ARTIKEL

CLUTERING DATA NILAI SISWA UNTUK MENENTUKAN JURUSAN PADA SISWA KELAS X SMAN 5 KEDIRI MENGGUNAKAN METODE *K-MEANS*

(Studi Kasus : SMAN 5 Kediri)

Oleh:

Jeffridin Indra Bachtiar

13.1.03.02.0362

Dibimbing oleh :

1. **Intan Nur Farida, M.Kom.**
2. **Danar Putra Pamungkas, M.Kom**

PROGRAM STUDI TEKNIK INFORMATIKA

FAKULTAS TEKNIK

UNIVERSITAS NUSANTARA PGRI KEDIRI

2017

SURAT PERNYATAAN
ARTIKEL SKRIPSI TAHUN 2017

Yang bertanda tangan di bawah ini:

Nama Lengkap	: Jeffridin Indra Bachtiar
NPM	: 13.1.03.02.0362
Telepun/HP	: 085784360510
Alamat Surel (Email)	: Jeffridinbachtiar@gmail.com
Judul Artikel	: *CLUSTERING* DATA NILAI SISWA UNTUK MENENTUKKAN JURUSAN PADA SISWA KELAS X SMAN 5 KEDIRI MENGGUNAKAN METODE *K-MEANS* (Studi Kasus: SMAN 5 KEDIRI)
Fakultas – Program Studi	: Fakultas Teknik – Teknik Informatika
Nama Perguruan Tinggi	: Universitas Nusantara PGRI Kediri
Alamat Perguruan Tinggi	: Jl. KH. Achmad Dahlan 76 Kediri

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, 09 Agustus 2017
Pembimbing I	Pembimbing II	Penulis,
Intan Nur Farida, M.Kom 0704108701	Danar Putra Pamungkas, M.Kom 0708028704	Jeffridin Indra Bachtiar 13.1.03.02.0362

CLUTERING DATA NILAI SISWA UNTUK MENENTUKAN JURUSAN PADA SISWA KELAS X SMAN 5 KEDIRI MENGGUNAKAN METODE *K-MEANS*
(Studi Kasus : SMAN 5 Kediri)

Jeffridin Indra Bachtiar
13.1.03.02.0362
FT – Teknik Informatika
Jeffridinbachtiar@gmail.com
Intan Nur Farida, M.Kom. dan Danar Putra Pamungkas, M.Kom.
UNIVERSITAS NUSANTARA PGRI KEDIRI

ABSTRAK

Penelitian ini dilatar belakangi oleh kurang tersrukturnya dalam menentukan jurusan siswa kelas X yang ada pada SMAN 5 Kediri, karena dalam penjurursan saat ini masih menggunakan sistem yang terpacu pada nilai UN SMP pada kurikulum 2013 dan angket minat siswa, hal ini menyulitkan siswa untuk menentukan jurusan sesuai kemampuanya.

Permasalahan dalam penelitian ini adalah bagaimana menerapkan data mining dengan metode *K-Means* dalam aplikasi penjurusan siswa di SMAN 5 Kediri. bagaimana merancang program aplikasi penentuan jurusan siswa SMAN 5 Kediri dengan metode *K-Means*.

Penelitian ini menggunakan metode *K-Means Clustering* untuk mendapat hasil penjurusan yang akan mempermudahkan guru atau sekolah dalam menentukan jurusan siswa. *K-means Clustering* metode *non hierarkial* adalah algoritma sederhana dan efektif untuk menemukan pengelompokan dalam data Pada proses penjurusan yang digunakan adalah data nilai dan data siswa tahun masuk 2016/2017. Adapun jurusan yang ada yaitu Ilmu Pengetahuan Alam (IPA) dan Ilmu Pengetahuan Sosial (IPS).

Kesimpulan hasil penelititan ini adalah menghasilkan sistem penentuan jurusan dengan menggunkan metode *K-Means Clustering* sesuai dengan kriteria penjurusan. Menghasilkan penjurusan siswa dari perhitungan menggunakan metode *K-means Clustering*.

