

Penerapan Algoritma K-Means Data Mining Untuk Clustering Hasil Test Urine Pasien

Ricky Syahbana Sitepu

Program Studi Teknik Informatika, STMIK Kaputama, Binjai, Sumatera Utara, Indonesia

Email: rickysitepu12@gmail.com

Email Penulis Korespondensi: rickysitepu12@gmail.com

Abstrak—Data Mining merupakan salah satu teknik pengolahan data dengan menggali berbagai informasi dari sekumpulan data yang tersimpan. Tes urin merupakan sarana untuk mengetahui ada atau tidaknya keterkaitan pasien terhadap penyalahgunaan narkotika. Badan Narkotika Nasional (BNN) Kota Binjai adalah sebuah lembaga pemerintah non kementerian (LPNK) Indonesia yang memiliki tugas di bidang pencegahan, pemberantasan penyalahgunaan dan peredaran gelap, Psikotropika, precursor, dan bahan adiktif lainnya kecuali bahan adiktif untuk tembakau dan alkohol. Penulisan laporan ini menggunakan metode clustering yang merupakan salah satu teknik data Mining untuk pengelompokan data Pasien Tes Urin Kantor BNN Kota Binjai. Dengan menggunakan metode clustering algoritma k-means. Dengan menerapkan 20 data alternatif pasien tes urin dan memberikan jumlah cluster sebanyak 3, dan memanfaatkan 3 kriteria utama sebagai penelitian di laporan ini, menghasilkan jumlah cluster 1 sebanyak 5 data, cluster 2 sebanyak 9 data. Dan cluster 3 sebanyak 6 data. Sistem pengelompokan data pasien tes urin ini dirancang dengan Bahasa pemrograman aplikasi MATLAB.

Kata Kunci: Algoritma K-Means; Clustering; Data Mining; Pasien Tes Urin; MATLAB

Abstract—Data Mining is a data processing technique by digging various information from a set of stored data. Urine test is a means to determine whether or not a patient is related to narcotic abuse. The National Narcotics Agency (BNN) of Binjai City is an Indonesian non-ministerial government agency (LPNK) that has duties in the fields of prevention, eradication of abuse and illicit trafficking, psychotropic substances, precursors, and other addictive substances except for tobacco and alcohol addicts. The writing of this report uses the clustering method which is one of the data mining techniques for grouping data on the Urine Test Patient at the Binjai City BNN Office. By using the k-means algorithm clustering method. By applying 20 alternative data for urine test patients and giving the number of clusters as many as 3, and utilizing the 3 main criteria as research in this report, the number of cluster 1 is 5 data, cluster 2 is 9 data. And cluster 3 of 6 data. This urine test patient data grouping system is designed with the MATLAB application programming language.

Keywords: K-Means Algorithm; Clustering; Data Mining; Patient Urine Test; MATLAB

1. PENDAHULUAN

Badan Narkotika Nasional (BNN) adalah sebuah lembaga Pemerintah Non Kementerian (LPNK) yang mempunyai tugas dan fungsi mengkoordinasikan instansi pemerintah terkait dalam penyusunan kebijakan dan pelaksanaan kebijakan operasional dalam pencegahan, pemberantasan, penyalahgunaan dan peredaran gelap narkotika (P4GN). BNN dipimpin oleh seorang Kepala BNN yang diangkat dan diberhentikan oleh Presiden. BNN berkedudukan dibawah dan bertanggung jawab kepada Presiden.

Peredaran dan penyalahgunaan Narkotika telah menjadi sebuah ancaman serius bagi masyarakat maupun pemerintah, oleh karena itu pemerintah membentuk sebuah badan khusus yang bertugas melakukan pencegahan dan penanggulangan bagi peredaran dan penyalahgunaan Narkotika, mulai dari tingkat nasional hingga kecamatan. Di seluruh wilayah Republik Indonesia, badan ini telah dibentuk dengan tujuan yang sama, yakni memerangi peredaran dan penyalahgunaan Narkotika.

Dalam Data Pasien *Test Urine* ini sering terjadi kesulitan dalam Penginputan/Pengentrian Data oleh Pihak BNN Kota Binjai dikarenakan banyaknya data yang kurang dimanfaatkan oleh BNN Kota Binjai dalam menghasilkan suatu informasi dikarenakan rentang usia yang bervariasi untuk setiap wilayah tempat tinggal pasien yang telah melakukan *Test urine*. Data mining dibutuhkan untuk menghasilkan informasi yang berguna terhadap pihak BNN Kota Binjai khususnya informasi tentang rentang usia pasien *Test urine* berdasarkan wilayah tempat tinggal pasien.

Dengan pemanfaatan data mining menggunakan data masukan berupa data pasien *Test urine* pada BNN Kota Binjai, data tersebut akan diolah dengan metode clustering menggunakan algoritma k-means dengan menggunakan variabel usia, variabel alamat dan variabel hasil *Test* pada data pasien *Test urine* pada BNN Kota Binjai. Berdasarkan hasil pengelompokan tersebut, pengolahan data yang dilakukan dapat menghasilkan informasi yang berguna terhadap pihak BNN Kota Binjai untuk memaksimalkan kinerja BNN Kota Binjai dalam hal Pemberantasan penyalahgunaan narkotika.

