

ارائه یک روش ابتکاری خوشه‌بندی در داده‌کاوی با استفاده از الگوریتم ژنتیک و حل یک مطالعه موردی

واقعی در صنعت حمل و نقل ریلی

مسعود یقینی^{۱۳۹}، سید فرید قنادپور^{۱۴۰}، سعید خدمتلو^{۱۴۱}

چکیده

هدف از انجام این پژوهش ارائه یک روش ابتکاری موثر با استفاده از الگوریتم قدرتمند ژنتیک، برای روش معروف خوشه‌بندی K-Means می‌باشد. روش K-Means به خوشه‌بندی نمونه‌هایی در K دسته می‌پردازد به نحویکه که مجموع فواصل اقلیدسی نمونه‌ها از مراکز دسته‌هایی که به آنها تخصیص یافته‌اند حداقل گردد. یکی از مشکلات این روش این است که در مجموعه‌هایی با حجم انبوه معمولاً با مشکل مواجه است و احتمال دارد که در دام نقاط بهیئ محلی قرار گیرد. بنابراین در این پژوهش سعی می‌گردد که براساس الگوریتم ژنتیک همراه با تعریف اپراتورهای جدید، به طراحی روشی پرداخته شود که به تولید جواب‌هایی با کیفیت بیشتر بپردازد. همچنین در این پژوهش به منظور تعیین مقدار مناسب برای دسته‌ها (K) نیز که یکی دیگر از مشکلات روش K-Means است، معیاری تعریف می‌گردد و به تحلیل نتایج پرداخته می‌شود. شایان ذکر است که از الگوریتم پیشنهادی برای حل یک مطالعه موردی واقعی از کلیه سفرهای مسافری انجام شده در سال 1385 که در راه‌آهن جمهوری اسلامی ایران صورت گرفته و توسط شرکت قطارهای مسافری رجا جایا گردیده است استفاده می‌شود. در این مطالعه موردی به ارزیابی سفرهای مسافری از نظر دو پارامتر نرخ بلیط و ضریب اشغال پرداخته می‌شود و به تحلیل نتایج حاصل شده از الگوریتم پیشنهادی در مقایسه با روش K-Means پرداخته می‌شود. به استناد نتایج حاصل شده، الگوریتم پیشنهادی از نظر کیفیت جواب‌های حاصل شده، نسبت به روش K-Means کاملاً برتر می‌باشد.

کلمات کلیدی

داده‌کاوی، خوشه‌بندی، الگوریتم K-Means، الگوریتم ژنتیک

¹³⁹ استادیار دانشکده مهندسی راه‌آهن، دانشگاه علم و صنعت ایران، yaghini@iust.ac.ir

¹⁴⁰ دانشجوی کارشناسی ارشد دانشگاه علم و صنعت ایران، کارشناس فنی شرکت قطارهای مسافری رجا، ghannadpour@rail.iust.ac.ir

¹⁴¹ دانشجوی کارشناسی ارشد دانشگاه علم و صنعت ایران، کارشناس شرکت قطارهای مسافری رجا، khedmatloo@rail.iust.ac.ir