

ارائه یک الگوریتم ترکیبی برای خوشه‌بندی با استفاده از K-means و روش کاهش گرادینان

محسن ضیائی¹، مهدی سیفی²، زینب شکوری³

¹استادیار، دانشکده مهندسی صنایع دانشگاه بجنورد؛ ziaee@iust.ac.ir

²کارشناسی ارشد مهندسی سیستم، دانشکده مهندسی صنایع دانشگاه صنعتی اصفهان؛ mehdesafy@gmail.com

³کارشناسی ارشد مهندسی مالی، دانشکده مهندسی صنایع دانشگاه صنعتی امیرکبیر؛ zeinabshakoori@aut.ac.ir

چکیده

خوشه‌بندی یکی از روش‌های بدون ناظر در تمیز داده‌ها و جزء پرکاربردی از علم داده‌کاوی، پردازش تصویر و شناسایی الگو است. یکی از الگوریتم‌های محبوب خوشه‌بندی، K-means است که دارای معایبی از جمله وابسته بودن به جواب اولیه و همگرا بودن به بهینه محلی، می‌باشد. به منظور برطرف کردن معایب K-means در این مقاله به ارائه یک روش ابتکاری نوین پرداخته می‌شود. این روش از K-means الگوبرداری و با روش کاهش گرادینان تصحیح شده است. از اینرو به نام روش K-Grad نامگذاری شد. روش پیشنهادی K-Grad برخلاف سایر الگوریتم‌های فراابتکاری موجود، فاقد تنظیم پارامتر است و از این حیث در پیاده‌سازی و اجرا، زمانبر نیست. نتایج عددی روی پنج بانک داده معیار، دال بر کارایی چشمگیر روش پیشنهادی K-Grad در سرعت بالای رسیدن به جواب با کیفیت نسبت به سایر روش‌های موجود می‌باشد.

کلمات کلیدی

داده‌کاوی، خوشه‌بندی، K-means، روش کاهش گرادینان.

A Hybrid Clustering Algorithm Based on K-means and Reduced Gradient Method

Mohsen Ziaee, Zeinab Shakoori, Mahdi Seifi

Department of Industrial Engineering University of Bojnord, 94531-55111 Bojnord, Iran

Department of Industrial Engineering, Amirkabir University of Technology, Tehran, Iran, 15875-4413

Department of Industrial Engineering Isfahan University of Technology Isfahan, Iran, 84156-83111

ABSTRACT

Clustering is a fundamental problem that frequently arises in a great variety of application such as data mining, image processing and pattern recognition domain. The K-Means algorithm is one of the most widely used techniques for clustering. However, K-Means has two shortcomings: dependency on the initial state and convergences to local optima. In order to overcome shortcomings of K-means, in this paper proposed a new method that modifies K-means by reduced gradient method, called K-Grad. Our proposed method unlike other existing meta-heuristic algorithms needs any parameter setting, therefore its complexity and execution time is less than meta-heuristic algorithms. Our proposed algorithm is test on several standard data sets from UCI Machine Learning Repository and its performance is compare with other well known. The simulation results are very promising in the terms of quality of solution and convergence speed of algorithm.

KEYWORDS

Data Mining, Clustering, K-means, Reduced Gradient Method.

¹ محسن ضیائی، دانشکده مهندسی صنایع دانشگاه بجنورد، +985832284611