2. METODOLOGI PENELITIAN

2.1 Data Mining

Data mining merupakan sebuah proses untuk mendapatkan informasi yang berguna dari gudang basis data yang besar. *Data mining* juga dapat diartikan sebagai pengekstrakan informasi baru yang diambil dari bongkahan data

besar yang membantu dalam pengambilan keputusan, istilah *data mining* kadang disebut juga *knowledge discovery*. [1]

Menurut Turban, dkk dalam Kusri (2009, h.3) menyatakan bahwa *data mining* adalah suatu istilah yang digunakan untuk menguraikan penemuan pengetahuan dalam *database*. *Data Mining* adalah proses yang menggunakan teknik statistik, matematika, kecerdasan buatan, dan pembelajaran mesin (*machine learning*) untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan berpengetahuan yang terkait dari berbagai *database* besar. [2]

2.2 Metode Clustering

Menurut Fajar Astuti (2013, h.16) *clustering* adalah “suatu metode mempartisi data-set menjadi beberapa sub-set atau kelompok sedemikian rupa sehingga elemen-elemen dari suatu kelompok tertentu memiliki *set property* yang *dishare* bersama, dengan tingkat similaritas yang tinggi dalam suatu kelompok dan tingkat similaritas antar kelompok yang rendah”. Sehingga *cluster* tidak harus sama akan tetapi pengelompokannya berdasarkan pada kedekatan dari suatu karakteristik sampel yang ada salah satunya dengan menggunakan rumus jarak *eclidean*. Tujuan utama dari metode clustering adalah mengelompokkan sejumlah data/obyek kedalam cluster (group) sehingga dalam setiap cluster akan berisi data yang semirip mungkin. [3] *Clustering* merupakan proses partisi satu set objek data ke dalam himpunan bagian yang disebut dengan *cluster*. Objek yang di dalam cluster memiliki kemiripan karakteristik antar satu sama lainnya dan berbeda dengan *cluster* yang lain. Partisi tidak dilakukan secara manual melainkan dengan suatu algoritma *clustering*. Oleh karena itu *clustering* sangat berguna dan bisa menemukan *group* atau kelompok yang tidak dikenal dalam data. *Clustering* banyak digunakan dalam berbagai aplikasi seperti misalnya pada *business intelligence*, pengenalan pola citra, *web search*, bidang ilmu biologi, dan untuk keamanan (*security*).

2.3 Algoritma K-Means

Algoritma *K-Means* merupakan algoritma yang relatif sederhana untuk mengklasifikasi atau mengelompokkan sejumlah besar obyek dengan atribut tertentu ke dalam kelompok (*cluster*) sebanyak *k* pada algoritma *K-Means*, jumlah *cluster* sudah ditentukan lebih dahulu.

Menurut Eko Prasetyo (2012, h.178) menyatakan bahwa “*K-Means* merupakan salah satu metode pengelompokan data non hierarki (sekatan) yang berusaha mempartisi data ke dalam *cluster* atau kelompok sehingga data yang memiliki karakteristik yang sama akan dikelompokkan ke dalam satu *cluster* yang sama dan data yang memiliki karakteristik yang berbeda dikelompokkan ke dalam kelompok yang lain”. [1] Adapun tujuan pengelompokkan data ini adalah untuk meminimalkan fungsi objektif yang diset dalam proses pengelompokkan, yang pada umumnya berusaha meminimalkan variasi di dalam suatu kelompok dan memaksimalkan variasi antar kelompok. Data *clustering* menggunakan metode *K-Means* ini secara umum dilakukan dengan algoritma dasar sebagai berikut :

1. Tentukan jumlah *cluster*
2. Alokasikan data ke dalam *cluster* secara random
3. Hitung *centroid* atau rata-rata dari data yang ada di masing-masing *cluster*.
4. Alokasikan masing-masing data ke *centroid* atau rata-rata terdekat
5. Kembali ke step 3, apabila masih ada data yang berpindah *cluster* atau apabila perubahan nilai *centroid*, ada yang diatas nilai *threshold* yang ditentukan atau apabila terjadi perubahan nilai pada *objective function* yang digunakan diatas nilai *threshold* yang ditentukan.

2.4 Pengertian Matlab

Menurut Etter and Kuncicky (Budi Santosa ,2007, h. 1) menyatakan bahwa *Matlab* adalah suatu bahasa pemrograman tingkat tinggi yang diperuntukkan untuk komputasi teknis. *Matlab* mengintegrasikan aspek komputasi, visualisasi dan dipergunakan untuk aplikasi. *Matlab* juga berisi *toolbox* yang berisi fungsi-fungsi tambahan untuk aplikasi khusus. [4]

Penggunaan *Matlab* meliputi bidang-bidang :

1. Pemodelan, Simulasi dan pembuatan *prototype*
2. Komputasi dan matematika
3. Data analisis, eksplorasi, visualisasi
4. Pembuatan grafik *scientific* dan *engineering*
5. Pembuatan *graphical user interface (GUI)* yang memudahkan pemakaian bagi kalangan yang awam dengan komputasi.

