

PENGELOMPOKAN DATA PASIEN TEST URINE DENGAN METODE *CLUSTERING* PADA KANTOR BADAN NARKOTIKA NASIONAL

WASIT GINTING¹⁾

Program Studi Sistem Informasi

¹⁾Fakultas Ilmu Komputer, Universitas Katolik Santo Thomas Medan
Jalan Setia Budi No. 179 F, Tanjung Sari, Medan
Email: wasitginting74@gmail.com

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) 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 BNN. 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*

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) 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. 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 di rancang dengan Bahasa pemrograman aplikasi MATLAB.*

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

1. PENDAHULUAN

Pengolahan data atau data processing merupakan manipulasi data kebentuk yang lebih informatif atau berupa informasi. Informasi merupakan hasil dari kegiatan pengolahan suatu data dalam bentuk tertentu yang lebih berarti dari suatu kegiatan atau suatu peristiwa untuk membantu memberikan informasi bagi suatu lembaga/perusahaan ataupun perorangan. Komputer merupakan sarana yang tepat untuk mempermudah segala urusan disegala bidang, terutama di instansi pemerintahan, di sekolah, ataupun di sebuah perusahaan.

Seiring dengan perkembangan teknologi yang begitu pesat, kebutuhan informasi sangat diperlukan, terlebih lagi informasi yang dihasilkan mengandung nilai yang benar, akurat, tepat, dan tentunya efisien dalam waktu. Sehingga siapapun dan apapun yang menggunakan teknologi informasi tersebut dapat menangani berbagai masalah dengan waktu yang singkat.

Semakin meningkatnya tuntutan masyarakat pada lembaga-lembaga pemerintahan untuk dapat memberikan pelayanan yang baik dan berkualitas, diperlukan suatu media informasi untuk memenuhinya. Dengan penerapan media informasi diharapkan sebuah lembaga pemerintahan dalam segala kegiatannya dapat menciptakan pelayanan kepada semua pihak.

Sebagai salah satu lembaga pemerintahan yang memiliki kewajiban dalam bidang narkotika perlu adanya suatu informasi data untuk memberikan kemudahan dalam pengelompokan data yang berkaitan dengan tes urin masyarakat, Badan Narkotika Nasional adalah lembaga masyarakat dalam pelayanan mengenai dunia narkoba. Banyak yang dapat kita lakukan dikantor BNN ini, mulai dari mengetahui berita tentang narkoba,

penyuluhan tentang narkoba serta melakukan tes urin yang akan digunakan oleh berbagai kepentingan. BNN bekerja melihat dari data masyarakat, data pejabat, data pelajar dan juga informasi berbagai kasus narkoba dikalangan masyarakat supaya masyarakat dapat mengetahui berbagai informasi mengenai narkoba. Untuk mengatasi permasalahan ini, maka BNN mengelompokkan beberapa kategori terkait dengan tes urin guna mempermudah pengelompokan data dengan akurat.

2. METODOLOGI PENELITIAN

2.1 Pengertian Data Mining

Data *mining* adalah suatu istilah yang digunakan untuk menguraikan penemuan pengetahuan di dalam *database*. Data *mining* adalah proses yang menggunkan teknik statistik, matematika, kecerdasan buatan, dan *machine learning* untuk menginteraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai *database* besar. (Turban, dkk.2005) Definisi umum dari data *mining* itu sendiri adalah proses pencarian pola – pola yang tersembunyi (*hidden patern*) berupa pengetahuan (*konowlegde*) yang diketahui sebelumnya dari suatu sekumpulan data yang mana data tersebut dapat berada didalam *database*, data *werehoude*, atau media penyimpanan informasi yang lain. Hal penting yang terkait dalam data *mining* adalah:

- 1.Data *mining* merupakan suatu proses otomatis terhadap data yang sudah ada.
- 2.Data yang akan diproses berupa data yang sangat besar.
- 3.Tujuan data *mining* adalah mendapatkan hubungan atau pola yang mungkin memberikan indikasi yang bermanfaat (Kusrini dan Emha Taufiq, 2009).