Kata Kunci : Penjurusan, SMAN 5 kediri, siswa, nilai siswa, *K-Means, Clustering.*

I. LATAR BELAKANG

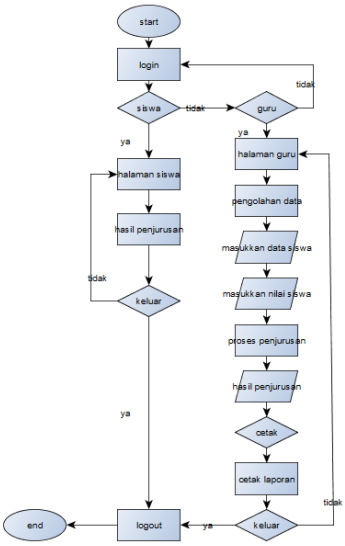
Kurikulum yang digunakan pada Sekolah Menengah Atas saat ini adalah kurikulum 2013, dalam kurikulum ini penjurusan dilakukan pada saat awal masuk, yaitu pada kelas 10. Perubahan kurikulum dimaksudkan agar memungkinkan penyesuaian program pendidikan pada satuan pendidikan dengan kondisi dan kekhasan potensi yang ada di daerah peserta didik. Penjurusan ini dimaksudkan agar siswa lebih mudah dalam memilih jurusan untuk pendidikan di kemudian hari dan mengarah ke profesi yang diinginkan.

Sistem penjurusan yang dilakukan pada awal siswa masuk SMA ini menggunakan sistem manual dengan menggunakan angket dan nilai UN SMP. Penjurusan IPA diambil dari pemilihan angket IPA, nilai IPA dan Matematika dan penjurusan IPS diambil dari pemilihan angket IPS, nilai IPS dan Matematika.

Untuk mengatasi permasalahan yang muncul pada penjurusan di SMAN 5 Kediri, peneliti memberikan solusi dengan membangun sistem penjurusan siswa dengan metode *K-Means*. Diharapkan dengan adanya sistem yang dibuat dapat membantu SMAN 5 Kediri dalam menentukan penjurusan siswa saat masuk SMA. Karena sesuai dengan kurikulum 2013, penjurusan dilakukan saat siswa baru masuk SMA.

II. METODE

A. *Flowchart* Sistem



Gambar 1 *Flowchart* Sistem

adalah proses dimulai dengan menu login sebgai guru atau siswa. Jika login sebagai siswa maka akan masuk ke dalam halaman siswa dan didalamnya akan bisa melihat hasil penjurusan. Jika login sebagai guru maka akan masuk ke halaman guru, dan akan bisa melakukan input data siswa, input nilai dan pengolahan data.

B. Simulasi Metode

K-Means merupakan metode *K-means* adalah teknik dari *clustering* yang bersifat *unsupervised leaning.*

Jarak dari data yang ada masing-masing *cluster* dihitung menggunakan rumus *Euclidean distance*[2] *:*

$$D_{L2} = (x2\,,\ x1) = \sqrt{\sum_{j-1}^{p}(x2j - x1j)^2}$$

D_{L2} = Jarak kuadrat *Euclidean* antar objek ke x2 dengan x1.

P = Jumlah Variabel *cluster.*

$x2j$ = Nilai atau data dari objek ke -2 pada *variable* ke – j.

$x1j$ = Nilai atau data dari objek ke -1 pada *variable* ke – j.

Berikut adalah perhitungan algoritma *K-Means* menggunakan dua puluh data yang berisi dua *variable* untuk tiap datanya dan dua *cluster*.

Tabel 1 *Dataset*

No	rata ipa dan UN	rata ips dan UN
1	8.135714286	8.228571429
2	8.457142857	8.178571429
3	8.035714286	8.057142857
4	8.064285714	8.078571429
5	8.635714286	8.45
6	8.25	8.25
7	8.614285714	8.635714286
8	8.142857143	8.064285714
9	8.385714286	8.435714286
10	8.342857143	8.335714286
11	8.207142857	8.25
12	8.307142857	8.2
13	8.307142857	8.321428571
14	8.535714286	8.25
15	8.05	8.128571429
16	8.15	8.121428571
17	8.028571429	8.021428571
18	8.078571429	8.192857143
19	8.292857143	8.214285714
20	8.107142857	8.028571429

Menentukkan *centroid* secara *random.*

Tabel 2 *Centroid* Awal

centroid awal		
c1	8.457142857	8.178571429
c2	8.035714286	8.057142857

Setiap data akan diukur jaraknya dengan tiap *centroid* awal menggunakan rumus *Euclidean distance* kemudian diambil jarak terpendek.

