عنوان مقاله:

Implementation of Fuzzy K - means clustering algorithm and comparison of its performance on different bioinformatics (data sets (gene expression

نویسندگان:

Hassan Nosrati Nahook - *Department of Computer Engineering, Higher Education Complex of Saravan, Saravan, Iran*

Shayesteh Tabatabaei - *Department of Computer Engineering, Higher Education Complex of Saravan, Saravan, Iran*

خلاصه مقاله:

Clustering or cluster analysis in statistics and machine learning is one of the branches of uncontrolled learning. In this paper, the implementation and comparison of the performance of K - means fuzzy clustering algorithm on four different bioinformatics datasets using MATLAB software is discussed. Since the performance of each algorithm varies according to the problem we are dealing with, the performance of the algorithm was compared to the biological data set published by KAGGLE. To evaluate this algorithm, the Silhouette Value was used and the best value for K (number of clusters) was obtained. The results show the efficiency of this algorithm for different values of the number of clusters according to the characteristics of each data set to the optimal value according to Specify and use input data.

کلمات کلیدی:

Kmeans, clustering, gene expression, bioinformatics, fuzzy