Data *mining* dilakukan dengan *tool* khusus yang mengeksekusi operasi data *mining* yang telah didefinisikan berdasarkan model analisa. Data *mining* merupakan proses

analisis terhadap data penekanan menemukan informasi yang tersembunyi pada sejumlah data besar yang disimpan ketika menjalankan bisnis perusahaan. Kemajuan luar biasa yang terus berlanjut dalam bidang data *mining* didorong oleh beberapa faktor antara lain:

1. Pertumbuhan yang cepat dalam kumpulan data.
2. Penyimpanan data dalam *warehouse*, sehingga sejumlah perusahaan memiliki akses kedalam *database* yang andal.
3. Adanya peningkatan akses data melalui navigasi web dan internet.
4. Tekanan kompetisi bisnis untuk meningkatkan penguasaan pasar dalam globalisasi ekonomi.
5. Perkembangan teknologi perangkat lunak untuk data *mining* (ketersediaan teknologi).
6. Perkembangan yang hebat dalam kemampuan komputasi dan kemampuan kapasitas media penyimpanan (Larose, 2005) istilah data *mining* dan *knowledge discovery in database* (KDD) sering digunakan secara bergantian untuk menjelaskan proses penggalian informasi tersembunyi dalam suatu basis data yang besar. Sebenarnya kedua istilah tersebut memiliki konsep yang berbeda, tetapi berkaitan satu sama lainnya. Salah satu ter seluruhnya proses KKD adalah data *mining*.

Proses KKD itu ada 5 tahapan yang dilakukan secara terurut, yaitu:

1. *Data selection* Pemilihan (seleksi) data dari sekumpulan data operasional perlu dilakukan sebelum tahap penggalian informasi dalam KKD dimulai. Data hasil seleksi yang digunakan untuk proses data *mining*, disimpan dalam surat berkas, terpisah dari data operasional.
2. *Pre-processing / cleaning* Sebelum proses data *mining* dapat dilaksanakan, perlu dilakukan proses *cleaning* pada data yang menjadi

fokus KKD Proses *cleaning* mencakup antara lain membuang duplikasi data, memeriksa data yang inkonsisten, dan memperbaiki kesalahan pada data.

3. Transformasi *Coding* adalah proses transformasi pada data yang telah dipilih, sehingga data tersebut sesuai untuk proses data *mining*. Proses *coding* dalam KKD merupakan proses kreatif dan tergantung pada jenis atau pada informasi yang akan dicari dalam basis data.
4. Data *mining* adalah proses mencari pada atau informasi menarik dalam data terpilih dengan menggunakan teknik, metode, atau algoritma dalam data *mining* sangat bervariasi. Pemilihan metode atau algoritma yang tepat sangat bergantung pada tujuan proses KKD secara keseluruhan.
5. *Interpretation / evaluation* pola informasi yang dihasilkan dari proses data *mining* perlu ditampilkan dalam bentuk yang mudah oleh pihak yang berkemungkinan. Tahap ini merupakan bagian proses KKD yang disebut *interpretation*. Tahap ini mencakup pemeriksaan apakah pola atau informasi yang ditentukan bertentangan dengan fakta atau hipotesis yang ada sebelumnya (Fayyad, 1996).

2.2 Pengertian Tes Urin

Tes urin dilakukan guna mengetahui adanya berbagai komponen dalam urin, hasil produk limbah yang dibuat oleh ginjal. Tes ini biasanya dilakukan untuk beberapa alasan. Tes urin sangat umum dilakukan dan dapat dilaksanakan diruang praktik dokter, instalasi gawat darurat, laboratorium, dan rumah sakit bahkan di kantor Badan Narkotika Nasional yang terdapat di seluruh kota di Indonesia untuk mengetahui apakah kita positif terhadap narkoba atau tidak.

Lebih dari 100 tes berbeda dapat dilakukan pada urin. Urin dapat diteliti berdasarkan penampilan fisiknya (warna, kejernihan, bau), pH (tingkat asam dan basa), adanya glukosa (gula), nitrit, sel darah putih dan merah, kristal, bakteri dalam urin, dan lain-lain.

Tes urin yang dilakukan di Kantor Badan Narkotika Nasional kepada masyarakat dilakukan untuk berbagai kepentingan seperti untuk masuk sekolah militer, melamar kerja ke sebuah perusahaan, masuk PNS atau hal lainnya yang dibutuhkan dalam keadaan tertentu.



Gambar 1. Alat Test urine

Cara kerja alat test urine 5 parameter :
Test narkoba 5 parameter kombinasi Mop, Met, Bzo, THC, Amp

1. *Morfin.*
2. *Metamfetamin.*
3. *Benzodiazepin.*
4. *Kokain.*
5. *Ampetamin.*

Cara menggunakan sebagai berikut:

1. Keluarkan alat uji dari sachetnya dan buka penutup pada bagian ujung alat uji. Perhatikan tanda panah dan celupkan kedalam spesimen urin selama 10-15 detik. Jangan celupkan melewati batas tanda panah.
2. Tempatkan alat uji pada permukaan datar yang tidak bersifat menyerap cairan, mulai menghitung waktu pengujian dan tunggu hingga muncul garis berwarna merah pada membran pengujian. Hasil dibaca pada menit kelima, jangan membaca hasil setelah 10 menit.

3. Pembacaan hasil berupa garis berwarna merah–ungu dengan arti: pada posisi C dan T: *negatif*, pada posisi C: *positif*, garis C tidak muncul *invalid*.

