ARTIKEL

PENERAPAN DATA MINING UNTUK MENENTUKAN KELOMPOK PENGAYAAN MATERI MATA PELAJARAN UJIAN NASIONAL MENGGUNAKAN METODE *K-MEANS* (STUDI KASUS : MTs SUNAN KALIJAGA MOJO)

Oleh:

Listina

13.1.03.02.0333

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 TAHUN2017

Yang bertanda tangan dibawah ini:

Nama Lengkap	: Listina
NPM	: 13.1.03.02.0333
Telepun/HP	: 089642260267
Alamat Surel (Email)	: listina1011@gmail.com
Judul Artikel	:PENERAPAN DATA MINING UNTUK MENENTUKAN KELOMPOK PENGAYAAN MATERI MATA PELAJARAN UJIAN NASIONAL MENGGUNAKAN METODE *K-MEANS.* (STUDY KASUS : MTs SUNAN KALIJAGA MOJO)
Fakultas – Program Studi	:Fakultas Teknik – Teknik Informatika
NamaPerguruan Tinggi	: Universitas Nusantara PGRI Kediri
Alamat PerguruanTinggi	: Jl. KH. Achmad Dahlan 76 Kediri

Denganinimenyatakanbahwa:

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 bertanggung jawab dan diproses sesuai dengan ketentuan yang berlaku.

Mengetahui		Kediri, 9 Agustus 2017
Pembimbing I	Pembimbing II	Penulis,
Intan Nur farida, M.Kom. 0704108701	Danar Putra Pamungkas, M.Kom. 0708028704	Listina 13.1.03.02.0333

PENERAPAN DATA MINING UNTUK MENENTUKAN KELOMPOK PENGAYAAN MATERI MATA PELAJARAN UJIAN NASIONAL MENGGUNAKAN METODE *K-MEANS* (STUDI KASUS : MTs SUNAN KALIJAGA MOJO)

Listina
13.1.03.02.0333
FT – Teknik Informatika
Listina1011@gmail.com
Intan Nur Farida, M. Kom. dan Danar Putra Pamungkas, M.Kom.
UNIVERSITAS NUSANTARA PGRI KEDIRI

ABSTRAK

Ujian Nasional merupakan sistem evaluasi standar pendidikan secara nasional dan persamaan mutu tingkat pendidikan antar daerah yang dilakukan oleh Pusat Penilaian Pendidikan. Untuk itu dalam menghadapi Ujian Nasional para siswa harus mempersiapkannya dengan baik agar dapat memperoleh hasil yang memuaskan dan dapat lulus dengan nilai yang memuaskan juga. Maka sebelum menghadapi Ujian Nasional perlu diadakan pengayaan materi Ujian Nasional bagi siswa yang memiliki nilai rendah pada salah satu mata pelajaran Ujian Nasional.

K-Means clustering merupakan metode yang digunakan untuk mengelompokkan sebuah data ke dalam dua atau lebih variasi atau kelompok. Pada sistem ini menggunakan data nilai mata pelajaran Bahasa Indonesia, Bahasa Inggris, Matematika, MTK, IPA. Dalam penelitian ini *K-Means clustering* digunakan untuk menentukan kelompok pengayaan materi mata pelajaran ujian nasional.

Hasil penelitian menggunakan metode *K-Means* adalah menghasilkan sistem yang dapat membantu untuk menentukan kelompok pengayaan secara efektif. Selanjutnya pada sistem ini dapat mengklasterisasi kelompok pengayaan ke dalam 4 kelompok yang dapat digunakan sebagai pemantauan kemampuan siswa.

Kata Kunci : data mining, pengayaan, mata pelajaran, *K-Means*.

I. LATAR BELAKANG

Ujian Nasional (UN) merupakan sistem evaluasi standar pendidikan secara nasional dan persamaan mutu tingkat pendidikan antar daerah yang dilakukan oleh Pusat Penilaian Pendidikan, Depdiknas di Indonesia berdasarkan Undang-Undang Republik Indonesia Nomor 20 Tahun 2003. Ujian nasional diselenggarakan untuk mengukur pencapaian kompetensi lulusan peserta didik jenjang satuan pendidikan dasar dan pendidikan menengah sebagai hasil dari proses pembelajaran sesuai dengan Standar Kompetensi Lulusan (SKL). Untuk itu dalam menghadapi Ujian Nasional para siswa harus mempersiapkannya dengan baik agar dapat memperoleh hasil yang memuaskan dan dapat lulus dengan nilai yang memuaskan juga. Maka sebelum menghadapi Ujian Nasional perlu diadakan pengayaan materi Ujian Nasional bagi siswa yang memiliki nilai rendah pada salah satu mata pelajaran Ujian Nasional.

