

Penggunaan Metoda K-Means Clustering pada Pemetaan Kualitas Pendidikan SMTA di Indonesia Berdasarkan Nilai UMPTN

Siti Aminah¹, Toemin A. Masoem²

¹Departemen Matematika FMIPA-UI, Kampus UI Depok 16424

²Fakultas Ilmu Komputer, Kampus UI Depok 16424
aminah@makara.eso.ui.ac.id

Abstrak

Makalah ini membahas tentang pengelompokan data melalui metoda K-Means Clustering. Hasil pengelompokan metoda tersebut digunakan untuk memetakan data rata-rata nilai UMPTN IPA per kabupaten/kota di seluruh Indonesia dengan menggunakan Arc-View GIS versi 3.1. Dari hasil pemetaan tersebut diharapkan data yang merupakan rataan nilai UMPTN IPA tersebut akan lebih mudah dan menarik untuk dibaca. Terutama bagi pengambil kebijakan dalam dunia pendidikan. Dengan melihat hasil pemetaan tersebut, diharapkan mereka bisa meningkatkan kualitas pendidikan atau melakukan perbaikan-perbaikan dalam dunia pendidikan di Indonesia secara global, sesuai dengan kualitas pendidikannya.

Keywords: Data clustering, K-Means Clustering Method, Arc-View GIS Version 3.1.

1. PENDAHULUAN

Data umumnya berbentuk sekumpulan angka-angka maupun karakter dengan jumlah yang tidak sedikit. Bagi beberapa orang, mengamati ribuan angka bukanlah suatu hal yang menyenangkan. Oleh karena itu ada berbagai cara untuk menyajikan data agar pembaca lebih mudah untuk membacanya, misalnya secara grafik.

Diagram garis, diagram batang, diagram lingkaran dan diagram titik adalah beberapa contoh bentuk penyajian data secara grafik. Seiring dengan pesatnya perkembangan grafika komputer dewasa ini, bentuk penyajian grafik untuk mempresentasikan data dengan menggunakan komputer semakin populer digunakan. Beberapa paket perangkat lunak Matematika dan Statistika menampilkan grafik untuk mempresentasikan data, seperti SPSS, Matematika dan lain sebagainya. Dengan paket-paket tersebut kita dapat menampilkan data dalam bentuk grafik dan menganalisa data berdasarkan grafik yang mempresentasikan data tersebut.

Seperti kita ketahui, kualitas pendidikan adalah sangat penting untuk kemajuan suatu bangsa. Sesuai dengan bunyi Alinea Keempat Pembukaan Undang-Undang Dasar 1945, yaitu Pemerintah Negara

Indonesia, antara lain, berkewajiban mencerdaskan kehidupan bangsa, maka pasal 31 ayat 1 UUD 1945 menetapkan bahwa: "Setiap warganegara berhak mendapat pendidikan". Untuk maksud itu, Undang-Undang Dasar 1945 mewajibkan Pemerintah mengusahakan dan menyelenggarakan satu sistem pendidikan nasional yang diatur dengan undang-undang (Pasal 31 ayat 3). Penyelenggaraan pendidikan dilaksanakan melalui dua jalur, yaitu jalur pendidikan sekolah dan jalur pendidikan luar sekolah. Yang akan dibahas dalam makalah ini adalah kualitas pendidikan sekolah pada jenjang SMTA di seluruh Indonesia. Penulis memilih jenjang pendidikan SMTA karena kualitas pendidikan SMTA lebih dapat menggambarkan kualitas pendidikan suatu daerah dibandingkan dengan jenjang pendidikan lain.

Toemin A. Masoem dalam bukunya yang berjudul "Mana Yang Lebih Dapat Diandalkan, Ebtanas atau UMPTN?", diantaranya membahas secara terinci hubungan antara nilai UMPTN dan nilai Ebtanas, hubungan antara nilai Ebtanas, nilai UMPTN dan IPK. Dari pembahasan tersebut disimpulkan bahwa Ebtanas tidak baku secara nasional dan UMPTN cukup baku dan dapat diandalkan sebagai alat ukur pendidikan di tingkat SMTA [1].

2. METODE PENELITIAN

Data yang dipergunakan dalam penelitian ini adalah rata-rata nilai UMPTN per kabupaten/kodya tahun 1999 yang merupakan hasil penelitian yang dilakukan oleh Prof. DR. Toemin A. Masoem [1].

