

PENGELOMPOKAN SEKOLAH DASAR DI KOTA SAMARINDA BERDASARKAN INDIKATOR STANDAR PELAYANAN MINIMAL PENDIDIKAN MENGGUNAKAN METODE *K-MEANS*

Tien Rahayu Tulili¹⁾, Karyo Budi Utomo²⁾, Dwi Lestari³⁾

¹⁾Program Studi Teknik Informatika Multimedia, Politeknik Negeri Samarinda

^{2,3)}Program Studi Teknik Informatika, Politeknik Negeri Samarinda

Email: ¹⁾tien.tulili@polnes.ac.id; ²⁾karyobudi@gmail.com; ³⁾dwi.lestari@gmail.com

Abstrak - Kualitas pelayanan pada sekolah menentukan tingkat keberhasilan pelayanan suatu pendidikan. Berdasarkan data dari dinas pendidikan tahun 2015, jumlah sekolah dasar di Kalimantan Timur khususnya kota Samarinda adalah 222 sekolah dasar. Standar pelayanan minimal pendidikan merupakan suatu acuan untuk penilaian kinerja pelayanan pendidikan. Terdapat enam kategori dalam SPM yaitu sarana dan prasarana, pendidik dan tenaga kependidikan, kurikulum, penjaminan mutu pendidikan, penilaian pendidikan, dan manajemen sekolah. Masing-masing kategori memiliki indikator pencapaian untuk dijadikan acuan menilai capaian SPM yang telah dilakukan oleh tiap sekolah. Jumlah data yang banyak dapat memperlambat proses untuk mengetahui kelompok sekolah dasar dengan tingkat capaian standar pelayanan minimal yang baik. Oleh karena itu, untuk mempercepat proses tersebut maka dapat dilakukan sebuah proses dalam pengelompokan data. Penelitian ini bertujuan untuk mengelompokkan capaian terhadap standar pelayanan pada Sekolah Dasar sehingga dapat diketahui sekolah dengan tingkat capaian yang kurang, sedang, dan baik. Pengelompokan terhadap capaian tersebut dilakukan dengan menggunakan metode *K-Means* yang merupakan algoritma *Clustering* non hirarki. Berdasarkan hasil penelitian yang telah dilakukan terhadap tingkat capaian standar pelayanan pada sekolah dasar, terdapat tiga *cluster* yaitu terdapat 194 sekolah yang masuk dalam kategori Baik, 17 sekolah yang masuk dalam kategori Sedang, dan 11 sekolah yang masuk dalam kategori kurang dengan nilai squared error sebesar 32.2505.

Kata kunci: Standar pelayanan minimal; *Clustering*; *K-Means*;

1. Pendahuluan

Menurut Undang-Undang Dasar 1945 Pasal 31 ayat (1), *setiap warga negara berhak mendapat pendidikan*. Undang-Undang Dasar Pasal 31 ayat (2)[9], *setiap warga negara wajib mengikuti pendidikan dasar dan pemerintah wajib membiayainya*. Untuk meningkatkan kualitas pendidikan yang memadai, maka diperlukan sebuah Standar Pelayanan Minimal (SPM). Menurut Peraturan Menteri Pendidikan Dan Kebudayaan Republik Indonesia Nomor 23 Tahun 2013 Pasal 6 ayat (1), SPM pendidikan merupakan acuan dalam perencanaan program dan penganggaran pencapaian target masing-masing daerah kabupaten/kota. Oleh karena itu, penting bagi setiap daerah kabupaten/kota untuk mencapai target SPM dengan meningkatkan kualitas pelayanannya.

Berdasarkan data dari dinas pendidikan tahun 2015, saat ini jumlah sekolah dasar di Kalimantan Timur khususnya kota Samarinda adalah 222 sekolah dasar. Pemerintah mencanangkan bahwa setiap sekolah dasar harus memenuhi SPM. Terdapat enam kategori jenis pelayanan yang terdiri atas: Sarana dan Prasarana, Pendidik dan Tenaga Kependidikan, Kurikulum, Penjaminan Mutu Pendidikan, Penilaian Pendidikan, dan

Manajemen Sekolah. Masing-masing kategori memiliki indikator pencapaian untuk dijadikan acuan menilai capaian SPM yang telah dilakukan oleh tiap sekolah. Jumlah data yang banyak dapat memperlambat proses untuk mengetahui kelompok sekolah dasar dengan tingkat capaian standar pelayanan minimal yang baik. Oleh karena itu, untuk mempercepat proses tersebut maka dapat dilakukan sebuah proses dalam pengelompokan data.