2.3 Pengertian Narkoba

Narkoba atau narkotika dan obat-obatan berbahaya adalah bahan kimia baik sintetik ataupun organik yang merusak kerja saraf. Pengertian narkoba oleh kementerian kesehatan diartikan sebagai NAPZA. Narkotika, Psikotropika, dan Zat Adiktif. Narkoba dapat menyebabkan ketagihan, terganggu pada bagian saraf dan atau mampu tidak sadarkan diri.

Pengertian Narkotika secara umum adalah obat-obatan yang mampu membius. Dengan kata lain, narkotika adalah obat-obatan yang mampu mengganggu sistem kerja saraf tubuh untuk tidak merasakan sakit atau rangsangan. Narkotika pada awalnya ada tiga yang terbuat dari bahan organik yaitu Candu (*Papaver Somniferum*), kokain (*Erythroxylon coca*) dan ganja (*Cannabis sativa*). Sekarang narkoba jenis narkotika adalah *Opium* atau *Opioid* atau *Opiat* atau Candu, *Codein*, *Methadone* (MTD), LSD, PC, *mescaline*, *barbiturat*, *demerol*, *petidin*, dan lainnya. Beberapa gambar dari jenis narkoba:



Gambar 2. Narkoba (Narkotika) Daun Ganja, Ganja olah, dan lenting ganja



Gambar 3. Narkoba jenis Kokain

Psikotropika adalah salah satu narkoba. Pengertian *psikotropika* adalah segala narkoba yang tidak mengandung narkotik atau narkoba yang tidak menyebabkan hilang rasa sakit akan tetapi memberikan efek ketagihan dan terganggunya saraf akan tetapi bukan pada bagian rasa sakit. Narkoba jenis ini mampu mengubah mental dan tingkah laku penggunanya. Contoh dari *psikotropika* adalah *Ekstasi* atau *Inex* atau *Metamphetamines*, *Demerol*, *Speed*, *Angel Dust*, Sabu-sabu (*Shabu/Syabu/ICE*), *Sedatif-Hipnotik Benzodiazepin/BDZ*, *BK*, *Lexo*, *MG*, *Rohip*, *Dum*, *Megadon*, *Nipam*. Akan tetapi, setelah munculnya UU No. 35 tahun 2009, narkoba jenis *psikotropika* hanya pada golongan III dan IV. Jenis *psikotropika* adalah *Sedatin* (Pil BK), *Rohypnol*, *Magadon*, *Valium*, *Mandrax*, *Amfetamine*, *Fensiklidin*, *Metakualon*, *Metifenidat*, *Fenobarbital*, *Flunitrazepam*, *Ekstasi*, *Shabu-shabu*, *LSD* (*Lycergic Syntetic Diethylamide*) dan sebagainya.

2.4 Pengertian Clustering

Clustering adalah metode penganalisis data, yang sering dimasukkan sebagai salah satu metode *data mining*, yang bertujuan adalah untuk mengelompokkan data dengan karakteristik yang sama ke suatu ‘wilayah’ yang sama dan data dengan karakteristik yang berbeda ‘wilayah’ yang lain. Ada beberapa pendekatan yang digunakan dalam pengembangan metode *clustering*.

Dua pendekatan utama adalah *clustering* dengan pendekatan partisi dan *clustering* dengan pendekatan hirarki (Oliviera et al, 2007) *Clusterring* pgnsn pendekatan partisi atau sering disebut dengan *partition-based clustering* mengelompokkan data dengan memilah – milah data yang dianalisa kedalam *clustering* yang ada. *Clustering* dengan pendekatan hirarki atau sering disebut dengan *hieratchical clustering* mengelompokkan data dengan membuat suatu hirarki berupa dengdogram dimana yang mirip akan ditempatkan pada hirarki yang berdekatan dan yang tidak pada hirarki yang berjauhan. Di samping kedua pendekatan tersebut, ada juga *clustering* dengan pendekatan *automatic mapping* (*Self-Organising Map/SOM*).

2.5 Clustering Dengan Pendekatan Partisi

Salah satu metode yang banyak digunakan dalam melakukan *clustering* dengan partisi ini dalah dengan metode *k-means*. 5 secara umum metode *k-means* ini melakukan proses pengelompokan dengan prosedur sebagai berikut (Maimon et al, 2010).

1. Tentuka jumlah *cluster*.
2. Alokasi data secara random ke *cluster* yang ada.
3. Harga rata – rata setiap *cluster* dari data yang tergabung di dalamnya.
4. Alokasi Kembali semua data ke *cluster* terdekat.
5. Ulang prose no 3, sampai tidak ada perubahan atau perubahan yang terjadi masih sudah dibawah *tes hold* prosedur dasar ini bisa berubah mendekati pendekatan pengelokasian data yang diterapkan, apakah *crisp* atau *fuzy*. Setelah meneliti *clustering* dari sudut yang lain, ditemukan bahwa *k-means clustering* mempunyai beberapa kelemahan.

