

عنوان مقاله:

مروری بر مقاله برآورد تعداد خوشه ها در مجموعه داده ها از طریق معیار آمار فاصله ای Gap Statistics

محل انتشار:

نخستین کنفرانس بین المللی شهر هوشمند چالش ها و راهبردها (سال: 1398)

تعداد صفحات اصل مقاله: 7

نویسندگان:

زهرا شکیب - دانشجوی ارشد، مهندسی نرم افزار کامپیوتر، موسسه آموزش عالی غیرانتفاعی آیدانا

هاله همایونی - موسسه آموزش عالی غیرانتفاعی آیدانا

کیمیا بازرگان لاری - موسسه آموزش عالی غیرانتفاعی آیدانا

خلاصه مقاله:

این یک مساله مهم و چالش برانگیز در یادگیری بدون نظارت برای تخمین تعداد خوشه ها در مجموعه داده ها است . ما یک روش آمار فاصله ای gap statistics برای تخمین تعداد خوشه ها (گروه ها) در مجموعه ای از داده ها را پیشنهاد می کنیم . این تکنیک از خروجی هر الگوریتم خوشه بندی (به عنوان مثال K - means یا سلسله مراتبی) استفاده می کند، و تغییر در پراکندگی در داخل خوشه با آن انتظار تحت یک توزیع صفر مرجع مناسب را مقایسه می کند . برخی تئوری برای این پیشنهاد توسعه یافته است و یک مطالعه شبیه سازی نشان می دهد که آمار فاصله ای gap statistics معمولا از روش های دیگری که در مقاله پیشنهاد شده اند بهتر عمل می کند

کلمات کلیدی:

خوشه بندی، دسته بندی، سلسله مراتبی، K-means، توزیع یکنواخت

لینک ثابت مقاله در پایگاه سیویلیکا:

<https://civilica.com/doc/998524>

