

Menentukan Efektivitas Pembelajaran Online Menggunakan Algoritma Agglomerative Hierarchical Clustering

MENENTUKAN EFEKTIVITAS PEMBELAJARAN ONLINE MENGGUNAKAN ALGORITMA ANGLOMERATIVE HIERARCHICAL CLUSTERING

I Kadek Dwi Nuryana

Program Studi Sistem Informasi, Universitas Negeri Surabaya

Email : dwinuryana@unesa.ac.id

Much. Zuyyinal Haqqul Barir

Magister Informatika, Fakultas Sains dan teknologi, UIN Maulana Malik Ibrahim Malang

Email : haqqulb3@gmail.com

Zainal Ikhwan Muhammad

Program Studi Sistem Informasi, Fakultas Teknologi Informasi, Universitas Hasyim Asy'ari

Email : zainalikhwan@unhasy.ac.id

Abstrak

Metode clustering adalah metode yang dapat mengelompokkan untuk menghasilkan proses pemodelan, serta menganalisis data dengan sistem partisi. Dalam mengklasifikasikan kepribadian siswa, kami menggunakan Algoritma Agglomerative Hierarchical Clustering (AHC). AHC adalah teknik pengelompokan data yang membagi data menjadi beberapa kelompok berdasarkan kesamaannya. Penelitian ini bertujuan untuk mengklasifikasikan kepribadian mahasiswa dalam perkuliahan berbasis E-Learning dengan metode AHC. Dalam penelitian ini, perhitungan skor dilakukan dalam tiga proses mengikuti proses kluster hirarkis yang dilakukan dalam tiga proses. Proses ini dilakukan tiga kali dengan tujuan mengamati perkembangan distribusi partikel selama proses cluster hierarkis. Dapat dilihat bahwa skor pertama menghasilkan nilai 0,54, kemudian skor kedua menghasilkan nilai 0,51, sedangkan yang terakhir menghasilkan nilai 0,54. Kemudian dibutuhkan nilai terakhir untuk menentukan seberapa baik penyebarannya. Dapat dilihat dari nilai terakhir bahwa nilai yang dihasilkan berada di atas 0 dan mendekati 1 sehingga dapat dikatakan bahwa distribusi partikelnya cukup baik.

Kata Kunci : AHC, Clustering, Optimasi, Kepribadian, PSO, Mahasiswa,

Abstract

The clustering method is a method that can group to produce modeling processes, as well as analyze data with a partition system. In classifying student personalities, we use the Agglomerative Hierarchical Clustering (AHC) Algorithm. AHC is a data grouping technique that divides data into groups based on their similarity. This study aims to classify student personalities in lectures based on E-Learning with the AHC method. In this study, the score calculation was carried out in three processes following the hierarchical cluster process which was carried out in three processes. This process is carried out three times with the aim of observing the development of particle distribution during the hierarchical cluster process. It can be seen that the first score produced a value of 0.54, then the second score produced a value of 0.51, while the last one produced a value of 0.54. Then it takes the last value to determine how well it is spread. It can be seen from the last value that the resulting value is above 0 and close to 1 so that it can be said that the distribution of the particles is fairly good.

Keywords : AHC, Clustering, Optimization, Personality, PSO, Student.

Menentukan Efektivitas Pembelajaran Online Menggunakan Algoritma Agglomerative Hierarchical Clustering

PENDAHULUAN

Tahun 2019 merupakan tahun dimana seluruh wilayah di dunia dilanda wabah baru, yaitu Coronavirus Disease 2019 (COVID-19). Wabah virus telah mengakibatkan perubahan besar dalam kehidupan. Pada awalnya, orang bisa hidup normal dan berinteraksi langsung satu sama lain, tetapi sekarang mereka harus dibatasi atau bahkan tidak dapat melakukannya sama sekali, salah satunya di Indonesia. Dengan merebaknya virus Covid-19, banyak negara telah mengambil kebijakan untuk memutus rantai penularan dengan memberlakukan lockdown (karantina wilayah). Kebijakan lockdown di Indonesia diberi nama Pembatasan Sosial Berskala Besar (PSBB) yang diberlakukan per daerah, baik provinsi maupun kabupaten/kota berdasarkan tingkat penyebaran virus Covid-19[12].