Salah satu metode pengelompokan yang dapat digunakan yaitu metode *K-Means Clustering*. *K-Means Clustering* merupakan salah satu metode berbasis partisi yang dapat mengelompokkan data berdasarkan titik pusat klaster (*centroid*) terdekat. Oleh karena itu, metode *K-Means Clustering* diterapkan dalam penelitian ini. Dengan adanya penerapan metode *K-Means Clustering* ini, diharapkan dapat membantu untuk mengetahui kelompok sekolah dasar dengan tingkat capaian standar pelayanan minimal yang baik, sedang dan kurang.

2. Tinjauan Pustaka

Beberapa penelitian terdahulu yang mendasari penelitian ini antara lain yang dilakukan oleh Agustin dkk[1], (2015) yaitu penelitian tentang

Implementasi Algoritma *K-Means* Untuk Menentukan Kelompok Pengayaan Materi Mata Pelajaran Ujian Nasional. Kesimpulan yang diperoleh adalah penelitian ini menggunakan 4 *cluster* sesuai dengan jumlah mata pelajaran Ujian Nasional dan 12 sampel data dari siswa yang memiliki nilai terendah. Hasil pengelompokan pada aplikasi ini selain dimanfaatkan untuk menentukan kelompok pengayaan materi, dapat juga digunakan untuk mengukur progres siswa terhadap materi yang sudah dikuasai.

Penelitian yang dilakukan oleh Darmi dan Setiawan[2] adalah Penerapan Metode *Clustering K-Means* Dalam Pengelompokan Penjualan Produk. Pada penelitian ini dihasilkan bahwa algoritma *K-Means* tidak terpengaruh terhadap urutan objek yang digunakan. Jumlah keanggotaan *cluster* yang dihasilkan berjumlah sama ketika peneliti menggunakan objek yang lain sebagai titik awal pusat *cluster* tersebut.

Selanjutnya, Suprawoto[3] dalam penelitiannya menyatakan bahwa kelompok mahasiswa dengan nilai rerata UN yang rendah memiliki pengaruh terhadap prestasi akademik mahasiswa yang rendah pada jenjang diploma-3 (D-3) dan strata-1 (S-1). Prestasi akademik mahasiswa juga tinggi pada semua jenjang jika mahasiswa memiliki nilai UN yang tinggi. Berdasarkan hasil pengelompokan daerah asal sekolah, IPK rata-rata yang tertinggi berasal dari propinsi Daerah Istimewa Yogyakarta (DIY) dan Jawa Tengah (Jateng).

Pada penelitian yang dilakukan oleh Irtawaty[4], 10 sampel data pasien yang memiliki 3 parameter uji yaitu *ureum*, *kritinien* dan *GFR* digunakan untuk dikelompokkan ke dalam 5 klaster yang terdiri dari kondisi ginjal normal, gejala ginjal stadium 1, stadium 2, stadium 3 dan stadium 4. Pada penelitian ini tingkat akurasi pengelompokan penyakit ginjal menggunakan metode *K-Means* cukup tinggi, yaitu sekitar 90%.

Penelitian yang dilakukan oleh Talakua dkk[5], diperoleh bahwa angka Indeks Pembangunan Manusia, Angka Harapan Hidup, Angka Melek Huruf, Angka Rata-rata Lama Sekolah dan angka Pengeluaran Per Kapita terdapat perbedaan yang signifikan pada tahun 2014.

