

Veri Madenciliğinde Kümeleme Teknikleri ile Ortaöğretim Öğrencilerinde Şiddet Eğiliminin Tespiti Üzerine Bir Analiz

Özet

Bu çalışmada, veri madenciliğinde önemli bir yere sahip olan kümeleme teknikleri tanıtılarak; İstanbul'da öğrenim gören 3468 ortaöğretim öğrencisinden elde edilen verilerle, öğrenciler şiddet eğilimlerine göre gruplandırılmaya ve bu eğilime yol açan sebepler ortaya çıkarılmaya çalışılmıştır. Araştırmanın kısıtlarına uyan merkeze dayalı bölümleyici ve hiyerarşik kümeleme tekniklerinden; k-ortalama, iki aşamalı kümeleme ve CLARA yaklaşımları ile elde edilen kümelermeler karşılaştırılarak, bu tekniklerin üstün ve zayıf yönleri incelenmiştir.

Anahtar Kelimeler: Veri madenciliği, kümeleme, k-means, two-step, CLARA

An Analysis on the Determination of the Tendency Toward Violence on Secondary-School Students with Clustering Techniques in Data Mining

Abstract

The selection of the most appropriate clustering technique in data mining depends on the structure of data set and the purpose of the study. Within the scope of this study, the tendency of students towards violence and the reasons causing this tendency were tried to be revealed with the use of data obtained from 3468 students at the secondary schools in Istanbul. Of the clustering techniques that were compatible with the constraints of the study, the ones obtained by the k-means, two-step and CLARA approaches were compared to each other and the strengths and weaknesses of each technique were revealed.

Keywords: Data mining, clustering, k-means, k-medoid, CLARA, two-step

İ. Esen YILDIRIM¹

Yasemin KOLDERE AKIN²

¹ Yrd.Doç.Dr., Marmara Üniversitesi İ.İ.B.F. Ekonometri Böl. İstatistik Anabilim Dalı,

eyildirim@marmara.edu.tr

² Yrd.Doç.Dr., Trakya Üniversitesi İ.İ.B.F. Ekonometri Böl. İstatistik Anabilim Dalı,

yaseminkoldere@trakya.edu.tr