Matlab merupakan perangkat lunak produk dari *The Math Works, Inc* yang memadukan kemampuan perhitungan, pencitraan, dan pemrograman dalam satu paket. Sebenarnya *Matlab* tidaklah berbeda dengan kalkulator *scientific* yang sehari-hari dikenal oleh orang-orang di bidang teknik. Secara garis besar lingkungan kerja *Matlab* terdiri atas beberapa unsur, yaitu :

1. *Command window* (Jendela Perintah), semua perintah *Matlab* dituliskan dan dieksekusi. Kita dapat menuliskan perintah perhitungan sederhana, memanggil fungsi, mencari informasi tentang sebuah fungsi dengan aturan penulisannya (*help*), demo program, dan sebagainya. Setiap penulisan perintah selalu diawali dengan prompt '>>'. Misal, mencari nilai sin 750, maka pada *command window* kita dapat mengetikkan :

- ```
>> sin(75)
ans = -0.38778
```
- Workspace* (Jendela Ruang Kerja), Jendela ini berisi informasi pemakaian variabel didalam memori matlab. Misalkan kita akan menjumlahkan dua buah bilangan, maka pada command window kita dapat mengetikkan :

```
>> bilangan1 = 10
Bilangan1 =10
>>bilangan2 = 5
Bilangan1=10
>>hasil= bilangan1 + bilangan2
Hasil=15
```
  - Command history* (Jendela *history*), Jendela ini berisi informasi tentang perintah yang pernah dituliskan sebelumnya. Kita dapat mengambil kembali perintah dengan menekan tombol panah keatas atau mengklik perintah pada jendela histori, kemudian melakukan *copy paste* ke *command window*.

### 3. HASIL DAN PEMBAHASAN

#### 3.1 Analisa

Langkah-langkah yang dilakukan untuk perhitungan data pasien *test urine* menggunakan metode *clustering* dengan algoritma *K-means* ini, agar dapat dihasilkan sebuah pengetahuan baru, mengenai berapa banyak kelompok data Usia pasien dan data Alamat pasien *test urine* berdasarkan Hasil *Test Urine*. Sehingga dapat diketahui hubungan terdekat antara kelompok data pasien *test urine*.

##### 3.1.1 Penerapan Algoritma K-Means

Untuk menentukan grup dari suatu objek, pertama yang harus dilakukan adalah mengukur jarak *dEuclidean* antara dua titik atau objek X dan Y yang didefinisikan sebagai berikut :

$$dEuclidean(X,Y) = \sqrt{\sum_i (X1 - Y2)^2} \quad (1)$$

Dengan rumus diatas maka dapat dilakukan perhitungan agar dapat menentukan jarak pada Pengelompokan Data Pasien *Test Urine*. Setelah data di-*import* kedalam *Matlab* dengan *syntax* yang sudah ditentukan berdasarkan jarak *centroid* terkecil berada. Hasil dari perhitungan group di bagi menjadi 3 *group* dengan data Usia (X), Alamat (Y), dan Hasil *Test* (Z), yaitu dapat dilihat pada tabel berikut :

Tabel 1. Hasil Penentuan *Group*

| No | X | Y | Z | Group | No | X | Y | Z | Group |
|----|---|---|---|-------|----|---|---|---|-------|
| 1  | 1 | 1 | 3 | 2     | 27 | 2 | 3 | 6 | 1     |
| 2  | 4 | 2 | 2 | 2     | 28 | 2 | 2 | 6 | 1     |
| 3  | 1 | 3 | 4 | 2     | 29 | 3 | 2 | 7 | 1     |
| 4  | 2 | 4 | 1 | 2     | 30 | 2 | 4 | 7 | 1     |
| 5  | 3 | 1 | 3 | 2     | 31 | 3 | 1 | 6 | 1     |
| 6  | 5 | 1 | 4 | 2     | 32 | 3 | 1 | 6 | 1     |
| 7  | 4 | 2 | 2 | 2     | 33 | 2 | 5 | 6 | 1     |
| 8  | 2 | 4 | 1 | 2     | 34 | 3 | 5 | 1 | 2     |
| 9  | 3 | 5 | 5 | 1     | 35 | 2 | 1 | 1 | 2     |
| 10 | 1 | 3 | 3 | 2     | 36 | 2 | 3 | 5 | 1     |
| 11 | 1 | 1 | 4 | 2     | 37 | 2 | 1 | 4 | 2     |
| 12 | 2 | 5 | 5 | 1     | 38 | 3 | 4 | 4 | 2     |
| 13 | 5 | 4 | 3 | 2     | 39 | 1 | 4 | 3 | 2     |
| 14 | 2 | 4 | 2 | 2     | 40 | 3 | 1 | 6 | 1     |
| 15 | 2 | 1 | 4 | 2     | 41 | 2 | 3 | 1 | 2     |
| 16 | 5 | 1 | 4 | 2     | 42 | 3 | 3 | 6 | 1     |
| 17 | 4 | 2 | 1 | 2     | 43 | 1 | 3 | 1 | 2     |
| 18 | 3 | 5 | 5 | 1     | 44 | 2 | 3 | 3 | 2     |
| 19 | 1 | 4 | 2 | 2     | 45 | 2 | 5 | 6 | 1     |
| 20 | 5 | 4 | 2 | 2     | 46 | 2 | 4 | 1 | 2     |
| 21 | 1 | 4 | 7 | 1     | 47 | 3 | 1 | 1 | 2     |
| 22 | 3 | 1 | 2 | 2     | 48 | 1 | 3 | 1 | 2     |
| 23 | 3 | 1 | 7 | 1     | 49 | 5 | 1 | 3 | 2     |
| 24 | 2 | 4 | 7 | 1     | 50 | 1 | 5 | 6 | 1     |
| 25 | 4 | 3 | 7 | 1     | 51 | 1 | 5 | 1 | 2     |
| 26 | 2 | 5 | 7 | 1     | 52 | 1 | 5 | 3 | 2     |