Mixture Modelling (*Mixture Modeling*) merupakan metode pengelompokan data yang mirip dengan *k-*

means dengan kelebihan penggunaan distribusi statistic dalam mendefinisikan setiap *cluster* yang ditemukan. Dibanding dengan *k-meas* yang hanya menggunakan *cluster center*, penggunaan distribusi statistic ini mengijinkan kita untuk (Hastie et al, 2010):

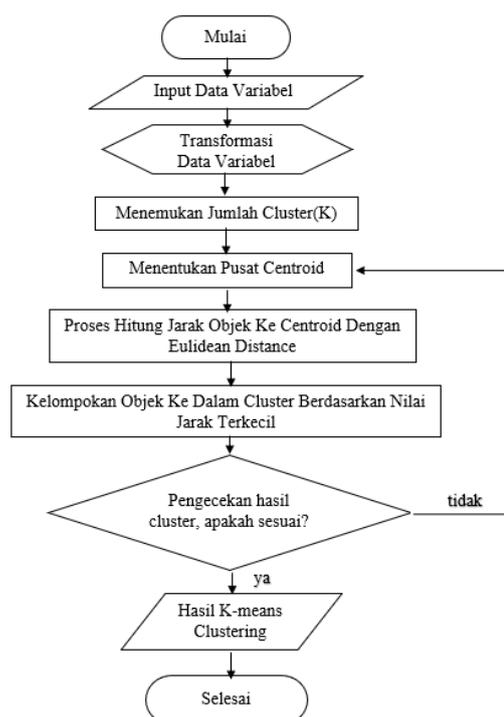
1. Metode data yang kita miliki dengan seting karakteristik yang berbeda – beda.
 2. Jumlah cluster yang sesuai dengan keadaan data bisa ditentukan seiring dengan proses pemodelan karakteristik dari masing – masing cluster
- 3 Hasil pemodelan *clustering* yang dilaksanakan bisa diuji tingkat keakuratannya. Distribusi statistic yang digunakan bisa bermacam – macam mulai dari yang digunakan untuk data *categorical* sampai yang digunakan untuk data *continuous*, termasuk diantaranya distribusi *binomial*, *multinomial*, normal dan lain – lain. Beberapa distribusi yang bersifat tidak normal seperti distribusi *Poisson*, *von-Mises*, *Gamma* dan *Student*, juga 6 diterapkan untuk bisa mengakomodasi berbagai keadaan data yang ada di lapangan. Beberapa pendekatan *multivariate* juga banyak diterapkan untuk memperhitungkan tingkat keterkaitan antara *variable* data yang satu dengan lainnya.

3. HASIL DAN PEMBAHASAN

3.1 Perancangan Sistem

Berdasarkan analisis yang dilakukan, maka penulis akan menggali data untuk pengelompokan dengan kriteria-kriteria (variabel) yang telah ditentukan.

Adapun kriteria (variabel) yang diambil dari penelitian tersebut adalah Usia , Alamat dan Hasil Tes. Kemudian dari data tersebut dimasukan kedalam matlab sehingga terdapat menjadi hasil pelompok serta dapat dilihat pada alur sistem yang dirancang. Adapun *flowchart* yang telah dirancang adalah seperti dibawah ini:

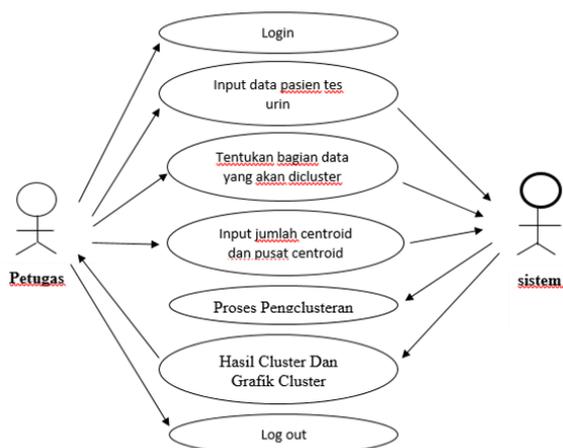


Gambar 4. Flowchart Algoritma K-means

Keterangan :

1. Input data variabel dan Transformasi data variabel
2. Tentukan jumlah cluster (K) sebagai input, apakah 2 cluster atau lebih.
3. Tentukan pusat/centroid secara random
4. Lakukan perhitungan jarak engan Eulidean Distance.
5. Lakukan pengelompokan grup berdasarkan nilai terdekat.
6. Selanjutnya adalah selisih pusat cluster baru dengan cluster lama? Jika ada maka proses akan berulang ke penentuan pusat/centroid perhitungan jarak dan pengelompokan group sampai tidak ada selisih.
7. Output hasil pehitungan dengan clustering *k-Means*, Selesai.