Pemberian pembelajaran pengayaan pada hakikatnya adalah pemberian bantuan bagi peserta didik yang memiliki kemampuan lebih, baik dalam kecepatan maupun kualitas belajarnya. Agar pemberian pengayaan tepat sasaran maka guru menempuh langkah-langkah sistematis, yaitu pertama mengidentifikasi kelebihan kemampuan peserta didik, dan kedua memberikan perlakuan pembelajaran pengayaan.Pada MTs Sunan Kalijaga Mojo dalam menghadapi Ujian Nasional para siswanya tidak di adakan pengayaan khusus untuk siswa yang memiliki nilai rendah pada mata pelajaran tertentu, tetapi pada MTs Sunan Kalijaga ini diadakan bimbel secara merata, sehingga kurang efektif rendah di salah sata mata pelajaran Ujian Nasional akan kesulitan mengikuti siswa yang sudah bisa.

Untuk menghadapi masalah tersebut maka perlu di adakannya sistem untuk mengelompokkan siswa yang memiliki nilai rendah pada salah satu mata pelajaran Ujian Nasional untuk diadakan pengayaan khusus materi Ujian Nasional yang nilainya kurang tersebut.

II. METODE

K-Means merupakan metode analisis kelompok yang mengarah pada pemartisian N objek pengamatan ke dalam K kelompok (*cluster*) di mana setiap objek pengamatan dimiliki oleh sebuah kelompok dengan mean (rata-rata) terdekat[8].

Jarakdaridatayangadamasing-masing*cluster*dihitungmenggunakan rumus *Euclideandistance*[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 15 data yang berisi empat *variable* untuk tiap datanya dan empat *cluster.*

Tabel 1 Dataset

No	Nama	Rata-Rata Semester 1-5			
		b.i	b.ing	mtk	ipa
1	Ah.Alfan Jauhari	83,4	79,4	80,4	79
2	Ahmad Maimun	84,2	78	79,8	78,8
3	Muh Misbahul	87	78	82	81
4	Ah. Nurul Ikhsan	83	76	79,4	78,8
5	Bagas Wahyu	81,8	77	77	78,8
6	Dzikrulloh Assifa	83,8	77,6	78,6	80
7	Dina Alfina	84,8	76,8	78,2	78,4
8	Galih AYU	80,8	79,4	79	79,8
9	Kuni Muslihah	84,2	78,6	76,6	81,4
10	Koni'atusSholihah	84,4	80,8	77,6	80,4
11	Nailatul Muna	84,2	78	76	77,4
12	Uswatun Hasanah	83,4	80,6	78,8	81,6
13	Ulfah Maulidah	84,8	81,8	80,2	80,6
14	Ummu Salma Nabila	81,2	80	77,6	78,4
15	Sri Mulyati	82,4	77,6	80,2	80,2

Menentukkan *centroid* secara *random.*

Tabel 2 *Centroid* Awal

Data No.	Nama	b.i	b.ing	mtk	ipa
5	Bagas Wahyu	81,8	77	77	78,8
8	Galih AYU	80,8	79,4	79	79,8
14	Ummu Salma Nabila	81,2	80	77,6	78,4
15	Sri Mulyati	82,4	77,6	80,2	80,2

Setiap data akan diukur jaraknya dengan tiap *centroid* awal menggunakan

rumus *Euclidean distance*kemudian diambil jarak terpendek.