| No  | X | Y | Z | Group |
|-----|---|---|---|-------|
| 53  | 1 | 5 | 6 | 1     |
| 54  | 1 | 4 | 6 | 1     |
| 55  | 1 | 1 | 2 | 2     |
| 56  | 2 | 3 | 7 | 1     |
| 57  | 2 | 5 | 6 | 1     |
| 58  | 3 | 5 | 7 | 1     |
| 59  | 2 | 3 | 6 | 1     |
| 60  | 3 | 3 | 6 | 1     |
| 61  | 2 | 4 | 7 | 1     |
| 62  | 4 | 3 | 6 | 1     |
| 63  | 2 | 3 | 2 | 2     |
| 64  | 3 | 4 | 7 | 1     |
| 65  | 2 | 3 | 1 | 2     |
| 66  | 2 | 5 | 7 | 1     |
| 67  | 2 | 4 | 7 | 1     |
| 68  | 3 | 3 | 2 | 2     |
| 69  | 2 | 1 | 7 | 1     |
| 70  | 3 | 5 | 1 | 2     |
| 71  | 2 | 5 | 2 | 2     |
| 72  | 3 | 5 | 7 | 1     |
| 73  | 3 | 4 | 6 | 1     |
| 74  | 3 | 1 | 2 | 2     |
| 75  | 2 | 3 | 6 | 1     |
| 76  | 6 | 5 | 7 | 1     |
| 77  | 4 | 2 | 7 | 1     |
| 78  | 2 | 5 | 2 | 2     |
| 79  | 3 | 4 | 7 | 1     |
| 80  | 2 | 4 | 7 | 1     |
| 81  | 3 | 1 | 7 | 1     |
| 82  | 2 | 3 | 7 | 1     |
| 83  | 3 | 4 | 3 | 2     |
| 84  | 2 | 5 | 6 | 1     |
| 85  | 3 | 5 | 3 | 2     |
| 86  | 2 | 3 | 3 | 2     |
| 87  | 3 | 3 | 3 | 2     |
| 88  | 1 | 3 | 7 | 1     |
| 89  | 3 | 2 | 5 | 1     |
| 90  | 2 | 2 | 4 | 2     |
| 91  | 2 | 2 | 5 | 1     |
| 92  | 2 | 3 | 7 | 1     |
| 93  | 2 | 4 | 4 | 2     |
| 94  | 2 | 5 | 6 | 1     |
| 95  | 2 | 4 | 2 | 2     |
| 96  | 2 | 1 | 6 | 1     |
| 97  | 2 | 3 | 5 | 1     |
| 98  | 2 | 1 | 6 | 1     |
| 99  | 2 | 5 | 7 | 1     |
| 100 | 2 | 5 | 6 | 1     |
| 101 | 2 | 5 | 5 | 1     |
| 102 | 2 | 5 | 6 | 1     |
| 103 | 3 | 5 | 6 | 1     |
| 104 | 2 | 4 | 7 | 1     |
| 105 | 2 | 3 | 4 | 2     |
| 106 | 2 | 5 | 6 | 1     |
| 107 | 2 | 4 | 5 | 1     |
| 108 | 2 | 1 | 6 | 1     |
| 109 | 2 | 3 | 6 | 1     |
| 110 | 3 | 1 | 5 | 1     |
| 111 | 2 | 5 | 2 | 2     |
| 112 | 2 | 5 | 7 | 1     |

| No  | X | Y | Z | Group |
|-----|---|---|---|-------|
| 113 | 2 | 5 | 5 | 1     |
| 114 | 2 | 5 | 6 | 1     |
| 115 | 2 | 4 | 3 | 2     |
| 116 | 2 | 1 | 5 | 1     |
| 117 | 2 | 3 | 7 | 1     |
| 118 | 2 | 5 | 6 | 1     |
| 119 | 2 | 5 | 5 | 1     |
| 120 | 2 | 3 | 4 | 2     |
| 121 | 2 | 3 | 5 | 1     |
| 122 | 2 | 4 | 7 | 1     |
| 123 | 2 | 5 | 3 | 2     |
| 124 | 2 | 5 | 6 | 1     |
| 125 | 2 | 4 | 5 | 1     |
| 126 | 2 | 1 | 6 | 1     |
| 127 | 4 | 3 | 7 | 1     |
| 128 | 2 | 1 | 4 | 2     |
| 129 | 2 | 2 | 7 | 1     |
| 130 | 2 | 5 | 6 | 1     |
| 131 | 2 | 5 | 3 | 2     |
| 132 | 2 | 2 | 6 | 1     |
| 133 | 2 | 4 | 1 | 2     |
| 134 | 2 | 1 | 1 | 2     |
| 135 | 2 | 3 | 2 | 2     |
| 136 | 3 | 5 | 6 | 1     |
| 137 | 2 | 3 | 7 | 1     |
| 138 | 2 | 5 | 6 | 1     |
| 139 | 2 | 5 | 7 | 1     |
| 140 | 2 | 1 | 7 | 1     |
| 141 | 2 | 5 | 6 | 1     |
| 142 | 2 | 3 | 1 | 2     |
| 143 | 5 | 3 | 7 | 1     |
| 144 | 3 | 3 | 3 | 2     |
| 145 | 2 | 5 | 1 | 2     |
| 146 | 2 | 4 | 4 | 2     |
| 147 | 2 | 1 | 1 | 2     |
| 148 | 3 | 3 | 7 | 1     |
| 149 | 2 | 1 | 6 | 1     |
| 150 | 2 | 5 | 2 | 2     |
| 151 | 2 | 5 | 7 | 1     |
| 152 | 2 | 2 | 3 | 2     |
| 153 | 2 | 2 | 7 | 1     |
| 154 | 3 | 4 | 4 | 2     |
| 155 | 3 | 1 | 1 | 2     |
| 156 | 1 | 3 | 4 | 2     |
| 157 | 2 | 5 | 3 | 2     |
| 158 | 2 | 5 | 7 | 1     |
| 159 | 2 | 3 | 1 | 2     |
| 160 | 2 | 3 | 7 | 1     |
| 161 | 2 | 4 | 1 | 2     |
| 162 | 2 | 3 | 4 | 2     |
| 163 | 2 | 3 | 7 | 1     |
| 164 | 2 | 4 | 1 | 2     |
| 165 | 3 | 3 | 4 | 2     |
| 166 | 5 | 2 | 1 | 2     |
| 167 | 2 | 4 | 7 | 1     |
| 168 | 2 | 5 | 3 | 2     |
| 169 | 2 | 4 | 4 | 2     |
| 170 | 2 | 1 | 7 | 1     |
| 171 | 2 | 3 | 1 | 2     |
| 172 | 3 | 1 | 7 | 1     |

