

عنوان مقاله:

بررسی کاربرد الگوریتم K-means در فرآیند خوشه بندی داده های بزرگ

محل انتشار:

اولین کنفرانس ملی مهندسی کامپیوتر و فناوری اطلاعات (سال: 1395)

تعداد صفحات اصل مقاله: 12

نویسنده:

فرزانه بنده پی - دانشجوی کارشناسی ارشد امنیت اطلاعات، دانشگاه مهرآستان

خلاصه مقاله:

عصر داده های بزرگ آغاز شده است. داده ها با سرعت فراوانی از نظر اندازه و تنوع در حال افزایش هستند. با رشد داده ها، چالش ها و تفاوت هایی برای این میزان داده به بار آمده است. داده های بزرگ نشان از ویژگی های متفاوت همچون حجم، تنوع، تغییرپذیری، ارزش، سرعت و پیچیدگی دارند که بر آن اساس آنالیز داده ها و کسب اطلاعاتی با فنون داده کاوی بسیار دشوار است. استخراج داده روشی است برای خارج کردن داده های مفید، اما رویکردهای کلاسیک استخراج داده ها به علت پیچیدگی زیاد نمی توانند به طور مستقیم برای داده های بزرگ مورداستفاده قرار گیرند. خوشه بندی از روش های مهم درداده کاوی است که به دلیل نزدیک بودن با مسائل طبیعی در بسیاری از زمینه ها موردتوجه قرار گرفته است. یکی از مهم ترین الگوریتم های خوشه بندی که به طور گسترده به کار می رود الگوریتم k-means است. الگوریتم فوق کاربردهای بسیاری در زمینه های مختلف علمی و صنعتی دارد. الگوریتم خوشه بندی K-means علی رغم سادگی یک روش پایه برای بسیاری از روش های خوشه بندی دیگر (مانند خوشه بندی فازی) محسوب می شود. این روش روشی انحصاری و مسطح محسوب می شود برای این الگوریتم شکل های مختلفی بیان شده است. ولی همه آن ها دارای روالی تکراری هستند.

کلمات کلیدی:

داده های بزرگ، خوشه بندی، الگوریتم k-means، داده کاوی

لینک ثابت مقاله در پایگاه سیویلیکا:

<https://civilica.com/doc/584671>