Dengan kebijakan ini, salah satunya berdampak pada sektor pendidikan, yaitu Pemerintah telah menetapkan kebijakan untuk mewajibkan siswa dan guru melakukan Proses Belajar Mengajar (PBM) dari kelas tatap muka hingga pembelajaran daring. Hal ini dilakukan untuk memutus rantai penularan Covid-19, di mana siswa dapat berperan sebagai pembawa dan penyebar penyakit tanpa gejala, dan hampir semua negara membatalkan kegiatan di sekolah. Kebijakan pemerintah yang dinilai mendadak sangat mempengaruhi kurangnya kesiapan sekolah, guru dan siswa[13]. Mulai dari kesiapan guru yang tidak menguasai teknologi internet atau media online yang tersedia sebagai media pembelajaran. Terutama bagi siswa yang harus belajar di rumah yang belum terbiasa dengan pembelajaran daring. Hal ini mengakibatkan pembelajaran menjadi kurang efisien, siswa tidak berkonsentrasi dalam memahami bahan ajar, kurangnya diskusi dan tanya jawab tentang materi yang diajarkan selama pembelajaran [14].

Salah satu metode clustering atau pengelompokan data mining dapat membantu mengetahui pola pengelompokan kepribadian siswa dalam proses pembelajaran daring, sehingga membuat cluster dan lainnya akan berada dalam cluster yang berbeda. Ini dapat membantu pendidik menemukan strategi yang efektif untuk pembelajaran online.

Dalam penelitian ini, peneliti tertarik untuk mengklasifikasikan minat dan kebutuhan literasi siswa menggunakan Agglomerative Hierarchical Clustering (AHC). Oleh karena itu, judul yang diambil dari penelitian ini adalah "MENENTUKAN EFEKTIVITAS PEMBELAJARAN ONLINE MENGGUNAKAN ALGORITMA ANGGLOMERATIVE HIERARCHICAL CLUSTERING".

Data Mining

Data mining contains the search for the desired pattern in a large database to assist future decision making. These patterns are recognized by certain tools which can provide a useful analysis of the data and can then be studied in more detail, which may be using other decision support tools." [15]. This technique involves using computer science, statistics, mathematics, and artificial intelligence to extract useful information from structured and unstructured data, such as relational databases, text documents, and multimedia data. Data mining is used in a variety of fields, including marketing, customer relationship management, fraud detection, scientific discovery, and homeland security.

Agglomerative Clustering

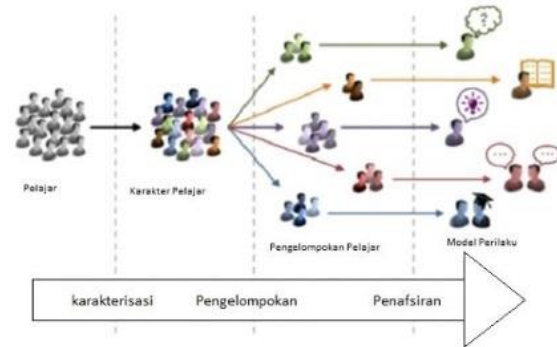
Clustering adalah metode analisis data yang sering dimasukkan sebagai metode yang bertujuan untuk mengelompokkan data dengan karakteristik yang sama ke dalam wilayah yang sama dan karakteristik yang berbeda ke wilayah lain. Pengelompokan aglomeratif adalah bagian dari pengelompokan berdasarkan hierarki. Pengelompokan aglomeratif dimulai dengan membuat setiap data menjadi kluster. Cluster ini akan bergabung berdasarkan jarak antar cluster. Penggabungan cluster akan terus terjadi sampai jumlah cluster yang diinginkan diproduksi [2].

Untuk menghitung jarak antar cluster dalam algoritma aglomeratif dapat dilakukan dengan beberapa cara. Salah satunya adalah menggunakan metode single linkage. Algoritma single linkage mencari jarak antara dua cluster sesuai dengan jarak terpendek antara dua anggota dalam dua cluster [5]. Dalam metode ini kedekatan antara dua kelompok ditentukan dari jarak terpendek antara pasangan dua data dari dua kelompok yang berbeda (satu dari kelompok pertama dan satu dari kelompok lain). Rumusnya dapat dilihat pada Persamaan 1 di bawah ini.

