

PEMBAGIAN KELAS HOMOGEN MENGGUNAKAN *K-MEANS CLUSTERING*

Nurhayadi

Program Studi Pendidikan Matematika

Email: nurhayadi@gmail.com

Abstrak: Kelas heterogen berisi siswa dengan kemampuan yang beragam. Bila guru menyampaikan materi dengan cepat, maka siswa yang kurang cerdas akan mengalami kesulitan. Sedangkan bila guru menyampaikan materi dengan lambat, maka siswa yang cerdas akan terasa kurang diuntungkan. Kelas homogen adalah kelas yang terdiri dari siswa dengan kemampuan yang setara. Kelas homogen memberikan kemudahan pada guru dalam menyampaikan materi sesuai dengan kemampuan siswa, dan hal ini berarti juga memberikan keuntungan bagi siswa itu sendiri. Pembagian kelas berdasar ranking jumlah nilai mata pelajaran sering digunakan untuk membentuk kelas yang homogen, akan tetapi tujuan ini sering tidak tercapai karena sebagian kelas yang terbentuk masih heterogen. Untuk mendapatkan kelas yang seluruhnya homogen, maka kelas disusun menggunakan *k-mean clustering*.

Kata kunci: kelas, homogen, peringkat, k-mean, clustering,

Abstract: Heterogeneous class provides students with diverse abilities. When teachers deliver material quickly, then the less intelligent students will have trouble. Meanwhile, when the teacher presenting the material at a slow pace, so the bright ones will feel disadvantaged. Homogeneous class is a class of students with similar capabilities. Homogeneous class provides convenience to teachers in presenting the material in accordance with the abilities of students, and it also provides benefits for the students themselves. The division of a class based on the ranking of the number of subjects values are often used to form a homogeneous class, but this goal is often not achieved because most classes are formed is still heterogeneous. To obtain a homogeneous whole class, then the class was compiled using k-mean clustering.

Keywords: class, homogeneous, ranking, k-mean, clustering,

Belajar adalah kegiatan yang dilakukan manusia untuk meningkatkan kemampuan. Winkel (1987) mengatakan bahwa bahwa belajar merupakan suatu proses pada manusia yang berlangsung dalam interaksi dengan lingkungannya dan menghasilkan perubahan-perubahan. Perubahan tersebut dapat berupa pengetahuan, pemahaman, ketrampilan, atau nilai sikap yang bersifat menetap. Belajar terjadi jika individu menerima rangsangan dari lingkungan, dan mengadakan respon terhadap rangsangan tersebut (Gagne, 1975).

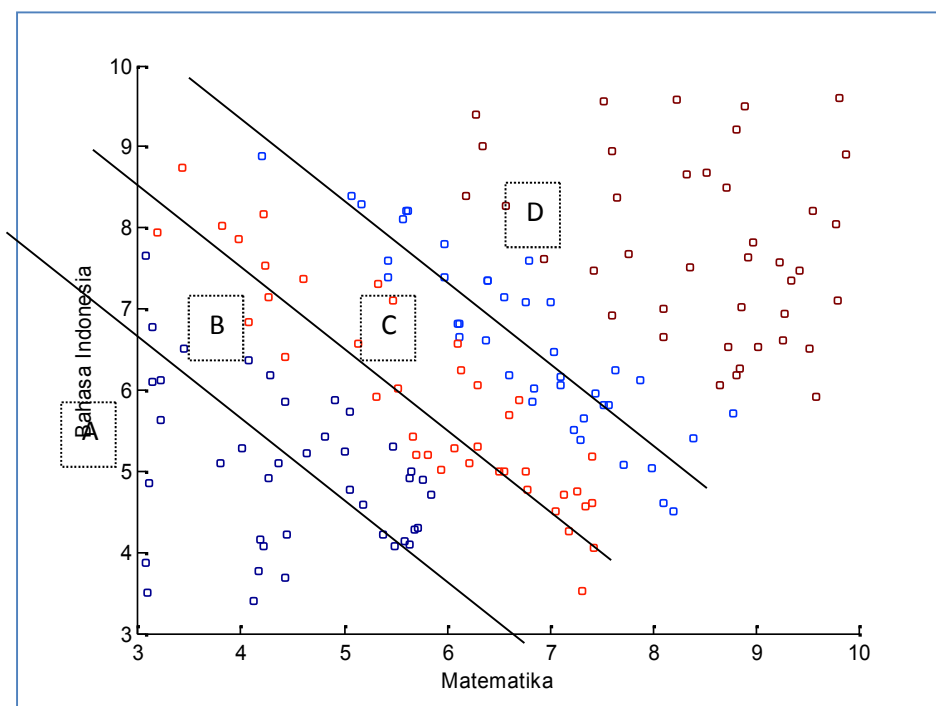
Belajar dapat dipandang sebagai suatu proses yang dipengaruhi oleh faktor eksternal seperti guru, bahan pelajaran, dan kondisi teman sekelas. Guru dapat memberikan materi pelajaran secara cepat bila siswa-siswanya mampu mengikuti. Akan tetapi seyogyanya penyampaian materi dilakukan dengan pelan bila siswanya mengalami kesulitan.