| No  | X | Y | Z | Group | No  | X | Y | Z | Group |
|-----|---|---|---|-------|-----|---|---|---|-------|
| 173 | 2 | 5 | 1 | 2     | 187 | 1 | 2 | 7 | 1     |
| 174 | 3 | 5 | 4 | 2     | 188 | 2 | 3 | 5 | 1     |
| 175 | 3 | 1 | 7 | 1     | 189 | 3 | 4 | 4 | 2     |
| 176 | 2 | 3 | 7 | 1     | 190 | 3 | 2 | 7 | 1     |
| 177 | 2 | 1 | 1 | 2     | 191 | 3 | 3 | 5 | 1     |
| 178 | 2 | 5 | 3 | 2     | 192 | 5 | 4 | 7 | 1     |
| 179 | 2 | 5 | 7 | 1     | 193 | 6 | 3 | 2 | 2     |
| 180 | 2 | 2 | 4 | 2     | 194 | 2 | 5 | 4 | 2     |
| 181 | 2 | 2 | 5 | 1     | 195 | 2 | 4 | 7 | 1     |
| 182 | 3 | 4 | 7 | 1     | 196 | 3 | 3 | 5 | 1     |
| 183 | 1 | 1 | 1 | 2     | 197 | 6 | 1 | 5 | 1     |
| 184 | 2 | 3 | 7 | 1     | 198 | 2 | 5 | 3 | 2     |
| 185 | 1 | 5 | 3 | 2     | 199 | 4 | 5 | 4 | 2     |
| 186 | 5 | 5 | 4 | 2     | 200 | 4 | 5 | 2 | 2     |

Dari hasil tabel diatas dapat dijumlahkan masing-masing *group* yaitu sebagai berikut :

1. *Group* 1 sebanyak 109 data
2. *Group* 2 sebanyak 91 data

Jadi total data = 200

### 3.1.2 Perhitungan Jarak Objek Ke Centroid

Dari hasil analisa diatas proses *Replicate* ditentukan sebanyak 5 kali perulangan dimana *cluster* ditentukan sebanyak 2 (X, Y, dan Z) maka total iterasi sebanyak 10 kali, hal ini menunjukkan bahwa proses iterasi berhenti jika total dan jarak dengan iterasi sebelumnya sampai pada *group* yang tidak berubah lagi. Hasil iterasi yang diperoleh dari perhitungan jarak objek ke *centroid* menggunakan perogran *Matlab* adalah sebagai berikut :

*Replicate 1, 3 iterations, total sum of distances = 798.967.*

*Replicate 2, 5 iterations, total sum of distances = 798.967.*

*Replicate 3, 5 iterations, total sum of distances = 798.967.*

*Replicate 4, 3 iterations, total sum of distances = 798.967.*

*Replicate 5, 5 iterations, total sum of distances = 798.967.*

*Best total sum of distances = 798.967*

*Centroid 1 = total group 1/banyak group 1*

$C1 = 2385 / 109 = 2,18$

$C2 = 3376 / 109 = 3,09$

$C3 = 6293 / 109 = 5,77$

*Centroid 2 = total group 2/banyak group 2*

$C1 = 2417 / 91 = 2,65$

$C2 = 3230 / 91 = 3,54$

$C2 = 2472 / 91 = 2,71$

Dari hasil perhitungan *centroid* diatas dapat dilihat pada tabel berikut :

**Tabel 2.** Perhitungan *Centroid*

| Jenis Barang<br>(X) | Wilayah<br>(Y) | Omset<br>(Z) | Keterangan        |
|---------------------|----------------|--------------|-------------------|
| 2,18                | 3,09           | 5,77         | <i>Centroid 1</i> |
| 2,65                | 3,54           | 2,71         | <i>Centroid 2</i> |

### 3.1.1 Hasil Akhir Penentuan Grup

Hasil akhir penentuan kelompok data pasien test urine pada mana yang termasuk dalam *group1*, dan *group2*, dapat dilihat dalam 2 bagian penjelasan dibawah ini :

- Untuk *group 1* memiliki nilai yang baik. Karena untuk hasil Usia (X), Alamat (Y), Hasil test (Z) jumlah data yang dimiliki sebanyak 109 data. Dari keterangan data menunjukkan bahwa usia pada data pasien test urine dengan range 2,18 kemudian alamat 3,09 dan hasil test ,5,77.