Menentukan Efektivitas Pembelajaran Online Menggunakan Algoritma Agglomerative Hierarchical Clustering

$$d_{uv} = \min \{d_{uv}\}, d_{uv} \in d$$

METODE



Gambar 1. Pemodelan Perilaku Mahasiswa Menggunakan Analisis Cluster

Deskripsi:

Penokohan, data mentah dari kegiatan siswa dikumpulkan terlebih dahulu dan hasil dari preprocessing ini adalah seperangkat fitur yang mewakili dan mencirikan siswa dalam hal perilaku mereka mengenai kedua jenis kegiatan dan data yang tersedia pemodelan kegiatan siswa secara online, yang terdiri dari berbagai tingkat partisipasi dan unit analisis yang telah ditentukan.

Clustering (Pengelompokan), siswa yang ditandai kemudian digunakan sebagai objek input untuk algoritma pengelompokan yang mengelompokkannya menurut kesamaan (atau kedekatannya), mereka adalah siswa dengan algoritma pengelompokan tertentu untuk diterapkan, memilih ukuran kedekatan, atau menentukan jumlah cluster optimal dalam data.

Interpretasi. Setiap kluster yang dihasilkan merupakan analisis subjek aktivitas siswa, kesamaan perilaku siswa dan ketidaksetaraan perilaku. Cluster dan menafsirkan perilaku belajar yang diwakili oleh masing-masing. Kluster mahasiswa dan maknanya sesuai dengan target kegiatan analisis. Interpretasi model ini dari perilaku yang dihasilkan harus dilakukan sesuai dengan fitur yang dipilih pada tahap pertama proses dan, dalam kasus khusus dari kegiatan pemodelan siswa dalam pembelajaran online.

HASIL DAN PEMBAHASAN

Implementasi algoritma Agglomerative Hierarchical Clustering menggunakan Google Colab dijelaskan sebagai berikut:

1. Informasi data terdiri dari 4 variabel, yaitu nomor ID siswa yang digunakan sebagai ID, nilai rata-rata penilaian, domain perilaku, dan parameter mata kuliah yang diberikan siswa. Dapat dilihat pada gambar berikut:

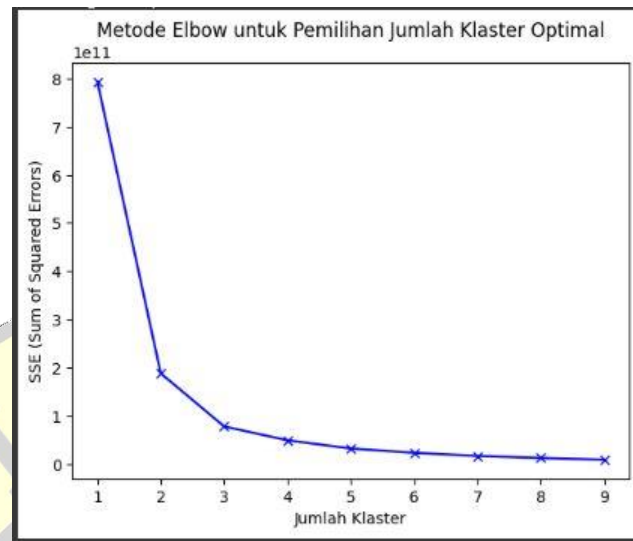
	id	Perhitunagn Kepribadian Akhir	Mean Penilaian	Parameter_Course
0	201500010504	3	90.666667	1
1	201500018336	2	66.666667	2
2	201500058841	1	74.000000	3
3	201500059447	1	79.333333	4
4	201500061735	2	90.666667	5

Gambar 2. Data Mahasiswa

2. Tentukan $n_{cluster}$ dengan metode Elbow

Penentuan $n_{cluster}$ untuk penelitian ini ditentukan dengan menggunakan metode siku dimana metode siku berfungsi untuk menentukan jumlah siku atau lekukan optimal pada grafik yang menunjukkan gugus dari dataset:

Menentukan Efektivitas Pembelajaran Online Menggunakan Algoritma Angglomerative Hierarchical Clustering



Gambar 3. Diagram Siku

Berdasarkan diagram pada Gambar 5, saya memilih nomor 4 sebagai jumlah cluster yang saya ambil. Pada angka 4 terdapat 'siku' yang signifikan pada cluster sum kuadrat, kemudian pembagian cluster diambil berdasarkan empat tipe kepribadian tersebut. Oleh karena itu, pemilihan nomor 4 sebagai jumlah cluster adalah yang paling tepat untuk dataset yang dianalisis.