Tabel 3 Iterasi 1

No	C1	C2	C3	C4	Jarak
1	3,73	1,73	3,05	3,21	1,73
2	2,12	0	3,89	4,26	0
3	5,24	4,18	7,13	4,76	4,18
4	2,34	2,36	4,19	6,38	2,34
5	3,26	3,82	3,42	6,74	3,26
6	2,08	1,78	3,52	4,64	1,78
7	0	2,12	5,03	5,81	0
8	5,03	3,89	0	4,88	0
9	3,89	4,16	4,52	4,91	3,89
10	4,52	3,90	4,15	2,82	2,82
11	2,76	4,04	5,31	6,53	2,76
12	5,19	4,02	3,38	2,52	2,52
13	5,81	4,26	4,88	0	0
14	4,85	4,24	2,10	5,27	2,10
15	3,69	2,34	2,72	4,85	2,34

Tabel 4 Kelompok *Cluster* iterasi 1

No.	C1	C2	C3	C4
1		1		
2		1		
3		1		
4	1			
5	1			
6		1		
7	1			
8			1	
9	1			
10				1
11	1			
12				1
13				1
14			1	
15		1		

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

Tabel 5 *Centroid* Baru

Centroid	Centroid ke-1	Centroid ke-1	Centroid ke-1	Centroid ke-1
C1	83,6	77,28	77,44	78,96
C2	84,16	78,12	80,2	79,8
C3	81	79,7	78,3	79,1
C4	84,2	81,06	78,86	80,86

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

Tabel 6 Iterasi 2

No	C1	C2	C3	C4	Jarak
1	4,38	2,19	3,94	2,46	2,19
2	2,53	0,90	3,92	3,70	0,90
3	6,48	4,01	7,80	5,32	4,01
4	2,42	2,73	4,35	5,62	2,42
5	1,88	4,25	3,11	5,48	1,88
6	1,60	1,73	3,62	3,60	1,60
7	1,60	2,84	4,83	5,00	1,60
8	3,93	3,79	1,05	3,93	1,05
9	2,96	3,96	4,43	3,39	2,96
10	3,88	3,78	3,86	1,39	1,39
11	2,32	4,83	4,61	5,44	2,32
12	4,45	3,45	3,61	1,18	1,18
13	5,67	3,81	4,97	80,61	3,81
14	3,67	4,58	1,05	4,22	1,05
15	3,27	1,87	3,34	4,18	1,87

Tabel 7 Kelompok *Cluster* Iterasi 2

No.	C1	C2	C3	C4
1		1		
2		1		
3		1		
4	1			
5	1			
6	1			
7	1			
8			1	
9	1			
10				1
11	1			
12				1
13		1		
14			1	
15		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.

Tabel 8 *Centroid* Baru

Centroid	Centroid ke-1	Centroid ke-1	Centroid ke-1	Centroid ke-1
C1	83,63	77,33	77,63	79,13
C2	84,36	78,96	80,52	80,64
C3	81	79,7	78,3	79,1
C4	83,9	80,7	78,2	81

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

Tabel 9 Iterasi 3

No	C1	C2	C3	C4	Jarak
1	4,13	1,30	3,94	2,63	1,37
2	2,34	2,03	3,92	3,73	2,03
3	6,25	3,45	7,80	5,68	3,45
4	2,32	3,90	4,35	5,40	2,32
5	1,99	5,11	3,11	4,93	1,99
6	1,33	2,50	3,62	3,28	1,33
7	1,58	3,90	4,83	4,77	1,58
8	3,82	3,98	1,05	3,65	1,05
9	2,85	4,01	4,43	2,68	2,68
10	3,76	3,45	3,86	0,98	0,98
11	2,53	5,64	4,61	5,01	2,53
12	4,26	2,73	3,61	0,98	0,98
13	5,48	2,89	4,97	2,48	2,48
14	3,68	4,96	1,05	3,86	1,05
15	3,05	2,44	3,34	4,06	2,44

Tabel 10 Kelompok *Cluster* Iterasi 3

No.	C1	C2	C3	C4
1		1		
2		1		
3		1		
4	1			
5	1			
6	1			
7	1			
8				
9				1
10				1
11	1			
12				1
13				1
14			1	
15		1		

Kemudian lakukan perbandingan lagi 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, kemudian ulangi langkah kedua hingga data tidak berubah atau sama.

III. HASIL DAN KESIMPULAN

A. Hasil

Berikut adalah hasil pengujian dari 6 skenario yang dilakukan.