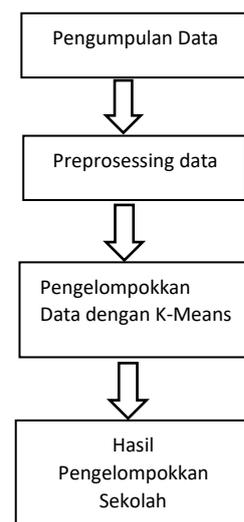
Penelitian yang dilakukan oleh Nasari dan Sianturi[6] adalah Penerapan Algoritma *K-Means Clustering* Untuk Pengelompokan Penyebaran Diare Di Kabupaten Langkat. Kesimpulan yang diperoleh adalah penelitian ini menggunakan data pada kecamatan-kecamatan yang berada di kabupaten langkat. Penelitian ini menghasilkan dua *cluster*, yaitu pusat *cluster* yang diperoleh pada *cluster* pertama berada pada kecamatan Batang serangan, Brandan Barat dan Pematang Jaya dan pusat *cluster* kedua berada pada kecamatan Hinai, Sei Bingai dan Sirapit. Pusat *cluster* pertama merupakan daerah penyebaran diare untuk

jumlah penderita tingkat menengah atau bukan merupakan pusat penyebaran diare, sedangkan pusat *cluster* kedua merupakan daerah-daerah pusat penyebaran diare, untuk itu pada daerah-daerah pusat *cluster* kedua harus menjadi daerah perhatian pemerintah untuk penanganan diare.

Selanjutnya, penelitian yang dilakukan oleh Nur[7] adalah Penerapan Algoritma *K-Means* Pada Siswa Baru Sekolah Menengah Kejuruan Untuk *Clustering* Jurusan. Kesimpulan yang diperoleh adalah penelitian ini menggunakan data siswa baru sekolah menengah kejuruan tahun ajaran 2014/2015. Dalam penelitian ini, pengelompokan disesuaikan kriteria-kriteria siswa seperti penghasilan orang tua, tanggungan anak orang tua dan nilai tes siswa. Kriteria-kriteria tersebut digunakan agar pengelompokan yang dihasilkan menjadi lebih optimal. Hasil dari pengelompokan tersebut diperoleh tiga kelompok yaitu kelompok tidak lulus, kelompok rekayasa perangkat lunak dan kelompok teknik komputer jaringan.

Penelitian yang kedelapan dilakukan oleh Putri[8] adalah Metode *Non Hierarchy* Algoritma *K-Means* Dalam Mengelompokkan Tingkat Kelarisan Barang (Studi Kasus : Koperasi Keluarga Besar Semen Padang). Kesimpulan yang diperoleh adalah penelitian ini menggunakan sampel data sebanyak 130 data, sedangkan atribut yang digunakan sebanyak 4 yaitu Kode Barang, Nama Barang, Total Stok, dan Stok Akhir. Jumlah *cluster* yang digunakan sebanyak 2 *cluster*, yaitu barang paling laris dan barang kurang laris. Berdasarkan hasil analisis menggunakan *cluster k-means*, perlengkapan mandi dan mencuci dapat dikelompokkan menjadi dua *cluster*, yaitu C1 adalah kelompok barang yang laris, dengan total sampel data 95 item dari 130 sampel data. Sedangkan C2 adalah kelompok barang yang kurang laris, dengan total sampel data 35 item dari 130 sampel data. Hasil yang dicari dengan cara manual ekuivalen dengan hasil yang diproses dengan menggunakan aplikasi *Tanagra* 1.4.50.

3. Metodologi Penelitian



Gambar 1. Metodologi penelitian

Pada penelitian ini, data yang digunakan adalah data standar pelayanan minimal (SPM) Pendidikan tahun 2015 yang bersumber dari Dinas Pendidikan dan Kebudayaan Kota Samarinda. Data yang digunakan berjumlah 222 sekolah meliputi daftar Sekolah Dasar dan hasil capaian dari masing-masing Indikator Pencapaian(IP). Berdasarkan data yang diperoleh terdapat ± 24 IP. Tiap IP akan dikelompokkan ke dalam jenis pelayanan. Pada SPM terdapat tujuh jenis pelayanan, antara lain: Sarana dan Prasarana, Pendidik dan Tenaga Kependidikan, Kurikulum, Penjaminan Mutu Pendidikan, Penilaian Pendidikan, Manajemen Sekolah. Berdasarkan kebijakan Dinas Pendidikan dan Kebudayaan Kota Samarinda, bobot kriteria setiap kategori dapat dilihat pada tabel 1.