3. Lakukan Hierarki

Proses pengelompokan hierarkis dilakukan secara bertahap tiga kali untuk menentukan cluster seperti yang dijelaskan pada gambar berikut:

```
from sklearn.cluster import AgglomerativeClustering
from sklearn.datasets import load_iris
import pandas as pd

# Contoh data
X = df_1[['id', 'Mean Penilaian', 'Perhitunagn Kepribadian Akhir', 'Parameter_Course']]

# Hierarki pertama: A1 dan A2
subset_data = X.iloc[:, :2]
clustering1 = AgglomerativeClustering(n_clusters=4)
clustering1.fit(subset_data)
labels1 = clustering1.labels_

# Hierarki kedua: Hasil clustering pertama (A1 dan A2) dan A3
subset_data2 = pd.concat([pd.Series(labels1, name='cluster'), X.iloc[:, 2]], axis=1)
clustering2 = AgglomerativeClustering(n_clusters=4)
clustering2.fit(subset_data2)
labels2 = clustering2.labels_

# Hierarki ketiga: Hasil clustering kedua (A1, A2, dan A3) dan A4
subset_data3 = pd.concat([pd.Series(labels2, name='cluster'), X.iloc[:, 3]], axis=1)
clustering3 = AgglomerativeClustering(n_clusters=4)
clustering3.fit(subset_data3)
labels3 = clustering3.labels_

# Menampilkan hasil klastering
print("Hasil Klastering:")
for i, label in enumerate(labels3):
    print("Data ke", i, ":", X.iloc[i].values, " Klaster:", label)
```

Gambar 4. Program Klaster Hirarki

Gambar 6 menunjukkan program clustering yang memiliki proses clustering dari 3 tahap clustering. Setiap tahap cluster sesuai dengan urutan data yang dimasukkan dalam array data. Dalam array data ada nama kolom dalam dataset yang akan dikelompokkan.

Menentukan Efektivitas Pembelajaran Online Menggunakan Algoritma Anglomerative Hierarchical Clustering

Pada proses pertama, pengelompokan dilakukan berdasarkan id cluster data dan peringkat rata-rata. Pada proses kedua, dilakukan clustering antara hasil cluster pertama dan data ketiga, yaitu perhitungan kepribadian akhir. Tahap terakhir dilakukan dengan clustering antara hasil cluster ke-2 dengan data keempat, yaitu mata kuliah penilaian. Dalam proses ini, algoritma cluster hirarkis dapat diproduksi sebagai berikut:

$$\Sigma = (((A1 + A2) + A3) + A4)$$

Dari pengelompokan tersebut dihasilkan data sampel sebagai berikut:

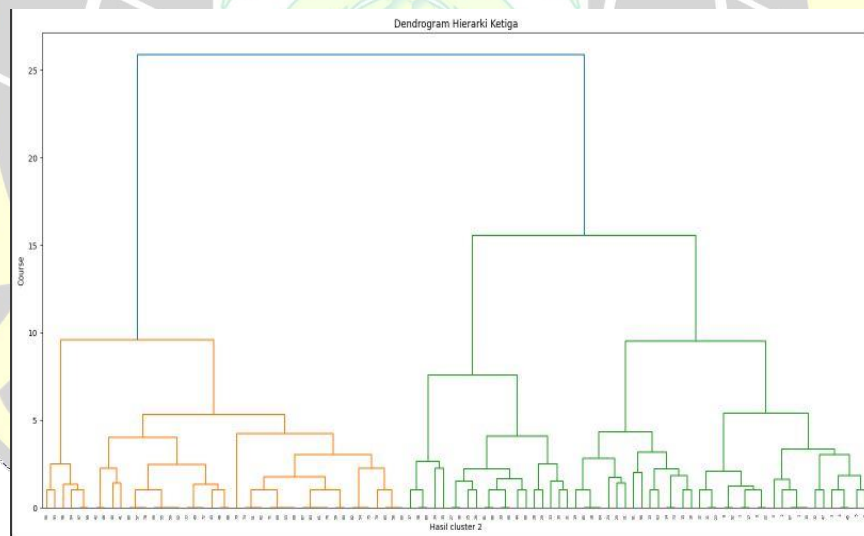
```
Data ke 21 : [2.01500128e+11 8.36666667e+01 3.00000000e+00 1.90000000e+01] Klaster: 3
Data ke 22 : [2.01500132e+11 9.30000000e+01 2.00000000e+00 9.00000000e+00] Klaster: 1
Data ke 23 : [2.01500133e+11 8.00000000e+01 1.00000000e+00 1.00000000e+01] Klaster: 1
Data ke 24 : [2.01500135e+11 7.36666667e+01 1.00000000e+00 2.00000000e+01] Klaster: 3
Data ke 25 : [2.01500156e+11 8.96666667e+01 2.00000000e+00 2.10000000e+01] Klaster: 0
Data ke 26 : [2.01500157e+11 8.46666667e+01 2.00000000e+00 2.20000000e+01] Klaster: 0
```