Berikut dibawah ini rancangan use case yang dirancang dapat di lihat dibawah ini:



Gambar 5. Use Case Sistem Data Mining Pengelompokan Pasien tes Urin

Tabel 1. Data Pengelompokan Pasien Tes Urin pada Kantor Badan Narkotik Nasional yang akan di clustering

| No | Nama Pasien Tes urin | Usia | Alamat | Hasil Tes |
|----|----------------------|------|----------------|---------------------|
| 1 | Abdul Rahman | 15 | Binjai Selatan | Met(Metamfetamin) |
| 2 | Heriyanto | 41 | Binjai Kota | Mop(Morfin) |
| 3 | Rahmat Setiawan | 17 | Binjai Barat | Bzo(Benzodiazepine) |
| 4 | Suriyadi Wijaya | 22 | Binjai Timur | Negatif |
| 5 | Robi Andika Purba | 31 | Binjai Selatan | Met(Metamfetamin) |
| 6 | Suhendri | 47 | Binjai Selatan | Bzo(Benzodiazepine) |
| 7 | Supriadi | 42 | Binjai Kota | Mop(Morfin) |
| 8 | Iskandar Dinata | 22 | Binjai Timur | Negatif |
| 9 | Agus Novianta | 33 | Binjai Utara | THC(Kokain) |
| 10 | Harun Pratama | 20 | Binjai Barat | Met(Metamfetamin) |
| 11 | Irwansyah Sembiring | 19 | Binjai Selatan | Bzo(Benzodiazepine) |
| 12 | Eko Siswanto | 24 | Binjai Utara | THC(Kokain) |
| 13 | Lilik Erawan | 48 | Binjai Timur | Met(Metamfetamin) |
| 14 | Dika Mahesa Ambarita | 27 | Binjai Timur | Mop(Morfin) |
| 15 | Jaka Perdana | 25 | Binjai Selatan | Bzo(Benzodiazepine) |
| 16 | Gilang Gumilar | 48 | Binjai Selatan | Bzo(Benzodiazepine) |
| 17 | Tito Afandy | 43 | Binjai Kota | Negatif |
| 18 | Putra Harun Siregar | 33 | Binjai Utara | Bzo(Benzodiazepine) |
| 19 | Niko Darmawan | 19 | Binjai Timur | Mop(Morfin) |
| 20 | Suwandi | 47 | Binjai Timur | Mop(Morfin) |

Tabel 2 Inisialisasi Kriteria Usia(X)

| Kode | Usia |
|------|---------------|
| 1 | 15 - 20 Tahun |
| 2 | 21 - 30 Tahun |
| 3 | 31 - 40 Tahun |
| 4 | 41 - 46 Tahun |
| 5 | 47 - 48 Tahun |

Tabel 3 Inisialisasi Alamat(Y)

| Kode | Alamat |
|------|----------------|
| 1 | Binjai Selatan |
| 2 | Binjai Kota |
| 3 | Binjai Barat |
| 4 | Binjai Timur |
| 5 | Binjai Utara |

Tabel 4 Inisialisasi Hasil Tes(Z)

| Kode | Hasil Tes |
|------|--------------------|
| 1 | Negatif |
| 2 | Mop(Morfin) |
| 3 | Met(Metamfetamin) |
| 4 | Bzo(Benzodiazepin) |
| 5 | Thec(Kokain) |

Tabel 5. Data Yang Telah Di Transformasikan Berdasarkan Pengkodean

| No | Nama | Usia | Alamat | Hasil Tes |
|----|------|------|--------|-----------|
| 1 | A | 1 | 1 | 3 |
| 2 | B | 4 | 2 | 2 |
| 3 | C | 1 | 3 | 4 |
| 4 | D | 2 | 4 | 1 |
| 5 | E | 3 | 1 | 3 |
| 6 | F | 5 | 1 | 4 |
| 7 | G | 4 | 2 | 2 |
| 8 | H | 2 | 4 | 1 |

| | | | | |
|----|---|---|---|---|
| 9 | I | 3 | 5 | 5 |
| 10 | J | 1 | 3 | 3 |
| 11 | K | 1 | 1 | 4 |
| 12 | L | 2 | 5 | 5 |
| 13 | M | 5 | 4 | 3 |
| 14 | N | 2 | 4 | 2 |
| 15 | O | 2 | 1 | 4 |
| 16 | P | 5 | 1 | 4 |
| 17 | Q | 4 | 2 | 1 |
| 18 | R | 3 | 5 | 5 |
| 19 | S | 1 | 4 | 2 |
| 20 | T | 5 | 4 | 2 |

Langkah-langkah selanjutnya meliputi:

1. Menentukan jumlah cluster

Jumlah cluster yang digunakan adalah 3.