Data yang merupakan rata-rata nilai per kabupaten/kodya di seluruh Indonesia tersebut, dikelompokkan menjadi 5 cluster dengan kriteria nilai sangat rendah, rendah, sedang, tinggi dan sangat tinggi. menggunakan metoda *K-Means Clustering* [2, 3].

Data rata-rata nilai UMPTN yang sudah dikelompokkan tersebut dipetakan pada peta Indonesia dengan menggunakan Arc-View GIS versi 3.1. Rataan nilai dinyatakan dengan warna [4, 5]. Makin tebal warnanya berarti makin tinggi rata-rata nilainya. Sebaliknya, makin tipis warnanya berarti makin rendah rata-rata nilainya. Untuk pemrograman, digunakan *Microsoft Visual Basic versi 6.0* [6].

3. TINJAUAN PUSTAKA

Metoda K-Means Clustering

Dalam mengelompokkan data melalui metoda *K-means clustering*, harus ditentukan terlebih dahulu partisi awal yaitu dengan memilih jumlah cluster yang dikehendaki dan pusat awal dari setiap cluster. Penentuan vektor pusat awal cluster dapat dilakukan dengan dua cara.

Cara yang pertama adalah secara sembarang. Bila jumlah cluster yang diambil adalah K , maka K vektor pertama dari data dapat diambil sebagai pusat awal dari cluster.

Cara yang kedua adalah dengan menggunakan rumus yang dapat mempersingkat proses iterasi yang diusulkan oleh K. S. Fu [2, 3]. Rumus penentuan harga awal tersebut adalah sebagai berikut:

$$m_{ik} = \mu_i + \sigma_i \cdot (2 \cdot (k - 1) / (K - 1) - 1)$$

dimana :

m_{ik} = rata-rata cluster untuk ciri ke- i dan cluster ke- k

μ_i = rata-rata ciri ke- i

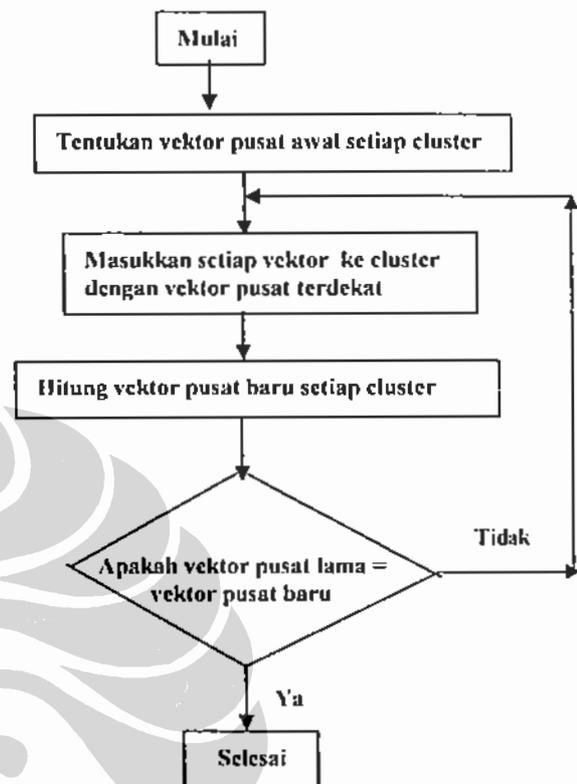
σ_i = variansi ciri ke- i

k = menunjukkan cluster ke- k

K = jumlah cluster

Selanjutnya dilakukan proses pemberian label pada setiap pola berdasarkan jarak terdekat atau jarak euclidean terhadap setiap pusat cluster. Pada akhir proses setiap iterasi dihitung pusat cluster yang baru. Proses ini berlangsung terus secara iteratif sampai diperoleh keadaan yang stabil, yaitu tidak ada pola yang pindah cluster lagi atau harga pusat cluster akhir sudah

tidak berubah lagi. Proses metoda *K-means clustering* dapat digambarkan seperti pada Gambar 1.



Gambar 1. Diagram metoda *K-means clustering*.

