JURNAL

IMPLEMENTASI DATA MINING PENGELOMPOKKAN MURID TAMAN KANAK-KANAK DENGAN METODE *K-MEANS*

Implementation Of Data Mining Grouping Kindergarten Pupils With K-Means Method

Oleh:

YOYOK BAGUS SAKSIKO

12.1.03.02.0132

Dibimbing oleh :

1. **M. Rizal Arief, S.T., M.Kom**
2. **RISKY ASWI RAMADHANI, M.Kom**

TEKNIK INFORMATIKA

FAKULTAS TEKNIK

UNIVERSITAS NUSANTARA PGRI KEDIRI

2017

SURAT PERNYATAAN
ARTIKEL SKRIPSI TAHUN 2017

Yang bertanda tangan di bawah ini:

Nama Lengkap	: YOYOK BAGUS SAKSIKO
NPM	: 12.1.03.02.0132
Telepun/HP	: 085790776816
Alamat Surel (Email)	: saiko.yb@gmail.com
Judul Artikel	: IMPLEMENTASI DATA MINING PENGEMPLOKAN MURID TAMAN KANAK-KANAK DENGAN METODE *K-MEANS*
Fakultas – Program Studi	: Teknik - Teknik Informatika
Nama Perguruan Tinggi	: UNIVERSITAS NUSANTARA PGRI KEDIRI
Alamat Perguruan Tinggi	: Jl. KH. Achmad Dahlan No.76 Kediri (64112)

Dengan ini menyatakan bahwa :

a. artikel yang saya tulis merupakan karya saya pribadi (bersama tim penulis) dan bebas plagiarisme;

b. artikel telah diteliti dan disetujui untuk diterbitkan oleh Dosen Pembimbing I dan II.

Demikian surat pernyataan ini saya buat dengan sesungguhnya. Apabila di kemudian hari ditemukan ketidaksesuaian data dengan pernyataan ini dan atau ada tuntutan dari pihak lain, saya bersedia bertanggungjawab dan diproses sesuai dengan ketentuan yang berlaku.

Mengetahui		Kediri, 31 Januari 2017
Pembimbing I	Pembimbing II	Penulis,
M. Rizal Arief, S.T., M.Kom. NIDN. 0716027504	RISKY ASWI R., M.KOM. NIDN. 0708049001	Yoyok Bagus Saksiko 12.1.03.02.0132

Implementasi Data Mining Pengelompokkan Murid Taman Kanak-Kanak Dengan Metode *K-Means*

YOYOK BAGUS SAKSIKO
12.1.03.02.0132
Teknik – Teknik Informatika
Saiko.yb@gmail.com
M. Rizal Arief, S.T., M.Kom[1] .dan RISKY ASWI R., M.Kom.[2]

UNIVERSITAS NUSANTARA PGRI KEDIRI

ABSTRAK

Taman kanak-kanak adalah jenjang pendidikan usia dini(yakni usia 6 tahun atau dibawahnya)dalam bentuk pendidikan formal.Untuk memudah pengajaran kelas dapat dibagi menjadi 2 .Banyak cara yang dilakukan untuk menentukan pembagian murid yang sesuai dengan kelas .Salah satu cara dengan menggunakan suatu metode data mining, yaitu *K-MEANS*.

K-means merupakan salah satu algoritma *clustering* . Tujuan algoritma ini yaitu untuk membagi data menjadi beberapa kelompok. Algoritma ini menerima masukan berupa data tanpa label kelas. Dengan metode ini murid bisa masuk kelas yang sesuai.

Hasil dari penelitian ini merupakan perhitungan untuk menentukan pilihan murid taman kanak-kanak masuk kelas A atau B. Sesuai dengan kebutuhan.

KATA KUNCI : Taman kanak-kanak, Data mining, *K-means*.