Tabel 1. Tabel Nilai Bobot Jenis Pelayanan

Jenis Pelayanan	Bobot
Sarana dan Prasarana	25%
Pendidik dan Tenaga Kependidikan	30%
Kurikulum	10%
Penjaminan Mutu Pendidikan	15%
Penilaian Pendidikan	5%
Manajemen Sekolah	15%

Sumber data: Dinas Pendidikan Kota Samarinda

Pada preprocessing data, pengelompokkan, terhadap tiap IP berdasarkan jenis pelayanannya. Didapatkan bahwa terdapat enam IP pada jenis pelayanan Sarana dan Prasarana, tujuh IP pada jenis pelayanan Pendidik dan Tenaga Kependidikan, dua IP pada jenis pelayanan Kurikulum, empat IP pada jenis pelayanan Penjaminan Mutu Pendidikan, satu IP pada jenis pelayanan Penilaian Pendidikan, dan empat pada jenis pelayanan Manajemen Sekolah. Nilai pada tiap IP pada masing-masing jenis pelayanan akan diakumulasikan untuk dicari nilai rata-rata dari tiap jenis pelayanan. Nilai dari tiap jenis pelayanan ini akan digunakan pada proses awal perhitungan K-Means.

Selanjutnya nilai rata-rata tiap IP akan dikalikan dengan nilai bobot masing-masing jenis pelayanan ada pada tabel 1. Hasil perkalian dari masing-masing jenis pelayanan ini selanjutnya akan dihitung menggunakan K-Means untuk menentukan kelompok dari masing-masing sekolah.

4. Hasil dan Pembahasan

Pada penelitian ini, perhitungan K-Means menggunakan perangkat lunak Wekadata. Formula jarak yang digunakan pada perhitungan ini yaitu formula metode Euclidean dengan jumlah kluster sebanyak tiga dan menggunakan jumlah seeds yang berbeda-beda.

Pada tabel 2, percobaan pertama yaitu dengan menggunakan jumlah seed 10 dan nilai centroid awal diambil secara acak. Tabel 2 ini merupakan hasil klusterisasi terakhir dimana jumlah iterasi yang dilakukan sebanyak delapan kali dengan nilai squared error sebesar 37.6023.

Tabel 2. Nilai kluster iterasi terakhir dengan nilai seeds sebesar 10 dan nilai centroid awal acak

Atribut	Full Data	Cluster #0	Cluster #1	Cluster #2
Sarana dan Prasarana	13.20370	11.32020	13.65430	12.63410
Pendidik dan Tenaga	22.4637	16.1472	25.5978	15.6517
Kurikulum	9.1834	2.9524	9.8719	9.7212
Penjaminan	11.0945	4.6291	11.8633	11.4882
Penilaian	4.3147	0.0183	4.9194	4.2929
Manajemen	13.3108	6.9048	14.4702	12.5

Pada tabel 3, jumlah seed yang digunakan adalah sepuluh dan nilai centroid awal diambil dengan menggunakan metode K-Means++. Hasil klusterisasi terakhir pada tabel 2 menunjukkan bahwa nilai tiap atribut pada cluster #2 memiliki nilai-nilai yang lebih besar dari nilai pada cluster #0 maupun cluster #1. Nilai squared error yang dimiliki sebesar 37.5972.

Tabel 3 Nilai klusterisasi akhir dengan jumlah seeds 10 dan nilai centroid awal menggunakan K-Means++

Atribut	Full Data	Cluster #0	Cluster #1	Cluster #2
Sarana dan Prasarana	13.20370	11.32020	12.50710	13.71230
Pendidik dan Tenaga	22.4637	16.1472	15.8039	25.6781
Kurikulum	9.1834	2.9524	9.732	9.8702
Penjaminan	11.0945	4.6291	11.4069	11.8967
Penilaian	4.3147	0.0183	4.3201	4.9183
Manajemen	13.3108	6.9048	12.5962	14.4631

Pada tabel 4, jumlah seeds yang digunakan adalah dua puluh dengan nilai centroid awal menggunakan metode K-Means++. Dapat dilihat pada tabel 3 bahwa nilai tiap atribut pada cluster #0 lebih besar dari pada dua kluster lainnya. Jumlah iterasi yang terjadi pada percobaan ini adalah delapan iterasi. Nilai squared errors yang dimiliki sebesar 38.1699.