**Tabel 3.** Hasil *Cluster 1*

| NO | X | Y | Z | Group | NO | X | Y | Z | Group |
|----|---|---|---|-------|----|---|---|---|-------|
| 1  | 1 | 1 | 3 | 1     | 6  | 3 | 1 | 7 | 1     |
| 2  | 3 | 5 | 5 | 1     | 7  | 2 | 4 | 7 | 1     |
| 3  | 2 | 5 | 5 | 1     | 8  | 4 | 3 | 7 | 1     |
| 4  | 3 | 5 | 5 | 1     | 9  | 2 | 5 | 7 | 1     |
| 5  | 1 | 4 | 7 | 1     | 10 | 2 | 3 | 6 | 1     |

| NO | X | Y | Z | Group | NO  | X | Y | Z | Group |
|----|---|---|---|-------|-----|---|---|---|-------|
| 11 | 2 | 2 | 6 | 1     | 63  | 3 | 1 | 5 | 1     |
| 12 | 3 | 2 | 7 | 1     | 64  | 2 | 5 | 7 | 1     |
| 13 | 2 | 4 | 7 | 1     | 65  | 2 | 5 | 5 | 1     |
| 14 | 3 | 1 | 6 | 1     | 66  | 2 | 5 | 6 | 1     |
| 15 | 3 | 1 | 6 | 1     | 67  | 2 | 1 | 5 | 1     |
| 16 | 2 | 5 | 6 | 1     | 68  | 2 | 3 | 7 | 1     |
| 17 | 2 | 3 | 5 | 1     | 69  | 2 | 5 | 6 | 1     |
| 18 | 3 | 1 | 6 | 1     | 70  | 2 | 5 | 5 | 1     |
| 19 | 3 | 3 | 6 | 1     | 71  | 2 | 3 | 5 | 1     |
| 20 | 2 | 5 | 6 | 1     | 72  | 2 | 4 | 7 | 1     |
| 21 | 1 | 5 | 6 | 1     | 73  | 2 | 5 | 6 | 1     |
| 22 | 1 | 5 | 6 | 1     | 74  | 2 | 4 | 5 | 1     |
| 23 | 1 | 4 | 6 | 1     | 75  | 2 | 1 | 6 | 1     |
| 24 | 2 | 3 | 7 | 1     | 76  | 4 | 3 | 7 | 1     |
| 25 | 2 | 5 | 6 | 1     | 77  | 2 | 2 | 7 | 1     |
| 26 | 3 | 5 | 7 | 1     | 78  | 2 | 5 | 6 | 1     |
| 27 | 2 | 3 | 6 | 1     | 79  | 2 | 2 | 6 | 1     |
| 28 | 3 | 3 | 6 | 1     | 80  | 3 | 5 | 6 | 1     |
| 29 | 2 | 4 | 7 | 1     | 81  | 2 | 3 | 7 | 1     |
| 30 | 4 | 3 | 6 | 1     | 82  | 2 | 5 | 6 | 1     |
| 31 | 3 | 4 | 7 | 1     | 83  | 2 | 5 | 7 | 1     |
| 32 | 2 | 5 | 7 | 1     | 84  | 2 | 1 | 7 | 1     |
| 33 | 2 | 4 | 7 | 1     | 85  | 2 | 5 | 6 | 1     |
| 34 | 2 | 1 | 7 | 1     | 86  | 5 | 3 | 7 | 1     |
| 35 | 3 | 5 | 7 | 1     | 87  | 3 | 3 | 7 | 1     |
| 36 | 3 | 4 | 6 | 1     | 88  | 2 | 1 | 6 | 1     |
| 37 | 2 | 3 | 6 | 1     | 89  | 2 | 5 | 7 | 1     |
| 38 | 6 | 5 | 7 | 1     | 90  | 2 | 2 | 7 | 1     |
| 39 | 4 | 2 | 7 | 1     | 91  | 2 | 5 | 7 | 1     |
| 40 | 3 | 4 | 7 | 1     | 92  | 2 | 3 | 7 | 1     |
| 41 | 2 | 4 | 7 | 1     | 93  | 2 | 3 | 7 | 1     |
| 42 | 3 | 1 | 7 | 1     | 94  | 2 | 4 | 7 | 1     |
| 43 | 2 | 3 | 7 | 1     | 95  | 2 | 1 | 7 | 1     |
| 44 | 2 | 5 | 6 | 1     | 96  | 3 | 1 | 7 | 1     |
| 45 | 1 | 3 | 7 | 1     | 97  | 3 | 1 | 7 | 1     |
| 46 | 3 | 2 | 5 | 1     | 98  | 2 | 3 | 7 | 1     |
| 47 | 2 | 2 | 5 | 1     | 99  | 2 | 5 | 7 | 1     |
| 48 | 2 | 3 | 7 | 1     | 100 | 2 | 2 | 5 | 1     |
| 49 | 2 | 5 | 6 | 1     | 101 | 3 | 4 | 7 | 1     |
| 50 | 2 | 1 | 6 | 1     | 102 | 2 | 3 | 7 | 1     |
| 51 | 2 | 3 | 5 | 1     | 103 | 1 | 2 | 7 | 1     |
| 52 | 2 | 1 | 6 | 1     | 104 | 2 | 3 | 5 | 1     |
| 53 | 2 | 5 | 7 | 1     | 105 | 3 | 2 | 7 | 1     |
| 54 | 2 | 5 | 6 | 1     | 106 | 3 | 3 | 5 | 1     |
| 55 | 2 | 5 | 5 | 1     | 107 | 5 | 4 | 7 | 1     |
| 56 | 2 | 5 | 6 | 1     | 108 | 2 | 4 | 7 | 1     |
| 57 | 3 | 5 | 6 | 1     | 109 | 3 | 3 | 5 | 1     |
| 58 | 2 | 4 | 7 | 1     | 110 | 6 | 1 | 5 | 1     |
| 59 | 2 | 5 | 6 | 1     |     |   |   |   |       |
| 60 | 2 | 4 | 5 | 1     |     |   |   |   |       |
| 61 | 2 | 1 | 6 | 1     |     |   |   |   |       |
| 62 | 2 | 3 | 6 | 1     |     |   |   |   |       |

- b. Untuk *group* 2 memiliki nilai yang cukup baik. Karena untuk hasil Usia (X), Alamat (Y), Hasil test (Z) jumlah data yang dimiliki sebanyak 109 data. Dari keterangan data menunjukkan bahwa usia pada data pasien test urine dengan *range* 2,65 kemudian alamat 3,54 dan hasil test ,2,71 .