2. Menentukan Pusat *Cluster* (Centroid) secara acak (*Random*)

Adapun proses pusat *Cluster* yang dipilih adalah:

Centroid 1=(2,4,1) diambil dari data No.4

Centroid 2=(5,4,3) diambil dari data No.13

Centroid 3=(1,4,2) diambil dari data No.19

ITERASI I :

Bagian A (1,1,3)

C1 (X=2,4,1)

$$= \sqrt{(1-2)^2 + (1-4)^2 + (3-1)^2} = 3.74$$

C2 (Y=5,4,3)

$$= \sqrt{(1-5)^2 + (1-4)^2 + (3-3)^2} = 5$$

C3 (Z=1,4,2)

$$= \sqrt{(1-1)^2 + (1-4)^2 + (3-2)^2} = 3.16$$

Bagian B (4,2,2)

C1 (X=2,4,1)

$$= \sqrt{(4-2)^2 + (2-4)^2 + (2-1)^2} = 3$$

C2 (Y=5,4,3)

$$= \sqrt{(4-5)^2 + (2-4)^2 + (2-3)^2} = 2.44$$

C3 (Z=1,4,2)

$$= \sqrt{(4-1)^2 + (2-4)^2 + (2-2)^2} = 3.60$$

Bagian C (1,3,4)

C1 (X=2,4,1)

$$= \sqrt{(1-2)^2 + (3-4)^2 + (4-1)^2} = 3.31$$

C2 (Y=5,4,3)

$$= \sqrt{(1-5)^2 + (3-4)^2 + (4-3)^2} = 4.24$$

C3 (Z=1,4,2)

$$= \sqrt{(1-1)^2 + (3-4)^2 + (4-2)^2} = 2.23$$

Bagian D (2,4,1)

C1 (X=2,4,1)

$$= \sqrt{(2-2)^2 + (4-4)^2 + (1-1)^2} = 0$$

C2 (Y=5,4,3)

$$= \sqrt{(2-5)^2 + (4-4)^2 + (1-3)^2} = 3.60$$

C3 (Z=1,4,2)

$$= \sqrt{(2-1)^2 + (4-4)^2 + (1-2)^2} = 1.41$$

Dan sampai selesai perhitungan sesuai data yang sudah di gunakan, dimana hasil interasi 1 dapat di simpulkan hasilnya adalah :

Setelah dilakukan perhitungan menggunakan rumus cluster dan hasil group dilihat dari nilai yang terkecil antara C1, C2, dan C3, maka group berdasarkan jarak minimal centroid terdekat adalah :

3, 1, 2, 2, 2, 1, 2, 1, 3, 3, 2, 1, 3, 2, 1, 2, 3, 2}

Grup Iterasi II : {3, 2, 3, 1, 2, 2, 2, 1, 2, 3, 3, 3, 2, 1, 3, 2, 1, 2, 1, 2}

Terjadi perubahan pada group, maka dilanjutkan ke perhitungan Iterasi berikutnya yaitu Iterasi 3.

ITERASI III :

Dalam hasil perhitungan Iterasi 2, maka untuk mencari Iterasi 3 ditentukan dengan perhitungan berikut ini.

Untuk Grup 1 ada 5 data yaitu :

$$C1 (X) = (2+2+2+4+1) / 5 = 2.2$$

$$C2 (Y) = (4+4+4+2+4) / 5 = 3.6$$

$$C3 (Z) = (1+1+2+1+2) / 5 = 1.4$$

Untuk Grup 2 ada 9 data yaitu :

$$C1 (X) = (4+3+5+4+3+5+5+3+5) / 9 = 4.11$$

$$C2 (Y) = (2+1+1+2+5+4+1+5+4) / 9 = 2.77$$

$$C3 (Z) = (2+3+4+2+5+3+4+5+2) / 9 = 3.33$$

Untuk Grup 3 ada 6 data yaitu :

$$C1 (X) = (1+1+1+1+2+2) / 6 = 1.33$$

$$C2 (Y) = (1+3+3+1+5+1) / 6 = 2.33$$

$$C3 (Z) = (3+4+3+4+5+4) / 6 = 3.83$$

Diperoleh centroid baru untuk literasi 3 yaitu :

$$C1 = (2.2, 3.6, 1.4)$$

$$C2 = (4.11, 2.77, 3.33)$$

$$C3 = (1.33, 2.33, 3.83)$$

Setelah didapat hasil dari C1, C2, dan C3 dalam perhitungan di atas, maka selanjutnya akan dihitung untuk Iterasi berikutnya, yaitu Iterasi 3. Berikut ini merupakan perhitungan dari Iterasi 3 antara lain :