Kelas heterogen terdiri dari siswa dengan kemampuan yang beragam. Bila guru menyampaikan materi dengan cepat, maka siswa yang kurang cerdas akan mengalami kesulitan. Sedangkan bila guru menyampaikan materi dengan lambat, maka siswa yang cerdas akan terlihat kurang diuntungkan. Pembentukan kelas homogen, yakni kelas yang terdiri dari siswa dengan kemampuan setara, akan memudahkan guru dalam menyampaikan

materi sesuai dengan kemampuan siswa. Selain memudahkan tugas guru, pembentukan kelas homogen juga akan menguntungkan bagi siswa.

Pembentukan kelas berdasar kemampuan homogen, paling mudah adalah dengan menjumlahkan nilai dalam ijazah atau buku raport, kemudian dilakukan pengelompokan menurut peringkat. Akan tetapi pembagian kelas dengan cara ini memiliki kelemahan karena sering terbentuk beberapa kelas homogen semu.

Sebagai ilustrasi, siswa sebanyak 160 akan dibagi menjadi 4 kelas berdasar peringkat jumlah nilai matematika dan nilai IPS, sebagaimana ditunjukkan pada Gambar-1. Kelas yang terbentuk berdasar peringkat ditunjukkan dengan warna yang berbeda, dibatasi dengan garis-garis dan ditandai sebagai kelas A, kelas B, kelas C, dan kelas D.



Gambar 1. Pembagian kelas menggunakan peringkat jumlah nilai

Dari Gambar 1 dapat dilihat bahwa pembagian kelas berdasar peringkat jumlah nilai mata pelajaran memiliki sifat sebagai berikut:

Kelas A: adalah siswa yang memiliki nilai tinggi dalam pelajaran matematika dan IPS.

Kelas B: adalah siswa yang memiliki jumlah nilai di atas rata-rata

Kelas C: adalah siswa yang memiliki jumlah nilai di bawah rata-rata

Kelas D: adalah siswa yang memiliki nilai kurang pada kedua mata pelajaran.

Seorang siswa memiliki nilai tinggi pada pelajaran matematika dan nilai yang rendah pada pelajaran IPS dapat berada dalam satu kelas dengan seorang siswa yang memiliki nilai rendah pada pelajaran matematika dan memiliki nilai tinggi pada pelajaran IPS apabila jumlah nilainya relatif sama, sebagaimana ditunjukkan dalam kelas B dan kelas C. Maka ada kelas yang di dalamnya terdapat siswa yang memiliki nilai matematika tinggi dan rendah, yang artinya heterogen untuk mata pelajaran matematika. Hal ini berarti pembentukan kelas homogen kurang berhasil. Kelas homogen yang terbentuk hanya kelas A dan kelas D.