I. LATAR BELAKANG

Taman Kanak-kanak merupakan salah satu bentuk pendidikan anak usia dini pada jalur pendidikan formal untuk rentang usia empat sampai dengan enam tahun. Pendidikan anak usia dini atau taman kanak-kanak, pada hakekatnya adalah pendidikan yang diselenggarakan dengan tujuan untuk memfasilitasi pertumbuhan dan perkembangan anak secara menyeluruh atau menekankan pada pengembangan seluruh aspek kepribadian anak. Oleh karena itu pendidikan untuk usia dini khususnya taman kanak-kanak perlu menyediakan berbagai kegiatan yang dapat mengembangkan berbagai aspek perkembangan yang meliputi kognitif, bahasa, fisik dan motorik.

Ketidaksamaan perkembangan kognitif dan motorik membuat murid susah belajar jika dibuat satu kelas maka dari itu saya mengelompokan menjadi 2 yaitu kelas A dan B. Dengan membuat 2 kelas maka berbeda pula dengan cara mengajarnya,kelas pertama kelas dengan murid yang belum bisa meng-eja, menghitung, mewarnai dan juga belum bisa beradaptasi. Sedangkan kelas satunya sudah bisa dan perlu pembelajaran lebih lanjut agar semakin pandai.

Pengelompokan ini mengunakan Data Mining yaitu metode *K-Means* . Metode *K-*

Means merupakan salah satu algoritma *clustering*. Tujuan algoritma ini yaitu untuk membagi data menjadi beberapa kelompok. Algoritma ini menerima masukan berupa data tanpa label kelas. Pada algoritma pembelajaran ini,komputer mengelompokkan sendiri data-data yang menjadi masukannya tanpa mengetahui terlebih dulu target kelasnya. Pembelajaran ini termasuk dalam *unsupervised learning*. Masukan yang diterima adalah data atau objek dan k buah kelompok *(cluster)* yang diinginkan. Algoritma ini akan mengelompokkan data atau objek ke dalam k buah kelompok tersebut. Pada setiap *cluster* terdapat titik pusat *(centroid)* yang merepresentasikan *cluster* tersebut (Kusrini, Emha Taufiq Luthfi, 2010). Dengan metode ini saya akan mengelompokan menJadi 2 dengan mengunakan angket yang telah saya berikan kepada kepala taman kanak-kanak untuk selanjutnya menproses hasil dari angket tersebut untuk di *clustering.*

II. METODE

A. *Action Research*

Action Research merupakan kegiatan penelitian yang difokuskan secara kolaboratif untuk menyelesaikan permasalahan yang ada, dimana permasalahan lebih kepada masyarakat secara langsung maupun tidak langsung.

A. Konsep Jarak

Jarak menjadi aspek penting dalam pengembangan metode pengklasifikasian. Untuk mengukur jarak dua titik x dan y, (d(x,y)), bisa dipergunakan konsep jarak. Ukuran jarak harus memenuhi syarat-syarat sebagai berikut :

1. $d(x,y) \geq 0$ (non negatif)

   tidak ada jarak yang mempunyai nilai negatif.

2. $d(x,y) = 0$ jika dan hanya jika x = y

   jarak antara suatu obyek antar titik dengan obyek atau titik itu sendiri adalah nol.

3. $d(x,y) \geq d(y,x)$

   jarak dari x ke y adalah sama dengan jarak y ke x.

Salah satu contoh konsep jarak adalah jarak *Euclidean* (L_2-norm). Jarak dua titik x dan y menurut *Euclidean* (Richards and Jia, 2006 dalam Syamani dan Agusta, Yudi), dirumuskan sebagai berikut :

$$d(x,y) = |x-y|^2 = \sqrt{\sum_{i=1}^{n}(x_i - y_i)^2}$$

B. *K-Means*

Metode pengklasteran secara partitioning yang memisahkan data ke dalam kelompok yang berbeda. Dengan partitioning secara iteratif, *K-Means* mampu meminimalkan rata-rata jarak setiap data ke klasternya. Dalam algoritma *K-Means*, setiap data harus termasuk ke *cluster* tertentu pada suatu tahapan proses, pada tahapan proses berikutnya dapat berpindah ke *cluster* yang lain. Pada dasarnya penggunaan algoritma K-*Means* dalam melakukan proses *clustering* tergantung dari data yang ada dan konklusi yang ingin dicapai (Agustin,2015).