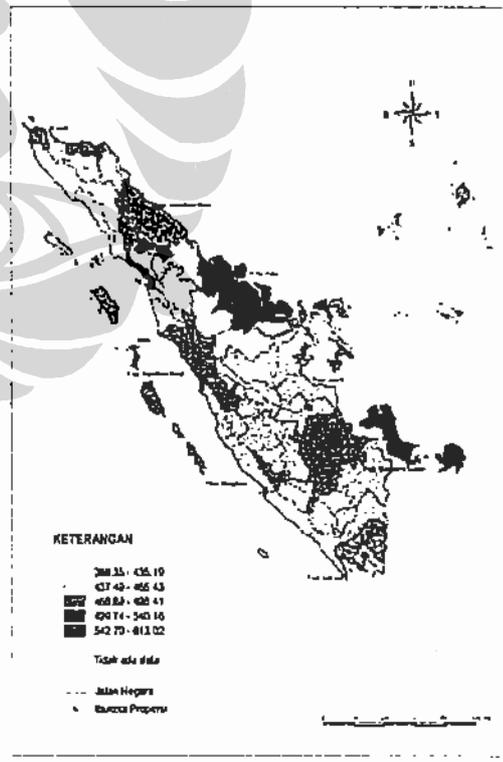
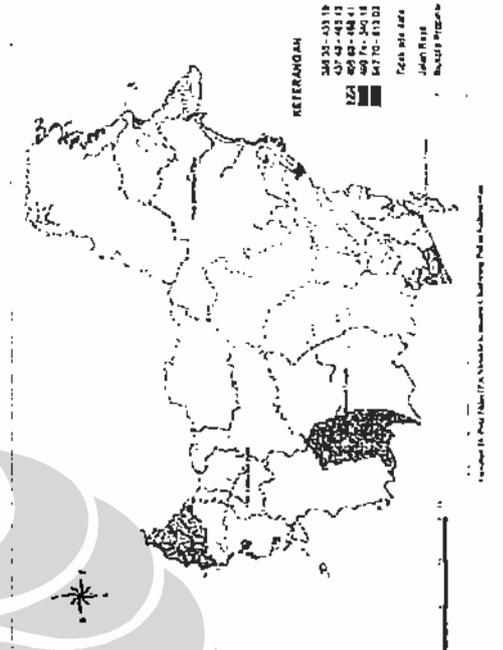
Dengan metoda *K-Means clustering*, data masukan berupa suatu set sampel yang terdiri dari pola-pola yang dapat mewakili sifat pola secara keseluruhan. Dari set sampel tersebut, dihitung vektor pusat untuk setiap cluster yang terbentuk. Pusat cluster tersebut kemudian digunakan sebagai ciri cluster dalam proses pengelompokan seluruh pola lainnya.