Sedangkan kelas B dan kelas C sebenarnya adalah heterogen. Agar didapat kelas yang seluruhnya homogen, maka diperlukan suatu cara lain untuk melakukan pembagian kelas.

Pembagian kelas pada contoh di atas hanya melibatkan dua mata pelajaran. Apabila melibatkan lebih dari dua mata pelajaran, maka pembagian kelas akan lebih rumit lagi. Untungnya, para peneliti statistik telah mengembangkan cara pembagian kelas yang sistematis, yakni menggunakan data mining. Data mining adalah proses penambangan informasi pada data yang tersedia dalam jumlah besar, gunanya untuk menemukan hubungan antar data yang belum diketahui. Selain itu, juga digunakan untuk menyederhanakan data agar lebih mudah dipahami dan lebih mudah untuk digunakan. Data mining sudah banyak digunakan di bidang bisnis, industri, kesehatan dan kebudayaan. Satu bagian ilmu dari data mining, yakni *k-means clustering* akan digunakan untuk pembagian kelas.

Proses pembagian kelas menggunakan *k-means clustering* adalah proses yang panjang. Akan tetapi, dengan berkembangnya teknologi informasi dan komunikasi yang sangat pesat, pembagian kelas menggunakan cara ini sudah bukan merupakan hal yang sulit dilakukan bila kita menggunakan komputer.

KAJIAN PUSTAKA

Tulisan ini membahas pembagian kelas menggunakan *k-means clustering*. Derajat homogenitas diukur menggunakan *Mean Square Error (MSE)*. Sebagai bahan evaluasi, homogenitas hasil simulasi pembagian kelas menggunakan *k-mean clustering* akan dibandingkan dengan pembagian kelas menggunakan peringkat.

Mean Square Error

Mean Square Error adalah rata-rata kuadrat jarak anggota kluster terhadap pusatnya. Semakin besar *MSE*, maka berarti semakin luas sebaran data dari pusatnya. Besarnya *MSE* dihitung dengan rumus:

$$MSE = \frac{1}{n_k} \sum_{i=1}^{n_k} (x_i - \bar{x})^2 \text{ dengan } \bar{x}_k = \text{titik pusat dari cluster ke-k, } n_k = \text{banyaknya data pada}$$

cluster ke-k dan $x_i = \text{data ke-i pada cluster ke-k}$

Jarak *Euclidean*

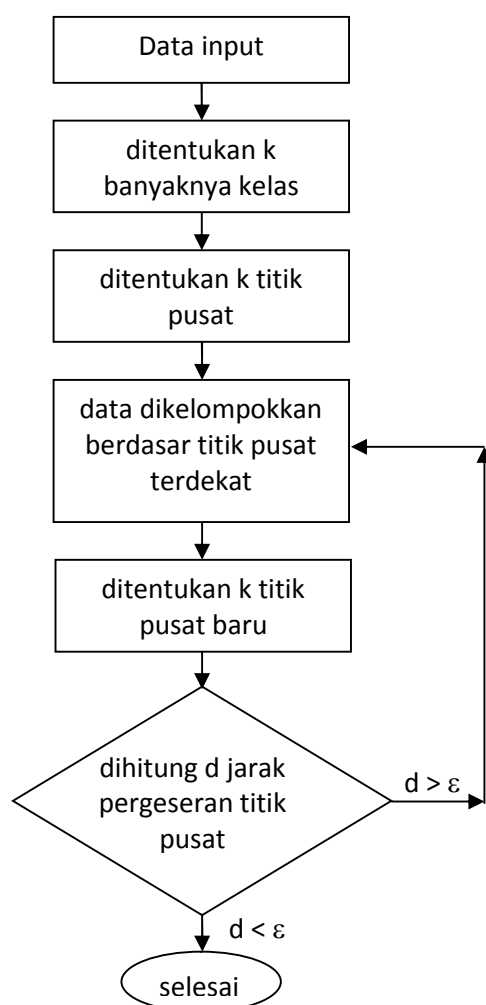
Proses pengelompokan data ke dalam suatu *cluster* dilakukan dengan cara menghitung jarak terdekat dari sebuah data ke sebuah titik pusat. Salah satu perhitungan jarak yang sering digunakan adalah jarak *Minkowski*. Russell S and Norvig P (2010) menyatakan bahwa perhitungan jarak antar dua buah data menggunakan cara *Minkowski* adalah menggunakan rumus:

$d(x_1, x_2) = \left(|x_{11} - x_{21}|^g + |x_{12} - x_{22}|^g + \dots + |x_{1p} - x_{2p}|^g \right)^{1/g}$ dengan x_1, x_2 adalah dua buah data koordinat yang akan dihitung jaraknya, p adalah dimensi dari sebuah data, dan g adalah suatu bilangan bulat positif

Jika $g = 2$, maka jarak *Minkowski* sering disebut dengan jarak *Euclidean*

k-means clustering

Pembagian kelas menggunakan *k-means clustering* adalah pengelompokan titik berdasar sifat yang saling berdekatan. Kedekatan antar titik diukur berdasar aturan jarak yang ditentukan. Prinsip kerja pembagian kelas dengan metoda *k-means clustering* sebagaimana ditunjukkan pada Gambar-2, dimulai dengan menentukan banyaknya kelas yang akan dibentuk, misal sebanyak k , kemudian diambil sebanyak k buah titik sembarang sebagai titik pusat kelas, dalam hal ini koordinat titik adalah nilai siswa. Setiap titik akan dihitung jaraknya terhadap sebanyak k titik pusat. Berdasar jarak *Minkowski*, maka masing-masing titik dikelompokkan pada salah satu pusat yang memiliki jarak terdekat.



Gambar-2. Algoritma *k-means clustering*

Setiap kelas yang terbentuk, diambil titik pusat baru. Pengambilan pusat yang baru, dilakukan menggunakan rumus

$$\bar{x}_k = \frac{1}{n_k} \sum_{i=1}^{n_k} x_i \text{ dengan: } \bar{x}_k = \text{titik pusat dari kelas ke-}k, n_k = \text{banyaknya data pada kelas ke-}$$

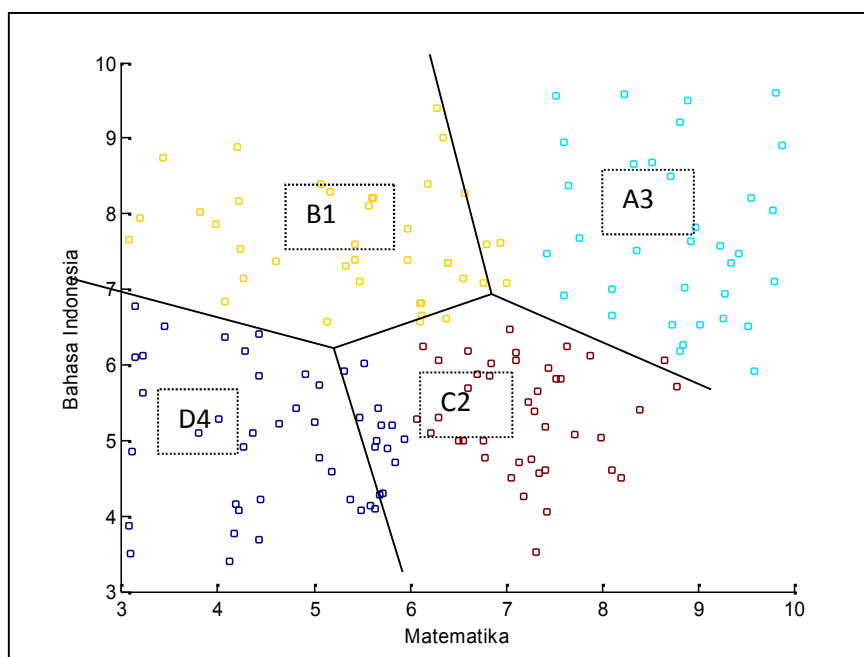
k , dan x_i = data ke- i pada kelas *cluster* ke- k

Berdasar titik pusat baru tersebut, disusun kembali kelas baru. Pada kelas yang baru, diambil titik pusat yang baru lagi, dan dibentuk lagi kelas yang baru. Hal ini dilakukan

berulang-ulang sampai terjadi jarak pergeseran titik pusat kurang dari suatu bilangan ε yang ditentukan.