memiliki nilai Pada tabel kriteria sebagai berikut A= meng-eja, B= penjumlahan, C= menulis, D= keaktifan, E= penguranagan, F= mewarnai, G= menggambar, H= mencocokan bentuk, seperti terlihat pada Tabel dibawah ini:

Tabel 2.1 Data Training

No	Nama	A	B	C	D	E	F	G	H
1	Aditya d	80	75	75	80	75	80	85	75
2	Adriano H	75	80	75	80	70	75	75	80
3	Dewi S	85	80	80	85	85	80	70	80
4	Fani R	90	85	80	85	80	85	85	85
5	Purti A	75	65	70	80	80	80	75	75
6	Lucky A	80	75	90	90	80	85	80	85
7	Tasya A	70	70	75	60	75	75	70	75
8	Safa D	90	75	80	85	80	70	80	90
9	Ilham S	60	65	75	60	80	75	75	80
10	Yuni N	75	70	80	75	80	85	70	80

Tabel 2.2 Nilai Standart Kelas Sementara

Kelas	A	B	C	D	E	F	G	H
A	75	65	70	80	75	80	75	80
B	90	85	80	85	85	90	85	85

Setelah melihat stadart tersebut mulai lah menghitung *clustering* dengan menggunakan metode *K-Means* , dengan rumus sebagai berikut:

$$d=\sqrt{(xi1-xj1)^2+(xi2-xj2)^2+(xi3-xj3)^2+(xi4-xj4)^2}$$

Di mana:

g = 2, untuk menghitung jarak Euclidean

g = ∞, untuk menghitung jarak Chebychev

x_i , x_j adalah dua buah data yang akan dihitung jaraknya

p = dimensi dari sebuah data

(Banyaknya update)= 2

Misal aditya =m1 dengan standart kelas A

M1=ska

$$=\sqrt{\begin{matrix}(xi1-xj1)^2+(xi2-xj2)^2+(xi3-xj3)^2+(xi4-xj4)^2\\+(xi5-xj5)^2+(xi6-xj6)^2+(xi7-xj7)^2+(xi8-xj8)^2\end{matrix}}$$

$$=\sqrt{\begin{matrix}(75-80)^2+(65-75)^2+(70-75)^2+(80-80)^2\\+(75-75)^2+(80-80)^2+(75-85)^2+(80-75)^2\end{matrix}}$$

$$=\sqrt{\begin{matrix}(-5)^2+(-10)^2+(-5)^2+(0)^2\\+(0)^2+(0)^2+(10)^2+(5)^2\end{matrix}}$$

$$=\sqrt{25+100+25+0+0+0+100+25}$$

$$=\sqrt{275}$$

=16,5

Misal aditya =m1 dengan standart kelas B

M1=skb

$$=\sqrt{\begin{matrix}(xi1-xj1)^2+(xi2-xj2)^2+(xi3-xj3)^2+(xi4-xj4)^2\\+(xi5-xj5)^2+(xi6-xj6)^2+(xi7-xj7)^2+(xi8-xj8)^2\end{matrix}}.1)$$

=

$$\sqrt{\begin{matrix}(90-80)^2+(85-75)^2+(80-75)^2+(85-8\ \ ^2\\+(85-75)^2+(90-80)^2+(85-85)^2+(85-75)^2\end{matrix}}$$

4

$$=\sqrt{\begin{matrix}(10)^2+(-10)^2+(5)^2+(5)^2\\+(10)^2+(10)^2+(0)^2+(10)^2\end{matrix}}$$

$=\sqrt{100+100+25+25+100+100+100}$

$=\sqrt{550}$

$=23{,}4$

Setelah selesai melakukan *clustering* dapat dilihat jika hasil tersebut lebih mendekati standart maka dipastikan peserta didik tersebut masuk kelas itu