Tabel 3 Iterasi 1

No	C1	C2	jarak terpendek
1	0.325294215	0.198463486	0.198463486
2	0	0.438573755	0
3	0.438573755	0	0
4	0.405384675	0.035714286	0.035714286
5	0.324901869	0.717172737	0.324901869
6	0.219112309	0.288291944	0.219112309
7	0.483397838	0.818223561	0.483397838
8	0.334419997	0.107380688	0.107380688
9	0.266879167	0.515573784	0.266879167
10	0.194306722	0.414655008	0.194306722
11	0.260003925	0.25803417	0.25803417
12	0.151522882	0.540219154	0.151522882
13	0.207142857	0.389203792	0.207142857
14	0.106186205	0.535904728	0.106186205
15	0.410201543	0.072843136	0.072843136
16	0.312413253	0.131125427	0.131125427
17	0.456472723	0.036421568	0.036421568
18	0.378840874	0.14232042	0.14232042
19	0.16812289	0.301357473	0.16812289
20	0.380788655	0.076930926	0.076930926

Tabel 4 Kelompok *Cluster* iterasi 1

C1	C2
	1
1	
	1
	1
1	
1	
1	
	1
1	
1	
	1
1	
1	
1	
	1
	1
	1
	1
1	
	1

Menghitung *centroid* baru dengan cara menjumlahkan rata-rata dari anggota *cluster* yang telah dibentuk.

Tabel 5 *Centroid* Baru

centroid baru		
c1	0.212157274	0.495917594
c2	0.370239362	0.105923411

Hitung kembali jarak dengan setiap *centroid* baru yang baru dibentuk kemudian diambil jarak terpendek.

Tabel 6 Iterasi 2

No	c1	c2	jarak terpendek
1	0.318243482	0.102877265	0.102877265
2	0.219770391	0.497728276	0.219770391
3	0.545159319	0.126052998	0.126052998
4	0.499123144	0.078514421	0.078514421
5	0.248324751	0.612928403	0.248324751
6	0.207742107	0.236849463	0.207742107
7	0.421251208	0.721232518	0.421251208
8	0.407319408	0.035848996	0.035848996
9	0.058145089	0.422488767	0.058145089
10	0.083200061	0.355341375	0.083200061
11	0.242647533	0.187855089	0.187855089
12	0.075094326	0.48626093	0.075094326
13	0.106831549	0.326876497	0.106831549
14	0.113264462	0.504586963	0.113264462
15	0.46713331	0.051877553	0.051877553
16	0.378318102	0.063079319	0.063079319
17	0.520410066	0.11075513	0.11075513
18	0.390914804	0.037399577	0.037399577
19	0.199480995	0.281150388	0.199480995
20	0.451648503	0.030852095	0.030852095

Tabel 7 Kelompok *Cluster* iterasi 2

ipa	ips
	1
1	
	1
	1
1	
1	
1	
	1
1	
1	
	1
1	
1	
1	
	1
	1
	1
	1
1	
	1

Kemudian lakukan perbandingan antara kelompok *cluster* sebelum dan setelahnya. Setiap data harus diukur jaraknya dengan tiap-tiap *centroid* baru. Jika data masih berubah atau tidak sama dengan kelompok *cluster* sebelumnya, maka dilakukan perhitungan lagi dimulai dari penentuan *centroid* baru, ulangi

langkah kedua hingga data tidak berubah atau sama.

Karena kelompok sudah sama atau tidak mengalami perpindahan, maka proses berhenti. Ada 10 siswa masuk IPA dan ada 10 siswa masuk IPS.

III. HASIL DAN KESIMPULAN

A. Hasil

Berikut adalah hasil pengujian dari 3 skenario yang dilakukan.