HASIL DAN PEMBAHASAN

Percobaan metoda *k-means clustering* untuk pembagian kelas, menggunakan data sebagaimana ditunjukkan pada Gambar-1. Diambil empat buah titik titik pusat sembarang, idealnya adalah titik yang berada di sekitar titik-titik yang akan dibagi kelas. Pembagian kelas menggunakan *k-means clustering* memberikan hasil sebagaimana ditunjukkan pada Gambar-3.



Gambar 3. Pembagian kelas menggunakan *k-means clustering*

Dari Gambar 3 dapat dilihat bahwa *k-means clustering* mengelompokkan siswa berdasar kemiripan kemampuan dalam dua mata pelajaran. Masing-masing kelas yang terbentuk memiliki sifat sebagai berikut:

- Kelas A: adalah siswa yang memiliki nilai tinggi dalam pelajaran matematika dan IPS.
- Kelas B: adalah siswa yang memiliki nilai tinggi untuk mata pelajaran matematika, tetapi agak kurang untuk mata pelajaran IPS
- Kelas C: adalah siswa yang memiliki nilai tinggi untuk mata pelajaran IPS, tetapi agak kurang untuk pelajaran matematika
- Kelas D: adalah siswa yang memiliki nilai kurang pada kedua mata pelajaran.

Kelas yang terbentuk pada pengelompokan menggunakan *k-means clustering* relatif lebih homogen dari pada pembagian kelas berdasar peringkat jumlah nilai. Bila kita amati lebih lanjut, MSE pada pembagian kelas menggunakan *k-means clustering* secara keseluruhan lebih kecil dari MSE pada pembagian kelas menggunakan peringkat jumlah nilai. Nilai MSE dari pembagian kelas masing-masing cara, ditunjukkan pada Tabel-1.

Nilai MSE yang kecil juga berarti bahwa kelas yang terbentuk mengumpul dekat pada titik pusat kelompok. Hal ini juga berarti bahwa kemampuan antar individu tidak jauh berbeda untuk seluruh mata pelajaran. Dengan demikian, ketika guru masuk di kelas dapat lebih fokus dalam menyampaikan mata pelajaran sesuai dengan kebutuhan siswa.

Tabel-1. *Mean Square Error* pembagian kelas

Kelas	K-means clustering	Peringkat
A	1.6817	2.2173
B	1.6871	2.2102
C	0.9052	3.1271
D	1.5701	1.7572
Jumlah	5.8441	9.3119

KESIMPULAN DAN SARAN

Pembentukan kelas homogen akan memudahkan guru dalam mengajar. Untuk membentuk kelas yang seluruhnya homogen, dapat dilakukan menggunakan metoda *k-means clustering*. Homogenitas kelas hasil pengelompokan dengan cara tersebut lebih baik dari pada pembagian kelas berdasar peringkat jumlah nilai.

Pembentukan kelas menggunakan metoda *k-means clustering* memiliki kelemahan, yakni kelas yang terbentuk sering tidak sama besar. Untuk mengatasi masalah ini, maka perlu dicari cara memodifikasi algoritma, agar kelas yang terbentuk bisa sama besar.

DAFTAR PUSTAKA

Gagne, R.M., 1975, *Essentials of Learning for Instruction*, New York

Russell. S dan Norvig. P., 2010, *Artificial Intelligence A Modern Approach. Upper Saddle River*, New Jersey

Winkel. W., 1987, *Psikologi Pengajaran*, Gramedia, Jakarta.