Tabel 5.3 Hasil *Clustering* Kelas Sementara

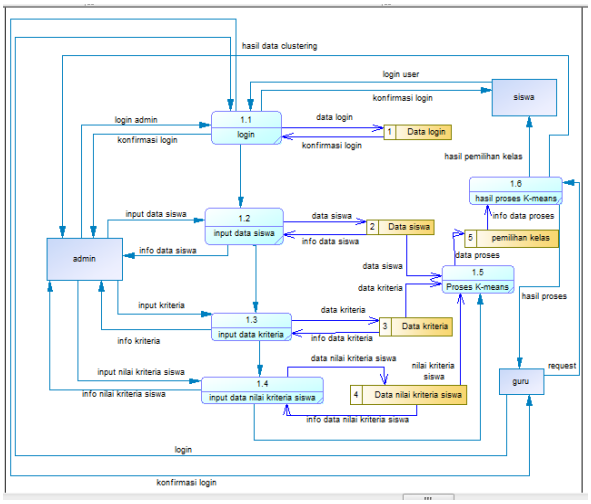
No	Nama	skA	skB	Kelas
1	Aditya D	16,5	23,4	A
2	Adriano H	18,0	29,5	A
3	Dewi S	23,9	20	B
4	Fani R	30,5	7,07	B
5	Purti A	7,07	32,7	A
6	Lucky A	26,9	20	B
7	Tasya A	23,4	43,8	A
8	Safa D	26,4	23,9	B
9	Ilham S	26,4	48,2	A
10	Yuni N	15	29,1	A

III. HASIL DAN KESIMPULAN

A. Perancangan Sistem

Dalam membangun sebuah sistem tentu diperlukan rancangan (desain) untuk menjelaskan tentang gambaran umum dari sistem untuk mengelompokan murid taman kanak-kanak sesuai kemampuannya. Berikut ini adalah perancangan dari sistem.

1. *Data Flow Diagram level* 1

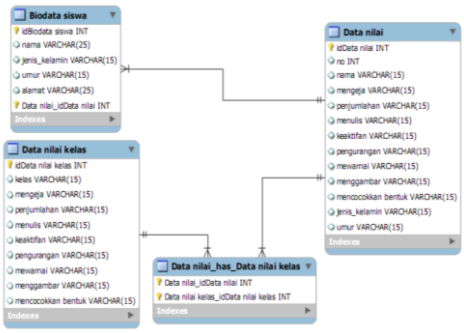


Gambar 1. ***Data Flow Diagram level* 1**

Pada DFD level 1 terdapat 5 proses yang meliputi proses *login*, *input* data siswa, *input* data nilai, proses *K-Means* dan proses pencarian data. Dalam proses *login*, TU, Guru dan Kepala Sekolah memberikan *input* ke sistem berupa *username* dan *password* kemudian inputan tersebut di proses oleh sistem apakah sudah sesuai dengan *database* hak akses *user* sistem atau belum. Jika sesuai makamenghasilkan *output* berupa hak akses sistem untuk Admin TU, siswa dan Kepala Sekolah. Selanjutnya dalam proses *input* data siswa TU memberikan inputan berupa data siswa dan jumlah *cluster* yang akan di hitung. Kemudian proses tersebut menghasilkan data siswa valid yang kemudian di simpan pada *database* data siswa dan menghasilkan *output* data siswa valid.

1. *Entity Relationship Diagram* (ERD)

Berikut adalah *Entity Relationship Diagram* (ERD) dari sistem pengelompokkan data siswa :



Gambar 5.9 ERD Aplikasi

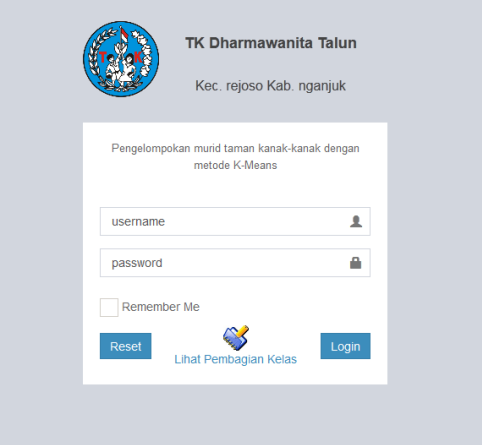
B. Tampilan Aplikasi

Tampilan program menjelaskan tentang gambaran umum dari system prediksi hasil penjualan barang pada toko sinar baru dengan *algoritma apriori*. Berikut ini adalah tampilan *input* dan *output* dari sistem.

