

Modifikasi algoritma maximum standard deviation reduction dan implementasinya pada graph clustering = On the modification of maximum standard deviation reduction algorithm and its Prastiwiimplementation on graph clustering

Diah Prastiwi, author

Deskripsi Lengkap: <https://lib.ui.ac.id/detail?id=20433791&lokasi=lokal>

Abstrak

Graph clustering adalah pengelompokkan simpul-simpul pada suatu graf menjadi cluster-cluster. Ada berbagai macam algoritma graph clustering yang dapat diterapkan pada graf berbobot, di antaranya adalah algoritma yang berbasis MST (Minimum Spanning Tree). Kelebihan dari algoritma graph clustering yang berbasis MST adalah mampu mendeteksi bentuk cluster yang tidak beraturan. Algoritma graph clustering berbasis MST yang dibahas dalam penelitian ini adalah Maximum Standard Deviation Reduction atau disingkat MSDR. Algoritma tersebut bersifat unsupervised, yang artinya banyaknya cluster tidak ditetapkan oleh pengguna. MSDR memiliki dua tujuan. Tujuan yang pertama adalah membentuk cluster-cluster dimana pada masing-masing cluster, busur-busurnya memiliki nilai bobot yang serupa. Tujuan yang kedua adalah menjaga agar banyaknya cluster yang terbentuk, tidak terlalu banyak. Pada algoritma MSDR, tujuan yang kedua dipenuhi melalui suatu proses yang menggunakan regresi polinomial. Dalam penelitian ini, dilakukan modifikasi MSDR dengan menggantikan proses yang menggunakan regresi tersebut dengan suatu proses baru yang disebut "Loncatan Terbesar", sehingga dihasilkan algoritma baru yang disebut Modifikasi MSDR atau disingkat MMSDR. Dilakukan implementasi MMSDR pada beberapa graf lengkap Euclidean, yaitu graf lengkap yang setiap simpulnya adalah titik di suatu ruang R_n dan setiap busurnya diberi bobot yang berupa jarak Euclidean antara dua titik. Kemudian dilakukan implementasi MMSDR pada graf yang memodelkan rute penerbangan domestik di Indonesia oleh suatu maskapai X.

.....

Graph clustering is the grouping of vertices in a graph into clusters. There are various clustering algorithms for weighted graphs, for example a graph clustering algorithm which is based on MST (minimum spanning tree). The advantage of MST-based clustering is that it allows the detection of clusters with irregular boundaries. An MST-based clustering known as MSDR (Maximum Standard Deviation Reduction) is studied in this research. MSDR is an unsupervised clustering, in which the number of clusters is not dictated by the user. There are two goals of MSDR. The first goal is to produce clusters such that in each cluster the edges have similar weights. The second goal is to prevent the formation of too many clusters. In MSDR, the second goal is met through a process that uses polynomial regression. In this research, MSDR is modified by replacing the process that uses polynomial regression with a new process that is called "largest jump". A new algorithm is produced, which is called Modified MSDR (MMSDR). MMSDR is implemented on several complete Euclidean graphs, where a Euclidean graph is a graph whose vertices are points in a space R_n and the weight of an edge is the distance between the endpoints. MMSDR is also implemented on a graph that models domestic flight routes in Indonesia of an unidentified airline X. As a validation, MMSDR is tested on several Euclidean data, and the result is compared to visually-identified clusters. Afterwards, MMSDR is implemented on a graph that models the domestic flight routes of an airline in Indonesia.