

عنوان مقاله:

خوشه بندی داده ها، مروری بر روش های موجود و مقایسه عملکرد آن ها

محل انتشار:

کنفرانس بین المللی پژوهش های کاربردی در فناوری اطلاعات، کامپیوتر و مخابرات (سال: 1394)

تعداد صفحات اصل مقاله: 25

نویسندگان:

محمود شاکری - دانشگاه جامع علمی کاربردی واحد جهاد دانشگاهی مشهد

محمد عبدالمهدی - دانشگاه جامع علمی کاربردی واحد جهاد دانشگاهی مشهد

خلاصه مقاله:

یکی از مهمترین اعمال در داده کاوی خوشه بندی داده های موجود در یک دیتاست می باشد. این تکنیک به دنبال کشف ساختارهایی بوده که منجر به گروه بندی نمونه های موجود در یک پایگاه داده بوده، به گونه ای که نمونه های مشابه درون دسته هایی که بیشترین شباهت را با هم داشته قرار گرفته، در حالی که دارای تفاوتی قابل قبول با نمونه های سایر گروه ها داشته باشند. الگوریتم های خوشه بندی را می توان به چند دسته کلی الگوریتم های خوشه بندی تفکیکی، سلسله مراتبی، مبتنی بر چگالی و مبتنی بر گرید تقسیم نمود. این تحقیق ابتدا به مرور روش های خوشه بندی مطرح پرداخته و چند الگوریتم از هر روش را معرفی کرده است. در ادامه همچنین به معرفی روش های دیگری از جمله روش های خوشه بندی مبتنی بر الگوریتم های فرا ابتکاری و خوشه بندی های آنلاین پرداخته شده و سپس چالش های موجود در خوشه بندی از قبیل انتخاب تعداد بهینه خوشه ها، کاهش ابعاد، مدیریت داده های پرت و روشهای مقابله با آنها معرفی شده اند. در نهایت روش ها و الگوریتم های موجود از نظر برخی از پارامترها مقایسه شده و به بررسی مزایا و معایب هر الگوریتم پرداخته شده است.

کلمات کلیدی:

خوشه بندی، روشهای خوشه بندی، مقایسه خوشه ها، معیار بهینگی

لینک ثابت مقاله در پایگاه سیویلیکا:

<https://civilica.com/doc/450800>

