Infocom Technologies and Optimization (ICRITO) (Trends and Future Directions)*, 2-4 Sept. 2015.
- [5] A. J. Hamid and T. M. Ahmed. "Developing Prediction Model Of Loan Risk In Banks Using Data Mining" *Machine Learning and Applications: An International Journal (MLAIJ)*, Vol.3, No.1, 2016.
- [6] M. Sood and S. Bansal, "K-Medoids Clustering Technique using Bat Algorithm" *International Journal of Advanced Research in Computer Science and Software Engineering*, 2013, Vol.5, No.8:20-22.
- [7] R. Hadi, I K. G. D. Putra dan I N. S. Kumara, "Penentuan Kompetensi Mahasiswa Dengan Algoritma Genetik Dan Metode Fuzzy C-Means", *Majalah Ilmiah Teknologi Elektro*, Vol. 15, No. 2, Juli - Desember 2016.
- [8] P. Suwirmayanti, I K. G. D. Putra dan I N. S. Kumara, "Optimasi Pusat Cluster K-Prototype Dengan Algoritma Genetika", *Majalah Ilmiah Teknologi Elektro*, Vol. 13, No. 2, Juli - Desember 2014.
- [9] M. Dorigo, L. M. Gambardella. "Ant Colony System: A Cooperative Learning Approach to the Traveling Salesman Problem", *IEEE Transactions on Evolutionary Computation*, 1:53-56, 1997.
- [10] M. Dorigo, T. Stutzle, *Ant Colony Optimization*. London: A Bradford Book The MIT Press, 2004.
- [11] J. Han, M. Kamber, *Data mining: Concept and Techniques second edition*. USA: Elsevier Inc, 2006.
- [12] F. Gorunescu, *Data mining Concepts, Models and Techniques*. Romania: Springer-Verlag Berlin Heidelberg, 2011.