No	data	*centroid*	Iterasi	Cl1	Cl2	Cl3	Cl4
1	60	4	10	18	7	11	24
2	60	3	9	14	21	25	0
3	120	4	12	34	27	19	40
4	120	3	8	17	49	54	0
5	150	4	12	45	32	24	49
6	150	3	16	27	67	56	0

Tabel 11 Hasil Pengujian

Skenario pengujian 1-6 dilakukan untuk mengetahui pengaruh jumlah data dan maksimum iterasi terhadap hasil klusterisasi. Hasil pengujian yang diperoleh, bahwa jumlah data tidak berpengaruh terhadap jumlah iterasi yang dihasilkan. Namun berpengaruh terhadap waktu yang dibutuhkan dan

keakuratan hasil klasterisasi. Semakin banyak iterasi maka akan semakin lama waktu yang dibutuhkan pula. Selain itu pengambilan *centroid* secara *random* sangat berpengaruh terhadap hasil klasterisasi dan hasil menjadi kurang akurat. Agar hasil bisa akurat maka *centroid* awal harus ditentukan.

Berikut ini merupakan presentasi hasil pengujian skenario 1-6 :

No	Data	Cl 1	Cl 2	Cl 3	Cl 4
1	60	0,3%	0,11%	0,18%	0,4%
2	60	0,23%	0,35%	0,41%	0
3	120	0,28%	0,22%	0,15%	0,33%
4	120	0,14%	0,40%	0,45%	0
5	150	0,3%	0,21%	0,16%	0,32%
6	150	0,18%	0,44%	0,37%	0

Tabel 12 Presentase Pengujian

Dari 1-6 skenario pengujian yang sudah dilakukan hasil terbaik dari skenario tersebut didapatkan pada skenario pertama dengan presentase 0,3% untuk cluster

pertama, 0,11% untuk cluster kedua, 0,18% untuk cluster ketiga dan 0,4% untuk cluster keempat.

A. KESIMPULAN

1. Telah dihasilkan sistem untuk menentukan kelompok pengayaan yang akan membantu guru dalam mengelompokkan suatu kelompok pengayaan untuk siswa agar dapat mendalami mata pelajaran tersebut secara efektif dan efisien.

1. Telah dihasilkan sistem untuk menentukan kelompok pengayaan menggunakan metode K-Means. Pada sistem ini terdapat empat *cluster* yang menjadi tujuan pengelompokkan. Dari hasil pengujian *clustering* yang dilakukan kurang akurat. Karena, pengambilan *centroid* awalnya dilakukan secara *random.* Pada scenario pengujian yang sudah dilakukan hasil terbaik dari skenario tersebut didapatkan pada skenario pertama dengan presentase 0,3% untuk cluster pertama, 0,33% untuk cluster kedua, 0,183% untuk cluster ketiga dan 0,4% untuk cluster keempat.

IV. DAFTAR PUSTAKA

[1] Agustin, Fenty. E.M., Ftria, Ardini., Anif, H.S. 2015. Implementasi Algoritma *K-Means* untuk Menentukan Kelompok

Pengayaan Materi Mata Pelajaran Ujian Nasional. *Jurnal Teknik Informatika*, Vol.8 No.(1) April 2015,tersedia:www.journal.uinjkt. ac.id, diunduh 12 Desember 2016.

[2]Agusta,Y.2007.K-Means-Penerapan,PermasalahandanMetodeTerkait.*JurnalSistem dan Informatika*, 3: 47-60.

[3]Kusrini., Luthfi,T.E., 2009. *Algoritma Data Mining*.Yogyakarta: Penerbit Andi.

[4] Prasetyo, E.2012.*Data Mining Konsep* dan *Aplikasi* Menggynakan *Matlab*. Yogyakarta:Andi offset.

[5] Mundzir, M.F. 2014. *PHP Tutorial Book For Beginner* (B. Waskitho, Ed). Yogyakarta: Notebook.

[6] Ramadhani, R.D. 2014. Data Mining Menggunakan Algoritma *K-MeansClustering* Untuk Menetukan Strategi Promosi Universitas Dian Nuswantoro. *Jurusan Sistem Informasi, Fakultas Ilmu Komputer, Universitas Dian Nuswantoro*, tersedia: http://eprints.dinus.ac.id, diunduh 21 September 2016.

[7] Rohmawati, W.N., Defiyanti, S., Jajuli, M. 2015. Implementasi Algoritma *K-Means* dalam Pengklasteran Mahasiswa Pelamar Beasiswa. *Jurnal Ilmiah Teknologi Informasi Terapan*, Vol.1 No.(2), 30 April 2015, tersedia: http://jitter.widyatama.ac.id, diunduh 20 November 2016.

[8] Usman, M., Uzer. 1993. *Upaya Optimalisasi Kegiatan Belajar Mengajar*. Bandung : Remaja Rosdakarya.