Bagian A (1, 1, 3)

$$C1 (2.2, 3.6, 1.4) =$$

$$\frac{\sqrt{(1 - 2.2)^2 + (1 - 3.6)^2 + (3 - 1.4)^2}}{2} = 3.28$$

$$C2 (4.11, 2.77, 3.33) =$$

$$\frac{\sqrt{(1 - 4.11)^2 + (1 - 2.77)^2 + (3 - 3.33)^2}}{2} = 3.59$$

$$C3 (1.33, 2.33, 3.83) =$$

$$\frac{\sqrt{(1 - 1.33)^2 + (1 - 2.33)^2 + (3 - 3.83)^2}}{2} = 1.60$$

Bagian B (4, 2, 2)

$$C1 (2.2, 3.6, 1.4) =$$

$$\frac{\sqrt{(4 - 2.2)^2 + (2 - 3.6)^2 + (2 - 1.4)^2}}{2} = 2.48$$

$$C2 (4.11, 2.77, 3.33) =$$

$$\frac{\sqrt{(4 - 4.11)^2 + (2 - 2.77)^2 + (2 - 3.33)^2}}{2} = 1.54$$

$$C3 (1.33, 2.33, 3.83) =$$

$$\frac{\sqrt{(4 - 1.33)^2 + (2 - 2.33)^2 + (2 - 3.83)^2}}{2} = 3.25$$

Bagian C (1, 3, 4)

$$C1 (2.2, 3.6, 1.4) =$$

$$\frac{\sqrt{(1 - 2.2)^2 + (3 - 3.6)^2 + (4 - 1.4)^2}}{2} = 2.92$$

$$C2 (4.11, 2.77, 3.33) =$$

$$\frac{\sqrt{(1 - 4.11)^2 + (3 - 2.77)^2 + (4 - 3.33)^2}}{2} = 3.19$$

$$C3 (1.33, 2.33, 3.83) =$$

$$\frac{\sqrt{(1 - 1.33)^2 + (3 - 2.33)^2 + (4 - 3.83)^2}}{2} = 0.76$$

Bagian D (2, 4, 1)

$$C1 (2.2, 3.6, 1.4) =$$

$$\frac{\sqrt{(2 - 2.2)^2 + (4 - 3.6)^2 + (1 - 1.4)^2}}{2} = 0.60$$

$$C2 (4.11, 2.77, 3.33) =$$

$$\frac{\sqrt{(2 - 4.11)^2 + (4 - 2.77)^2 + (1 - 3.33)^2}}{2} = 3.37$$

$$C3 (1.33, 2.33, 3.83) =$$

$$\frac{\sqrt{(2 - 1.33)^2 + (4 - 2.33)^2 + (1 - 3.83)^2}}{2} = 3.35$$

Bagian E (3, 1, 3)

$$C1 (2.2, 3.6, 1.4) =$$

$$\frac{\sqrt{(3 - 2.2)^2 + (1 - 3.6)^2 + (3 - 1.4)^2}}{2} = 3.15$$

$$C2 (4.11, 2.77, 3.33) =$$

$$\frac{\sqrt{(3 - 4.11)^2 + (1 - 2.77)^2 + (3 - 3.33)^2}}{2} = 2.12$$

$$C3 (1.33, 2.33, 3.83) =$$

$$\frac{\sqrt{(3 - 1.33)^2 + (1 - 2.33)^2 + (3 - 3.83)^2}}{2} = 2.19$$

Kemudian akan di hitung Sampai proses iterasi ke 3 selesai.

Setelah di lakukan perhitungan menggunakan rumus cluster yang ada, pada

iterasi ke-2 dan iterasi ke-3 posisi *cluster* tidak berubah atau tidak ada data yang berpindah grup lagi, maka perhitungan dapat dihentikan. Adapun hasil grup yang diperoleh dari perhitungan iterasi ke-2 dan ke-3 adalah sebagai berikut :

Grup Iterasi II : {3, 2, 3, 1, 2, 2, 2, 1, 2, 3, 3, 3, 2, 1, 3, 2, 1, 2, 1, 2}

Grup Iterasi III : {3, 2, 3, 1, 2, 2, 2, 1, 2, 3, 3, 3, 2, 1, 3, 2, 1, 2, 1, 2}

Berikut contoh hasil iterasi 2 dan iterasi ke 3 dapat dilihat dibawah ini:

| No | Object | X | Y | Z | C1 | C2 | C3 | Group |
|----|--------|---|---|---|------|------|------|-------|
| 1 | A | 1 | 1 | 3 | 3.28 | 3.59 | 1.60 | 3 |
| 2 | B | 4 | 2 | 2 | 2.48 | 1.54 | 3.25 | 2 |
| 3 | C | 1 | 3 | 4 | 2.92 | 3.19 | 0.76 | 3 |
| 4 | D | 2 | 4 | 1 | 0.60 | 3.37 | 3.35 | 1 |
| 5 | E | 3 | 1 | 3 | 3.15 | 2.12 | 2.29 | 2 |
| 6 | F | 5 | 1 | 4 | 4.62 | 2.09 | 3.90 | 2 |
| 7 | G | 4 | 2 | 2 | 2.48 | 1.54 | 3.25 | 2 |
| 8 | H | 2 | 4 | 1 | 0.60 | 3.37 | 3.35 | 1 |
| 9 | I | 3 | 5 | 5 | 3.94 | 2.99 | 3.35 | 2 |
| 10 | J | 1 | 3 | 3 | 2.08 | 3.13 | 1.11 | 3 |
| 11 | K | 1 | 1 | 4 | 3.86 | 3.64 | 1.38 | 3 |
| 12 | L | 2 | 5 | 5 | 3.86 | 3.49 | 2.98 | 3 |
| 13 | N | 5 | 4 | 3 | 3.25 | 1.54 | 4.11 | 2 |
| 14 | M | 2 | 4 | 2 | 0.74 | 2.78 | 2.56 | 1 |
| 15 | O | 2 | 1 | 4 | 3.68 | 2.83 | 1.50 | 3 |
| 16 | P | 5 | 1 | 4 | 4.62 | 2.09 | 3.90 | 2 |
| 17 | Q | 4 | 2 | 1 | 2.44 | 2.46 | 3.90 | 1 |
| 18 | R | 3 | 5 | 5 | 3.94 | 2.99 | 3.35 | 2 |
| 19 | S | 1 | 4 | 2 | 1.40 | 3.59 | 2.50 | 1 |
| 20 | T | 5 | 4 | 2 | 2.89 | 2.01 | 4.42 | 2 |

4.KESIMPULAN

Berdasarkan hasil analisa diatas penulis membuat kesimpulan antara lain sebagai berikut :

1. Proses pengelompokan data pasien tes urin pada Badan Narkotika Nasional menggunakan metode *K-means* dengan 3 pusat centroid menghasilkan data yang di proses sebanyak 3 Iterasi kedalam grup yang mempunyai objek yang sama.

2. Penerapan *clustering* pada pengelompokan data pasien tes urin Badan Narkotika dengan data yang dikelompokan dalam *cluster-cluster* berdasarkan centroid yang telah di tentukan dengan perhitungan *k-means* menghasilkan data yang diperoleh telah memiliki tingkat keakurasian yang cukup baik dimana hasil *cluster* pada masing masing grup memiliki hasil yang berbeda sehingga data yang diberikan sangat bervariasi.

5. SARAN

Berikut saran-saran yang dapat menjadi pertimbangan dalam mengembangkan penelitian ini kedepannya.

1. Perlu dilakukan penelitian lebih lanjut terhadap penggunaan algoritma klasifikasi dan *Clustering* yang lebih baik, yakni dengan akurasi yang lebih baik pula
2. Perlu dilakukan proses penggalian variasi variabel data agar hasil *clustering* yang dihasilkan dapat lebih maksimal.
3. Sebagai bahan referensi untuk penelitian kedepan tentang *clustering* terutama menggunakan algoritma *K-Means*.

DAFTAR PUSTAKA

- [1] Adrian, A. R. (2015). Penerapan Metode K-Means Untuk Clustering Mahasiswa Berdasarkan Nilai Akademik Dengan Weka Interface Studi Kasus Pada Jurusan teknik Informatika UMM Magelang. *JURNAL ILMIAH SEMESTA TEKNIKA, Vol 18.No 01*, 76-82.
- [3]. Kusriani. (2009). *Algoritma Data Mining*. Yogyakarta: Andi.
- [4]. Mega, W. (Desember 2015). Clustering Menggunakan Metode K-means Untuk Menentukan Gizi Balita.

- Jurnal Informatika vol 15, No 2, 160-174.*
- [5]. Prasetyo, E. (2012). *Data Mining Konsep dan Aplikasi menggunakan Matlab.*
- [6]. Setiawan, Y. (September 2016). Penerapan Metode Clustering K-Means Dalam Pengelompokan Penjualan Produk. *Jurnal Media Ifotama Vol, 12, 148-157.*
- [7]. Sugiarti, Y. (2013). Analisis Dan Perancangan UML (Unified Modelling language). *Generate VB.6 Yogyakarta.*
- [8]. Taslim, F. (November 2016). Penerapan Algoritma K-Means Untuk Clustering Data Obat Pada Puskesmas Rumbai. *Jurnal Informasi & Komunikasi Digital Zone Vol 7.No2, 108-114.*