

عنوان مقاله:

روشی برای خوشه بندی داده های ترکیبی عددی و رسته ای

محل انتشار:

یازدهمین کنفرانس سراسری سیستم های هوشمند (سال: 1391)

تعداد صفحات اصل مقاله: 8

نویسندگان:

فائزه حسینی نژاد - دانشجوی کارشناسی ارشد مهندسی نرم افزار دانشگاه آزاد اسلامی واحد تهران جنوب

مریم خادمی - استادیار گروه ریاضی کاربردی دانشگاه آزاد اسلامی واحد تهران جنوب

میرمحسن پدram - استادیار گروه مهندسی کامپیوتر دانشگاه خوارزمی

خلاصه مقاله:

خوشه بندی یکی از اعمال اصلی در داده کاوی است و به معنای گروه بندی نمونه های مشابه در خوشه ها است. این تکنیک به دنبال کشف ساختار در داده ها از طریق بررسی شباهت ها یا تفاوت های میان آنها است. بسیاری از الگوریتم های خوشه بندی بر روی مجموعه داده های عددی یا رسته ای به خوبی عمل می کنند اما بیشتر آنها بر روی داده های ترکیبی عملکرد ضعیفی دارند و این در حالی است که داده ها در دنیای واقعی ترکیبی از داده های رسته ای و عددی هستند. به همین دلیل توسعه الگوریتم هایی با قابلیت عملکرد بر روی این داده ها از اهمیت به سزایی برخوردار است. در این مقاله روش جدیدی برای خوشه بندی داده های ترکیبی ارائه می دهیم. در روش پیشنهادی ابتدا داده های رسته ای را به عددی تبدیل کرده سپس با استفاده از الگوریتم بهینه سازی cross entropy ، خوشه بندی را بر روی مجموعه داده نهایی که کاملاً عددی است، انجام خواهیم داد. در پایان این روش بر روی دو مجموعه داده استاندارد اعمال شده و کارایی آن با روش k-prototype که یکی از مهمترین الگوریتم ها در زمینه داده های ترکیبی است، مقایسه خواهد شد

کلمات کلیدی:

داده های ترکیبی، ویژگی های رسته ای، الگوریتم بهینه سازی cross entropy

لینک ثابت مقاله در پایگاه سیویلیکا:

<https://civilica.com/doc/214713>