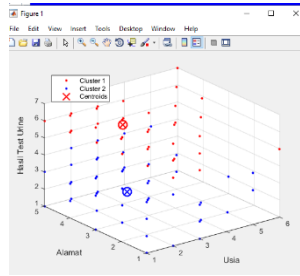
Tabel 3. Hasil Cluster 2

| NO | X | Y | Z | Group | NO | X | Y | Z | Group |
|----|---|---|---|-------|----|---|---|---|-------|
| 1  | 1 | 1 | 3 | 2     | 3  | 1 | 3 | 4 | 2     |
| 2  | 4 | 2 | 2 | 2     | 4  | 2 | 4 | 1 | 2     |

| NO | X | Y | Z | Group |
|----|---|---|---|-------|
| 5  | 3 | 1 | 3 | 2     |
| 6  | 5 | 1 | 4 | 2     |
| 7  | 4 | 2 | 2 | 2     |
| 8  | 2 | 4 | 1 | 2     |
| 9  | 1 | 3 | 3 | 2     |
| 10 | 1 | 1 | 4 | 2     |
| 11 | 5 | 4 | 3 | 2     |
| 12 | 2 | 4 | 2 | 2     |
| 13 | 2 | 1 | 4 | 2     |
| 14 | 5 | 1 | 4 | 2     |
| 15 | 4 | 2 | 1 | 2     |
| 16 | 1 | 4 | 2 | 2     |
| 17 | 5 | 4 | 2 | 2     |
| 18 | 3 | 1 | 2 | 2     |
| 19 | 3 | 5 | 1 | 2     |
| 20 | 2 | 1 | 1 | 2     |
| 21 | 2 | 1 | 4 | 2     |
| 22 | 3 | 4 | 4 | 2     |
| 23 | 1 | 4 | 3 | 2     |
| 24 | 2 | 3 | 1 | 2     |
| 25 | 1 | 3 | 1 | 2     |
| 26 | 2 | 3 | 3 | 2     |
| 27 | 2 | 4 | 1 | 2     |
| 28 | 3 | 1 | 1 | 2     |
| 29 | 1 | 3 | 1 | 2     |
| 30 | 5 | 1 | 3 | 2     |
| 31 | 1 | 5 | 1 | 2     |
| 32 | 1 | 5 | 3 | 2     |
| 33 | 1 | 1 | 2 | 2     |
| 34 | 2 | 3 | 2 | 2     |
| 35 | 2 | 3 | 1 | 2     |
| 36 | 3 | 3 | 2 | 2     |
| 37 | 3 | 5 | 1 | 2     |
| 38 | 2 | 5 | 2 | 2     |
| 39 | 3 | 1 | 2 | 2     |
| 40 | 2 | 5 | 2 | 2     |
| 41 | 3 | 4 | 3 | 2     |
| 42 | 3 | 5 | 3 | 2     |
| 43 | 2 | 3 | 3 | 2     |
| 44 | 3 | 3 | 3 | 2     |
| 45 | 2 | 2 | 4 | 2     |
| 46 | 2 | 4 | 4 | 2     |
| 47 | 2 | 4 | 2 | 2     |
| 48 | 2 | 3 | 4 | 2     |
| 49 | 2 | 5 | 2 | 2     |
| 50 | 2 | 4 | 3 | 2     |
| 51 | 2 | 3 | 4 | 2     |
| 52 | 2 | 5 | 3 | 2     |
| 53 | 2 | 1 | 4 | 2     |
| 54 | 2 | 5 | 3 | 2     |
| 55 | 2 | 4 | 1 | 2     |
| 56 | 2 | 1 | 1 | 2     |
| 57 | 2 | 3 | 2 | 2     |
| 58 | 2 | 3 | 1 | 2     |
| 59 | 3 | 3 | 3 | 2     |
| 60 | 2 | 5 | 1 | 2     |
| 61 | 2 | 4 | 4 | 2     |
| 62 | 2 | 1 | 1 | 2     |
| 63 | 2 | 5 | 2 | 2     |
| 64 | 2 | 2 | 3 | 2     |

| NO | X | Y | Z | Group |
|----|---|---|---|-------|
| 65 | 3 | 4 | 4 | 2     |
| 66 | 3 | 1 | 1 | 2     |
| 67 | 1 | 3 | 4 | 2     |
| 68 | 2 | 5 | 3 | 2     |
| 69 | 2 | 3 | 1 | 2     |
| 70 | 2 | 4 | 1 | 2     |
| 71 | 2 | 3 | 4 | 2     |
| 72 | 2 | 4 | 1 | 2     |
| 73 | 3 | 3 | 4 | 2     |
| 74 | 5 | 2 | 1 | 2     |
| 75 | 2 | 5 | 3 | 2     |
| 76 | 2 | 4 | 4 | 2     |
| 77 | 2 | 3 | 1 | 2     |
| 78 | 2 | 5 | 1 | 2     |
| 79 | 3 | 5 | 4 | 2     |
| 80 | 2 | 1 | 1 | 2     |
| 81 | 2 | 5 | 3 | 2     |
| 82 | 2 | 2 | 4 | 2     |
| 83 | 1 | 1 | 1 | 2     |
| 84 | 1 | 5 | 3 | 2     |
| 85 | 5 | 5 | 4 | 2     |
| 86 | 3 | 4 | 4 | 2     |
| 87 | 6 | 3 | 2 | 2     |
| 88 | 2 | 5 | 4 | 2     |
| 89 | 2 | 5 | 3 | 2     |
| 90 | 4 | 5 | 4 | 2     |
| 91 | 4 | 5 | 2 | 2     |