Tabel 4 Nilai klusterisasi akhir dengan jumlah seeds 20 dan nilai centroid awal menggunakan K-Means++

Atribut	Full Data	Cluster #0	Cluster #1	Cluster #2
Sarana dan Prasarana	13.20370	17.68000	10.42160	11.44180
Pendidik dan Tenaga	22.4637	23.5674	22.8724	16.1581
Kurikulum	9.1834	9.7174	9.9144	3.2727
Penjaminan	11.0945	11.9538	11.6149	5.1005
Penilaian	4.3147	4.7069	4.8433	0.0175

Tulili dkk., Pengelompokan Sekolah Dasar di Kota Samarinda Berdasarkan Indikator Standar Pelayanan Minimal Pendidikan Menggunakan Metode *K-Means*

Manajemen	13.3108	14.4512	13.6865	7.0455
-----------	---------	---------	---------	--------

Pada tabel 5, jumlah seeds yang digunakan adalah tujuh dengan nilai centroid awal menggunakan metode K-Means++. Jumlah iterasi yang dilakukan sebesar empat iterasi dan memiliki nilai error sebesar 32.2505

Tabel 5 Nilai klasterisasi akhir dengan jumlah seeds 7 dan nilai centroid awal menggunakan K-Means++

Atribut	Full Data	Cluster #0	Cluster #1	Cluster #2
Sarana dan Prasarana	13.20370	13.36520	13.93060	9.23180
Pendidik dan Tenaga	22.4637	23.2188	16.5616	18.268
Kurikulum	9.1834	9.8445	7.2871	0.4545
Penjaminan	11.0945	11.7418	10.4452	0.6818
Penilaian	4.3147	4.9024	0.3994	0
Manajemen	13.3108	13.9691	14.1177	0.4545

Pada tabel 6, jumlah seeds yang digunakan yaitu sebesar lima belas dengan nilai centroid awal didapatkan dengan menggunakan metode K-Means++. Jumlah iterasi yang dilalui sebanyak enam belas iterasi dengan nilai error sebesar 37.9779

Pada tabel 7 dapat dilihat hasil klasterisasi K-Means. Pada tabel, Hasil 1-5 menggunakan jumlah seed secara berturut-turut yaitu sebesar 10,10,20,7, dan 15. Hanya pada Hasil 1, pemilihan centroid awal dilakukan secara acak sedangkan empat hasil yang lain menggunakan K-Means++.

Tabel 7. Hasil Klasterisasi dengan metode K-Means

	Hasil 1	Hasil 2	Hasil 3	Hasil 4	Hasil 5
Kategori Baik	151	149	82	194	113
Kategori Sedang	50	52	118	17	87
Kategori Kurang	21	21	22	11	22
Sum Squared Errors selama Klasterisasi	37.6023	37.5972	38.1699	32.2505	37.9779
Jumlah Iterasi	8	8	8	4	16
Jumlah Seeds	10	10	20	7	15

Dapat dilihat dari tabel 7 bahwa, penentuan nilai centroid awal tidak berpengaruh besar terhadap hasil akhir klasterisasi. Hal ini dapat dilihat pada Hasil 1 dan Hasil 2, dengan jumlah seeds yang sama, jumlah instans yang berhasil dikelompokkan ke dalam kelas Baik, Sedang, dan Kurang tidak menunjukkan nilai yang signifikan, yaitu jumlah instans pada kategori

kurang sama-sama bernilai 21 instans; sedangkan pada dua kategori lainnya selisih nilai hanya terpaut dua instans. Selanjutnya bahwa, jumlah seeds mempengaruhi jumlah squared error. Semakin besar jumlah seeds yang diberikan, semakin besar jumlah squared error yang terjadi. Tetapi jumlah seeds ini tidak berpengaruh pada jumlah iterasi yang terjadi. Hasil 4 menunjukkan nilai error yang paling kecil dan jumlah iterasi yang paling kecil yaitu masing-masing bernilai 32.2505 dan 4.

Dari lima percobaan yang dilakukan, Dapat dikatakan bahwa berdasarkan percobaan ini, maka dapat dikatakan bahwa terdapat 194 sekolah yang masuk dalam kategori Baik, 17 sekolah yang masuk dalam kategori Sedang, dan 11 sekolah yang masuk dalam kategori kurang.