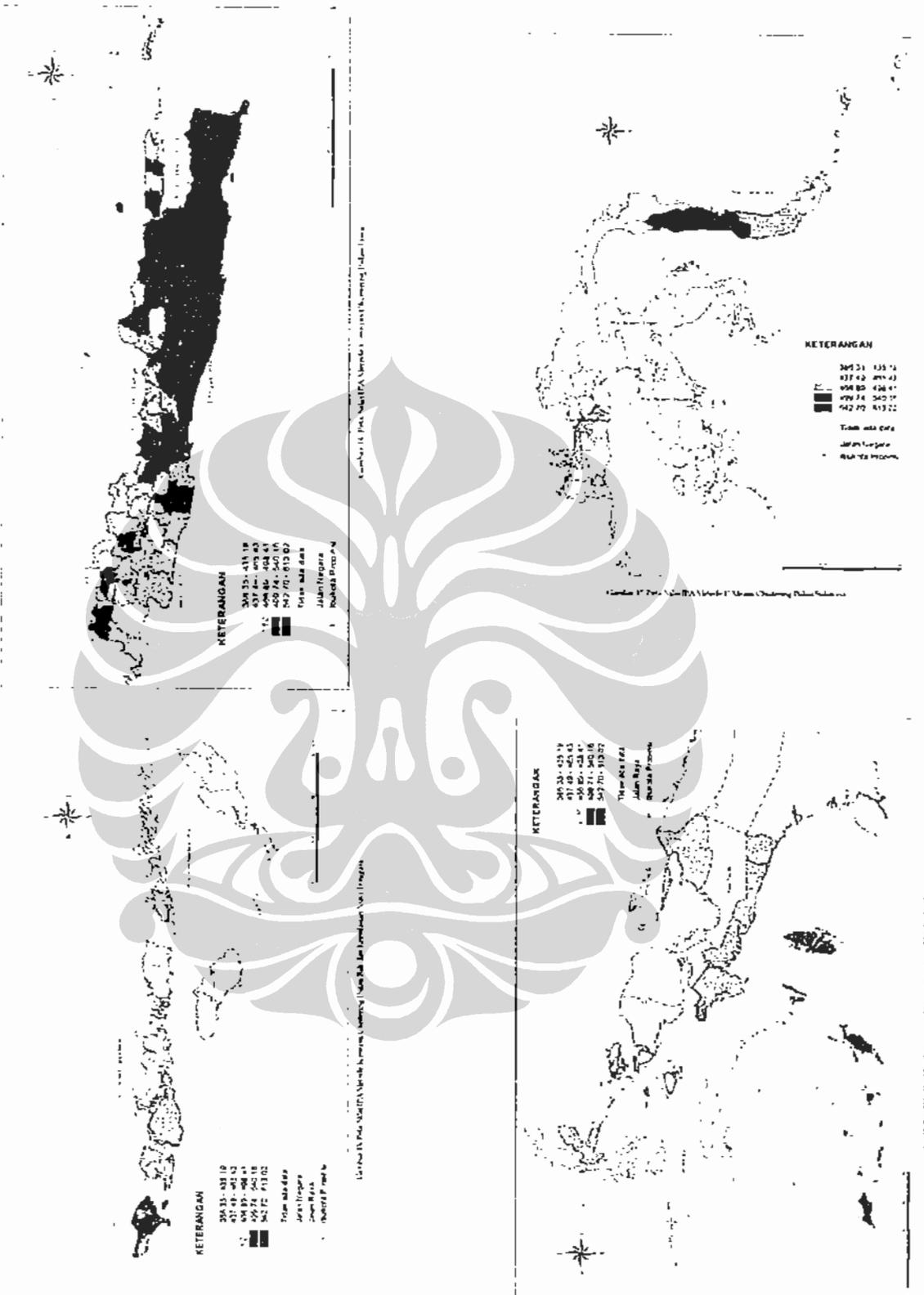
Hasil Pengelompokan Dan Hasil Pemetaan

Data yang digunakan pada penelitian ini adalah hasil seleksi UMPTN tahun 1999/2000. Hasil pengelompokan dengan Metoda *K-Means Clustering* tersebut adalah sebagai berikut:

Cluster 1:	386.35 - 435.19
Cluster 2:	437.49 - 465.43
Cluster 3:	466.89 - 498.41
Cluster 4:	499.74 - 540.16
Cluster 5:	542.70 - 613.02

Sedangkan hasil pemetaannya dapat dilihat pada serangkaian gambar berikut:





4. KESIMPULAN

Dari hasil pemetaan tersebut, dapat dilihat bahwa daerah-daerah di luar Pulau Jawa sebagian besar masih memiliki kualitas pendidikan yang rendah. Hanya beberapa daerah yang sudah memiliki kualitas pendidikan yang agak lumayan.

Bagi pengambil kebijakan dalam dunia pendidikan, hendaknya lebih memperhatikan daerah-daerah yang memiliki kualitas pendidikan yang sangat rendah (warna putih).

UCAPAN TERIMAKASIH

Bersama ini penulis mengucapkan terima kasih yang sebesar-besarnya kepada rekan-rekan dari 'Buana Katulistiwa' yang telah membantu penulis dalam menggunakan Arc-View GIS versi 3.1.

DAFTAR ACUAN

- [1] Arymurthy, Aniati Murni & Suryana Setiawan, *Pengantar Pengolahan Citra*. Elex Media Komputindo. Jakarta, 1992.
- [2] Duda, R & P. Hart, *Pattern Classification and Scene Analysis*. Wiley. New York, 1973.
- [3] Jain, Anil K. & Richard C. Dubes, *Algorithms for Clustering data*. Prentice Hall, Inc. New Jersey, 1988.
- [4] Jerke, Noel, *Visual Basic 6: The Complete Reference*. Mc. Graw Hill. New York, 1999.
- [5] Masoem, Toemin A, *Ebtanas atau UMPTN, Mana Yang Lebih Dapat Diandalkan?*, UI-Press. Jakarta, 1997.
- [6] Spath, Helmuth, *Cluster Analysis Algorithms*. John Wiley & Sons. New York, 1982.