Tabel 8 Hasil Pengujian

No	Jumlah data	iterasi	Centroid 1	Centroid 2
1	50	2	19	31
2	100	8	37	63
3	150	5	65	85

Skenario pengujian 1-3 dilakukan untuk mengetahui pengaruh jumlah data, pengambilan *centroid*, dan maksimum iterasi terhadap hasil klasterisasi dan waktu yang dibutuhkan untuk mengolah data. Hasil pengujian yang diperoleh, bahwa jumlah data dan jumlah *centroid* tidak berpengaruh terhadap jumlah iterasi yang dihasilkan namun berpengaruh terhadap waktu yang dibutuhkan dan keakuratan hasil klasterisasi. Selain itu pengambilan *centroid* secara *random* sangat berpengaruh terhadap hasil klasterisasi. Namun hanya saat jarak antara *centroid* 1 dengan *centroid* 2 cukup jauh, karena jika terlalu dekat bisa jadi nilai *centroid* 2 pada iterasi selanjutnya akan bertambah. Pengambilan *centroid* juga berpengaruh terhadap banyaknya jumlah iterasi yang dihasilkan, semakin banyak iterasi maka akan semakin lama waktu yang dibutuhkan.

B. Kesimpulan

1. Telah didapatkan hasil dari perhitungan dengan menggunakan *algortima K-Means Clustering* yaitu menentukan jurusan pada siswa berdasarkan nilai-nilai kelas 3 SMP semester 1 dan 2, dan nilai UN SMP. Jumlah penjurusan ada 2 yaitu Ilmu pengetahuan alam (IPA) dan Ilmu Pengetahuan Sosial (IPS). Tiap nilai siswa yang masuk akan dihitung jarak terpendek dengan centroid jurusan. Setelah itu akan dihasilkan data siswa yang masuk dalam penjurusan yang sudah di tentukan.
2. Telah dihasilkan aplikasi untuk menentukan jurusan pada siswa SMAN 5 Kediri

menggunakan metode *K-Means Clustering* yang dimula dari centroid awal kemudian hitung jarak data nilai dengan data centroid awal. Dihitung terus sampai rasio jarak perhitungan terbaru sama dengan rasio jarak perhitungan sebelumnya. Jika sudah maka perhitungan akan terhenti dan didapatkan hasil penjurusan siswa SMAN 5 Kediri.

IV. DAFTAR PUSTAKA

[1] Akbar, F. 2015. *Clustering Data Nilai Siswa SMA Untuk Penentuan Jurusan Menggunakan Algoritma Fuzzy C-Means. Jurnal Techno Nusa Mandiri.* (Online), XII (1): 56-63, tersedia: http://ejournal.nusamandiri.ac.id, diunduh 20 Oktober 2016.

[2] Andrian, R. A. 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 Tehnika. 18 (1): 76-82, tersedia: http//journal.umy.ac.id, diunduh 21 Oktober 2016.

[3] Kadir, A. 2008. *Dasar Pemograman Web Dinamis Menggunakan Php.* Yogyakarta: Cv Andi Offiset.

[4] Kristanto, N. Harianto. 2016. *Implementasi K-Means Clustering untuk Pengelompokan Analisis Rasio Profitabilitas dalam Working Capital,* tersedia: http//journal.uc.ac.id, diunduh 20 Oktober 2016.

[5] Kusrinu & Lutfhi, E.T. 2009. *Algoritma Data Mining.* Yogyakarta: Cv Andi Offiset.

[6] Nasari, Fina. 2015. *Penerapan K-Means Clustering Pada Data Penerimaan Mahasiswa Baru.* Jurnal ini disajikan dalam Seminar Nasional Teknologi Informasi dan Multimedia. 73-78, tersedia: http//ojs.amikom.ac.id, diunduh 20 Oktober 2016.

[7] Peraturan Pemerintah RI No. 17 Tahun 2010 tentang Pengelolaan dan Penyelenggaraan Pendidikan. (Online), tersedia: http://kelembagaan.ristekdikti.go.id diunduh 15 November 2016.

[8] Sulistiani. 2015. *Sistem Penentuan Jurusan Sekolah Menengah Atas Negeri 1 Karangmojo.* Seminar nasional teknologi informasi dan multimedia STMIK AMIKOM Yogyakarta.

[9] Utomo. 2015. *Implementasi Metode K-Means Clustering untuk Analisa Prestasi Siswa Berdasarkan Data Siswa di SMA Negeri Grogol*.simki.unpkediri.ac.id