5. Simpulan

Pengelompokan 222 sekolah dasar di kota Samarinda menggunakan metode K-Means dengan Euclidean Distance dapat dilakukan dengan hasil dimana 194 sekolah masuk dalam kategori baik, 17 sekolah masuk dalam kategori sedang, dan 11 sekolah masuk dalam kategori kurang. Nilai squared error yang paling minimum yaitu sebesar 32.2505. Selanjutnya, bahwa jumlah seeds yang digunakan berpengaruh pada besarnya nilai squared error yang didapatkan selama proses klasterisasi.

Untuk penelitian selanjutnya, klasterisasi dapat dilakukan dengan menggunakan jumlah klaster yang beragam dan menggunakan metode jarak yang berbeda untuk melihat seberapa besar pengaruh jumlah klister terhadap nilai squared error yang terjadi selama klasterisasi. Demikian halnya juga untuk penelitian dimasa mendatang, metode klasterisasi yang berbebeda dapat digunakan sebagai perbandingan terhadap hasil klasterisasi ini.

Referensi

- [1] Agustin, F. E., Fitria, A., & S, A. H. (2015). Implementasi Algoritma K-Means Untuk Menentukan Kelompok Pengayaan Materi Mata Pelajaran Ujian Nasional. *JURNAL TEKNIK INFORMATIKA VOL. 8, NO. 1*, 73-78.
- [2] Darmi, Y., & Setiawan, A. (2016). Penerapan Metode Clustering K-Means Dalam Pengelompokan Penjualan Produk. *Jurnal Media Infotama Vol. 12, No.2*, 148-157
- [3] Suprawoto, T. (2016). Klasifikasi Data Mahasiswa Menggunakan Metode K-Means Untuk Menunjang Pemilihan Strategi Pemasaran. *Jurnal Informatika dan Komputer (JIKO) Vol. 1, No. 1*, 12-18.
- [4] Irtawaty, A. S. (2017). Klasifikasi Penyakit Ginjal dengan Metode K-Means. *JURNAL TEKNOLOGI TERPADU Vol. 5, No. 1, April 2017*, 49-53.
- [5] Talakua, M. W., Leleury, Z. A., & Talluta, A. W. (2017). Analisis Cluster Dengan Menggunakan Metode K-Means Untuk Pengelompokkan Kabupaten/Kota Di Provinsi Maluku Berdasarkan

- Indikator Indeks Pembangunan Manusia Tahun 2014. *Jurnal Ilmu Matematika dan Terapan Vol. 11, No. 2*, 119-128
- [6] Nasari, F., & Sianturi, C. J. (2016). Penerapan Algoritma K-Means Clustering Untuk Pengelompokan Penyebaran Diare Di Kabupaten Langkat. *Cogito Smart Journal Vol. 2, No. 2, Desember 2016*, 108-119.
- [7] Nur, F., Zarlis, P., & Nasution, D. B. (2017). Penerapan Algoritma K-Means Pada Siswa Baru Sekolah Menengah Kejuruan Untuk Clustering Jurusan. *Jurnal Nasional Informatika dan Teknologi Jaringan Vol 1, No.2, Maret 2017*, 100-105.
- [8] Putri, D. E. (2015). Metode Non Hierarchy Algoritma K-Means Dalam Mengelompokkan Tingkat Kelarisan Barang (Studi Kasus : Koperasi Keluarga Besar Semen Padang). *Seminar Ilmiah Nasional Teknologi Komputer (SENATKOM 2015) Vol.1, Oktober 2015*, 36-41.
- [9] Peraturan Menteri Pendidikan dan Kebudayaan Nomor 23 Tahun 2013 Tentang Perubahan Atas Peraturan Menteri Pendidikan Nasional Nomor 15 Tahun 2010 Standar Pelayanan Minimal Pendidikan Dasar di Kabupaten/Kota.
- [10] Kamus Besar Bahasa Indonesia. (2018, Juli 19). Diambil kembali dari <https://kbbi.kemdikbud.go.id/entri/kelompok>
- [11] Asroni, & Adrian, 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 Teknik Vol. 18, No. 1*, 76-82.
- [12] Suyanto. (2017). *Data Mining Untuk Klasifikasi Dan Klasterisasi Data*. Bandung: Informatika Bandung.