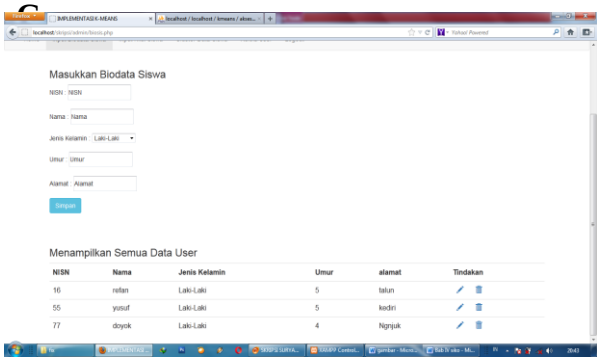
a. *Form Login*

Gambar 5.11 Halaman *Login*

Halaman ini digunakan untuk keamanan data. *User* yang akan



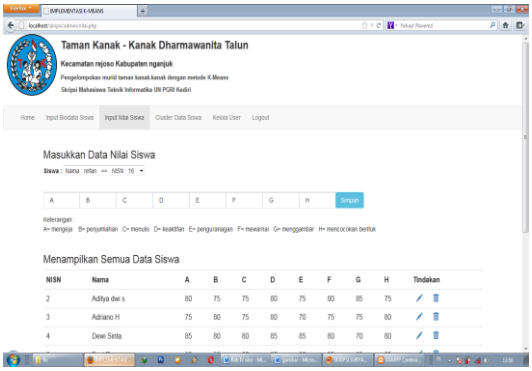
menggunakan sistem harus login terlebih dahulu dengan mengisi *username*, *password* dan *level* sesuai dengan data saat mendaftar. Penginputan data dari *form login* harus sesuai dapat dilanjutkan ke menu berikutnya.



b. *Halaman Tambah Siswa*

Gambar 5.13 Halaman Tambah Siswa

Terdapat 5 *textbox* dan 1 tombol *button*. *Textbox* NIS, Nama, jenis kelamin, umur dan alamat digunakan untuk menginputkan keterangan siswa. Terdapat *button* Simpan yang digunakan untuk menyimpan hasil inputan.Lihat Data untuk melihat hasil data



siswa yang sudah di *input*kan.

c. *Halaman Input Nilai Siswa*

Gambar 5.15 Halaman Input Nilai Siswa

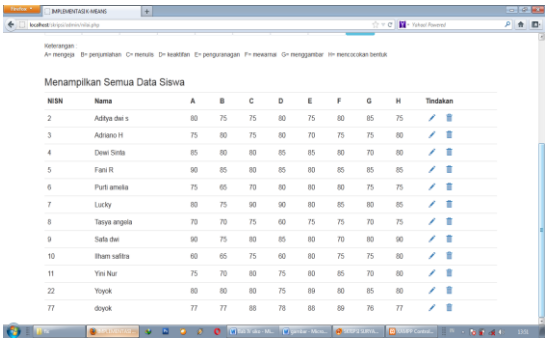
Halaman ini adalah halaman input nilai tiap siswa. Nilai yang diinputkan adalah nilai tiap mata pelajaran.

d. *Halaman Data Siswa*

Gambar 5.17 Halaman Data Siswa

Halaman ini berisi tentang data NIS dan Nama siswa yang telah diinputkan.

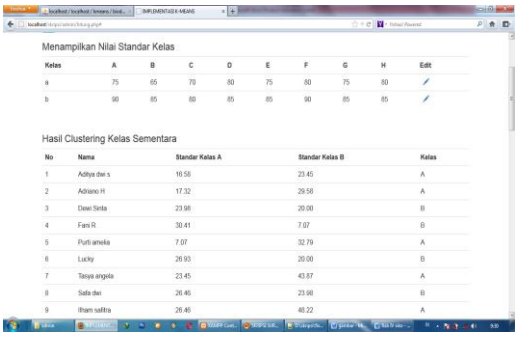
e. Halaman Nilai Pelajaran Siswa



Gambar 5.19 Halaman Nilai Pelajaran Siswa

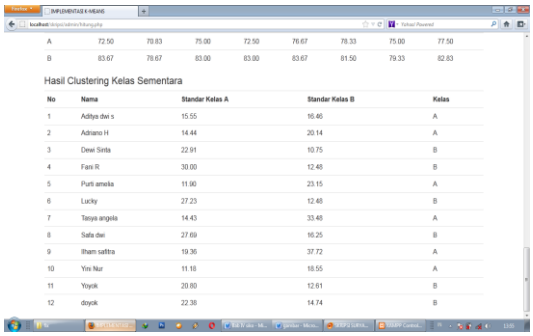
Halaman ini berisi tentang data NIS, Nama siswa dan Nilai siswa yang telah diinputkan.