Gambar 5. Pengujian tautan rata-rata

Dalam contoh hasil, clustering menghasilkan data dengan kondisi data dalam dataset baris ke-n memiliki id, nilai rata-rata, tipe kepribadian, dan course yang dimiliki dimasukkan dalam cluster n.

4. Menampilkan Dendrogram

Setelah menghasilkan cluster dari setiap dataset, dendrogram ditampilkan untuk melihat distribusi clustering yang telah dibuat seperti yang ditunjukkan pada gambar berikut.



Gambar 6. Cluster Dendrogram

5. Pengujian Skor Siluet.

Uji Silhouette Score bertujuan untuk mengetahui seberapa baik partikel data didistribusikan pada hasil pengelompokan. Dalam penilaian skor Silhouette, dapat dilihat dengan rentang nilai 1 hingga -1. Semakin dekat nilai ke 1, semakin baik dispersi partikel, sedangkan jika nilai menghasilkan nilai di bawah 0, maka dispersi partikel mungkin memiliki masalah. Proses perhitungan skor siluet dapat dilihat pada gambar berikut.

Menentukan Efektivitas Pembelajaran Online Menggunakan Algoritma Agglomerative Hierarchical Clustering

```
from sklearn.metrics import silhouette_score

# Menghitung Silhouette Score untuk hierarki pertama
silhouette_avg1 = silhouette_score(subset_data, labels1)
print("Silhouette Score Hierarki Pertama:", silhouette_avg1)

# Menghitung Silhouette Score untuk hierarki kedua
silhouette_avg2 = silhouette_score(subset_data2, labels2)
print("Silhouette Score Hierarki Kedua:", silhouette_avg2)

# Menghitung Silhouette Score untuk hierarki ketiga
silhouette_avg3 = silhouette_score(subset_data3, labels3)
print("Silhouette Score Hierarki Ketiga:", silhouette_avg3)

Silhouette Score Hierarki Pertama: 0.5462854684508085
Silhouette Score Hierarki Kedua: 0.5187308288253318
Silhouette Score Hierarki Ketiga: 0.5456137662432242
```

Gambar 7. Program Kluster Hirarki

Dalam proses penghitungan skor, tiga proses dihitung mengikuti proses hirarki cluster yang dilakukan sebanyak tiga proses. Proses ini dilakukan tiga kali dengan tujuan mengamati perkembangan distribusi partikel selama proses cluster hierarkis. Dapat dilihat bahwa skor pertama menghasilkan nilai 0,54, kemudian skor kedua menghasilkan nilai value of 0.51, sedangkan skor terakhir menghasilkan nilai 0,54. Kemudian nilai terakhir diambil untuk menentukan seberapa baik spreadnya. Hal ini dapat dilihat dari nilai terakhir bahwa nilai yang dihasilkan berada di atas 0 dan mendekati 1 sehingga dapat dikatakan bahwa distribusi partikel terbilang baik

PENUTUP

Simpulan

Dalam penelitian ini, pengelompokan kepribadian siswa dalam proses pembelajaran daring menggunakan metode agglomerative hierarchical clustering. Perhitungan skor dilakukan dengan menghitung sebanyak tiga proses berikut dari proses Hierarchical Cluster yang dilakukan oleh tiga proses. Proses ini dilakukan tiga kali dengan tujuan mengamati perkembangan distribusi partikel selama proses cluster hierarkis. Dapat dilihat bahwa skor pertama menghasilkan nilai 0,54, kemudian skor kedua menghasilkan nilai 0,51, sedangkan yang terakhir menghasilkan nilai 0,54. Kemudian nilai terakhir diambil untuk menentukan seberapa baik spreadnya. Dapat dilihat dari nilai terakhir bahwa nilai yang dihasilkan berada di atas 0 dan mendekati 1 sehingga dapat dikatakan bahwa distribusinya cukup baik.