Untuk lebih jelasnya hasil *cluster* 2 dapat dilihat pada grafik dibawah ini :



**Gambar 1.** Grafik Hasil Perhitungan *Cluster* 2 Berdasarkan Data Pasien Test Urine

Pusatnya =        2 (2,18)    3 (3,09)    6 (5,77)  
                       2 (2,65)    3 (3,54)    2(2,71)

**Keterangan :**

Dari 200 data pasien test urine diperoleh 2 *cluster*, dimana *cluster* 1 berjumlah 109 Pasien, dan *cluster* 2 berjumlah 91 Pasien. Berikut adalah keterangan dari pusat *cluster* pada grafik :

2 (2,18)    3 (3,09)    6 (5,77)

Dapat diketahui bahwasanya pada *cluster* 1 Kelompok dengan Rentang Usia 21-30 Tahun dengan Alamat/Wilayah Binjai Barat dan Hasil Test Urine adalah Extacy/Inex

2 (2,65)    3 (3,54)    2(2,71)

Dapat diketahui bahwasanya pada *cluster* 2 Kelompok dengan Rentang Usia 21-30 Tahun dengan Alamat/Wilayah Binjai Barat dan Hasil Test Urine adalah Mop(Morfin)

```

Command Window
Replicate 1, 3 iterations, total sum of distances = 798.967.
Replicate 2, 5 iterations, total sum of distances = 798.967.
Replicate 3, 5 iterations, total sum of distances = 798.967.
Replicate 4, 3 iterations, total sum of distances = 798.967.
Replicate 5, 5 iterations, total sum of distances = 798.967.
Best total sum of distances = 798.967

centroid =

 2.3853 3.3761 6.2936
 2.4176 3.2308 2.4725

MatriksDenganIndeks =

 1.0000 1.0000 3.0000 18.4128 7.2641 2.0000
 4.0000 2.0000 2.0000 22.9358 4.2421 2.0000
 1.0000 3.0000 4.0000 7.3211 4.3960 2.0000
 2.0000 4.0000 1.0000 28.5596 2.9344 2.0000
 3.0000 1.0000 3.0000 16.8716 5.5938 2.0000
 5.0000 1.0000 4.0000 17.7431 13.9784 2.0000
 4.0000 2.0000 2.0000 22.9358 4.2421 2.0000
 2.0000 4.0000 1.0000 28.5596 2.9344 2.0000
 3.0000 5.0000 5.0000 4.6881 9.8575 1.0000
 1.0000 3.0000 3.0000 12.9083 2.3410 2.0000
 1.0000 1.0000 4.0000 12.8257 9.3190 2.0000
 2.0000 5.0000 5.0000 4.4587 9.6927 1.0000
 5.0000 4.0000 3.0000 18.0734 7.5388 2.0000

```

**Gambar 2.** Jendela *Matlab* Menampilkan Iterasi Yang Terjadi Dan Hasil Pembahasan Perhitungan *Group 3 Cluster Equilidien*

**4. KESIMPULAN**

Setelah melakukan peneitian ini, dapat disimpulkan bahwa Pengelompokan Data Pasien *Test Urine* Menggunakan Metode *Clustering* Pada Kantor Badan Narkotika Nasional Kota Binjai merupakan sebuah perangkat lunak yang efektif untuk mengetahui informasi tentang wilayah mana yang paling banyak menggunakan narkoba dengan rentang usia yang yang paling banyak menggunakan narkoba berdasarkan hasil *test urinenya* dibandingkan dengan sistem yang sedang berjalan saat ini yang masih dilakukan secara manual. Program ini sangat membantu Pihak Badan Narkotika Nasional Kota Binjai, karena mempermudah kinerja Badan Narkotika Nasional Kota Binjai dalam memaksimalkan kinerjanya khususnya dalam hal pemberantasan penyalahgunaan narkoba.

**UCAPAN TERIMAKASIH**

**REFERENCES**

[1] Eko Prastyo, *Data Mining Konsep Dan Aplikasi Menggunakan Matlab*. Yogyakarta: CV Andi, 2012.  
 [2] E. T. L. Kusriani, *Algoritma Data Mining*. Yogyakarta: CV Andi, 2009. [Online]. Available: [https://www.google.co.id/books/edition/Algoritma\\_Data\\_Mining/-Ojclag73O8C?hl=id&gbpv=1&pg=PP1&printsec=frontcover](https://www.google.co.id/books/edition/Algoritma_Data_Mining/-Ojclag73O8C?hl=id&gbpv=1&pg=PP1&printsec=frontcover)  
 [3] Fajar Santuti Hermawati, *Data Mining*. CV Andi, 2013.  
 [4] Budi Santoso, *Data Mining Terapan Dengan Matlab*. Yogyakarta: Graha Ilmu, 2007.  
 [5] Indra Yatini B, *Flow Chart Algoritma dan Pemograman Bahasa C++*. Graha Ilmu, 2010.