



Intisari

Student counseling memiliki peran penting dalam keberhasilan akademik siswa, salah satu alasannya adalah *student counseling* dapat membantu siswa mencapai prestasi terbaik dan menyelesaikan masalah yang mengganggu performa akademisnya. Salah satu universitas di Yogyakarta saat ini masih menggunakan metode konvensional dalam proses *student counseling*, dan menemui banyak kendala dalam memahami karakter siswa sehingga berakibat pada banyaknya kasus keterlambatan kelulusan siswa yang pada akhirnya akan merugikan institusi pendidikan tersebut.

Clustering merupakan salah satu tugas utama dalam *data mining* dan telah banyak diaplikasikan di berbagai bidang, salah satunya adalah bidang pendidikan. Penelitian ini mencoba mengimplementasikan teknik *clustering* sebagai salah satu solusi dalam proses *student counseling* dengan menggunakan R. Dua buah algoritme *clustering*, yaitu k-means dan DBSCAN, digunakan untuk mengelompokkan siswa ke dalam beberapa kelompok atau *cluster*. Kemudian tiap *cluster* dianalisis untuk memahami karakteristik umum dari siswa dalam tiap *cluster*. Selanjutnya performa dari kedua algoritme dievaluasi berdasarkan hasil *clustering* dengan menggunakan WSS dan *silhouette width* dan dibandingkan untuk menentukan algoritme yang lebih baik dan cocok digunakan dalam membantu mengenali karakter siswa.

Berdasarkan hasil penelitian yang telah dilakukan ditemukan bahwa algoritme k-means bekerja lebih baik dibanding DBSCAN dalam mengelompokkan dan membantu mengenali karakter siswa yang dibuktikan dengan nilai Total WSS yang lebih tinggi 16,15% dari DBSCAN, namun berdasarkan evaluasi yang dilakukan kedua algoritme tersebut masih belum optimal untuk membantu mengenali karakter siswa karena hasil *clustering* dari kedua algoritme tersebut memiliki nilai rata-rata *silhouette width* yang rendah. Proses *clustering* yang dilakukan juga direpresentasikan pada antarmuka khusus yang dibangun dengan *shiny*, sebuah package dari R. Melalui antarmuka ini, pengguna dapat mencoba melakukan *clustering* menggunakan kedua algoritme yang digunakan dalam penelitian ini dengan data yang diinputkan melalui antarmuka tersebut.

Kata Kunci: pemodelan karakter siswa, *educational data mining*, *clustering*, k-means, DBSCAN.



Abstract

Student counseling has significant role in student's academic achievement, one of the the reason is student counseling help students achieve the highest achievement and help to solve problem interfering with students' academic performance. As of today, one of university in Yogyakarta still uses conventional method in student counseling process, and therefore meet many issues in understanding students' characteristic resulting in many cases of prolonged students' period of study.

Clustering or cluster analysis is one of the main task in data mining and have been applied to many fields, one of them is educational data mining. This research tried to implement clustering technique as an alternative solution in student counseling process with the use of R. Students were grouped into clusters by using 2 clustering algorithms, k-means and DBSCAN. Then each cluster was analyzed in order to understand a general characteristic of students of each clusters. Furthermore, the performance of both algorithms were evaluated by their clustering results using WSS and silhouette width, and compared to determine which of them is more suited to be used to help in understanding students' characteristic.

Based on this research's result, it is found that k-means performed better than DBSCAN in grouping and helping to understand students' characteristic which is proven with higher Total WSS score with 16.15% difference, while the evaluation test shown that both algorithms was not an optimal solution in order to support student counseling process because clustering results of both algorithm have low average silhouette width. The process of clustering was represented in an interface made by using shiny package for R. Through the interface, user can do the clustering using both of the clustering algorithm with their own data.

Keywords : *student's characteristic modeling, educational data mining, clustering, k-means, DBSCAN.*