Saran

Adapun saran untuk pengembang selanjutnya adalah digabungkan dengan metode lain agar menemukan keakuratan yang lebih baik dan keterbaruan dalam penelitian.

DAFTAR PUSTAKA

- [1] T. Alfina, B. Santosa, and J. A. R. Hakim, "Analisa Perbandingan Metode Hierarchical Clustering, K-means dan Gabungan Keduanya dalam Cluster Data (Studi kasus : Problem Kerja Praktek Jurusan Teknik Industri ITS)," vol. 1, p. 5, 2012.
- [2] C. C. Anggarwal, Data Clustering Alorithm and Applications, vol. 20121229. Chapman and Hall/CRC, 2012.
- [3] G. Gan, C. Ma, and J. Wu, Data clustering: theory, algorithms, and applications. Philadelphia, Pa. : Alexandria, Va: SIAM, Society for Industrial and Applied Mathematics; American Statistical Association, 2007.
- [4] N. P. E. Merliana, Ernawati, and Alb. J. Santoso, "Analisa Penentuan Jumlah Cluster Terbaik Pada Metode K-Means Clustering," pp. 1–7.
- [5] E. Mooi and M. Sarstedt, A concise guide to market research: the process, data, and methods using IBM SPSS statistics. Berlin: Springer, 2011.

Menentukan Efektivitas Pembelajaran Online Menggunakan Algoritma Agglomerative Hierarchical Clustering

- [6] A. F. Muhammad, “KLASTERISASI PROSES SELEKSI PEMAIN MENGGUNAKAN ALGORITMA K-MEANS,” p. 5.
- [7] E. Muningsih, “Optimasi Jumlah Cluster K-Means Dengan Metode Elbow Untuk Pemetaan Pelanggan,” pp. 105–114, Sep. 2017.
- [8] E. Prasetyo, Data Mining - Konsep dan Aplikasi Menggunakan MATLAB. Yogyakarta: Andi Offset, 2012.
- [9] P. Sari and B. Pramono, “IMPROVE K-MEANS TERHADAP STATUS NILAI GIZI PADA BALITA,” p. 6.
- [10] M. A. Syakur, B. K. Khotimah, E. M. S. Rochman, and B. D. Satoto, “Integration K-Means Clustering Method and Elbow Method For Identification of The Best Customer Profile Cluster,” IOP Conf. Ser.: Mater. Sci. Eng., vol. 336, p. 012017, Apr. 2018.
- [11] Puspitasari, A., Mahmudah, I., Maharani, C. A., Pratiwi, I. D., & Darmadi. (2022). Implementasi Kurikulum 2013 Saat Pembelajaran Daring di MAN 2 Madiun Melalui Tahapan Perencanaan, Pelaksanaan, dan Evaluasi.
- [12] Jagad, A. (2021). Efektivitas Pembelajaran Daring di Masa Pandemi Covid 19 Bagi Mahasiswa Universitas PGRI Yogyakarta. Jurnal Basicedu, 9 (2).
- [13] Anung, A. (2020). Pengaruh Kebijakan Sosial Distancing Pada Wabah Covid-19 Terhadap Kelompok Rentan di Indonesia. Jurnal Kebijakan Kesehatan Indonesia, 9 (2).
- [14] Harli Trimulya Suandi As, B., & Zahrotun, L. (2021). Penerapan Data Mining Dalam Mengelompokkan Data Riwayat Akademik Sebelum Kuliah Dan Data Kelulusan Mahasiswa Menggunakan Metode Agglomerative Hierarchical Clustering (Implementation Of Data Mining In Grouping Academic History Data Before Students And Stud. Jurnal Teknologi Informasi, Komputer Dan Aplikasinya, 3(1), 62–71. <http://jtika.if.unram.ac.id/index.php/JTIKA/>.