f. Halaman* Centroid *Awal



Halaman ini berisi nilai centroid awal yang ditentukan secara acak dengan menganalisa data nilai siswa.

g. Halaman Hasil Perhitungan



Gambar 5.23 Halaman hasil Perhitungan

Halaman ini berisi hasil perhitungan nilai siswa menggunakan *Algoritma K-Means clustering*. Pada halaman ini dapat dilihat siswa mana saja yang masuk dalam *cluster* yang sudah ditentukan.

C. Kesimpulan

Berdasarkan pembahasan dan implementasi program yang mengacu pada rumusan masalah yang ada yaitu bagaimana merancang sistem untuk mengelompokkan data siswa menggunakan Algoritma *K-Means clustering* dan bagaimana membuat program aplikasi sistem untuk mengelompokkan data siswa menggunakan Algoritma *K-Means clustering*, maka dapat diambil beberapa kesimpulan sebagai berikut :

1. Telah dihasilkan hasil akhir dari perhitungan Algoritma *K-Means clustering* yaitu pengelompokkan murid taman kanak-kanak menjadi dua kelas yaitu A dan kelas B yang sesuai dengan keahlian murid didik.
2. Telah dihasilkan program aplikasi untuk mengelompokkan data siswa dengan menggunakan perhitungan *Algoritma K-Means clustering* yang

dimulai dari penentuan nilai centroid awal lalu dihitung jarak data nilai dengan data centroid awal sesuai dengan kelompoknya. Dihitung terus sampai rasio jarak perhitungan terbaru sama dengan rasio jarak perhitungan sebelumnya. Jika sudah sama maka perhitungan di hentikan dan sudah di dapatkan hasilnya.

IV. DAFTAR PUSTAKA

Agustin. 2015. *Implementasi Algoritma K-Means untuk Menentukan Kelompok Pengayaan Materi Mata Pelajaran Ujian Nasional (Studi Kasus : SMPN 101 Jakarta).* Jurnal Teknik Informatika Vol. 8 No. 1, April 2015.

Astuti, padmi puji.2012. *Peningkatan Kemampuan Berbahasa Anak Melalui Bercerita Dengan Papan Flannel Pada Anak Kelompok A Di Taman Kanak-Kanak Jurangjero Ii Klaten Tahun 2012* . Surakarta

Hasibuan, Zainal A. 2007. *Metodologi Penelitian pada Bidang Ilmu Komputer Dan Teknologi Informasi*. Diunduh 09 November 2015 pukul 13.18 WIB.

Indrajani. 2015. *database Design*. Jakarta : PT. Elex Media Komputindo.

Jogiyanto, Hartono.2009. *analisis dan desain sistem informasi,edisi III*. Yogyakarta: ANDI.

Kusrini, dan Emha Taufiq Luthfi.2010.*algoritma data mining*. Yogyakarta: ANDI.

Prasetyo. 2014. *Data Mining Mengolah Data Menjadi Informasi menggunakan Matlab*. Yogyakarta : Andi.

Setiawan, Nugraha.2007. *penentuan ukuran sampel memakai rumus slovin dan tabel krejicie-morgan.* Bandung.

Tampubolon, K., Saragih, H., & Reza, B. 2013. *Implementasi Data Mining Algoritma Apriori Pada Sistem Persediaan Alat-Alat Kesehatan*. Makalah disajikan dalam Majalah Ilmiah Informasi dan Teknologi Ilmiah (INTI), Volume : I, Nomor : 1, Oktober 2013.

Kuncoro, Widigdo Anom.2006. *dasar pemrograman php dan MySQL*: ilmukomputer.com.

.