

عنوان فارسی مقاله :

بهبود دقت و کارایی الگوریتم خوشه بندی K میانگین

عنوان انگلیسی مقاله :

Improving the Accuracy and Efficiency of the k-means Clustering

Algorithm

توجه !

این فایل تنها قسمتی از ترجمه میباشد.



برای تهیه مقاله ترجمه شده کامل با فرمت ورد (قابل ویرایش) همراه با نسخه انگلیسی

مقاله، [اینجا](#) کلیک نمایید.

II. THE K-MEANS CLUSTERING ALGORITHM

This section describes the original k-means clustering algorithm. The idea is to classify a given set of data into k number of disjoint clusters, where the value of k is fixed in advance. The algorithm consists of two separate phases: the first phase is to define k centroids, one for each cluster. The next phase is to take each point belonging to the given data set and associate it to the nearest centroid. Euclidean distance is generally considered to determine the distance between data points and the centroids. When all the points are included in some clusters, the first step is completed and an early grouping is done. At this point we need to recalculate the new centroids, as the inclusion of new points may lead to a change in the cluster centroids. Once we find k new centroids, a new binding is to be created between the same data points and the nearest new centroid, generating a loop. As a result of this loop, the k centroids may change their position in a step by step manner. Eventually, a situation will be reached where the centroids do not move anymore. This signifies the convergence criterion for clustering. Pseudocode for the k-means clustering algorithm is listed as Algorithm 1 [7].

2. الگوریتم خوشه بندی میانگین K

این بخش به توصیف الگوریتم خوشه بندی میانگین k اصلی می پردازد. ایده اصلی رده بندی مجموعه داده های معلوم در k خوشه مجزا می باشد، که مقدار k از قبل ثابت می باشد. این الگوریتم از دو فاز جداگانه تشکیل می شود: فاز اول تعریف k مرکز (مرکز واره)، برای هر خوشه یکی می باشد. فاز بعدی، انتخاب هر نقطه متعلق به مجموعه داده های معلوم و ارتباط بین آن با نزدیک ترین مرکز. عموماً این گونه در نظر گرفته شده است که فاصله اقلیدسی فاصله بین نقاط داده و مراکز را تعیین می کند. به محض حضور کلیه نقاط در برخی از خوشه ها، مرحله اول به پایان رسیده و گروه بندی اولیه انجام شده است. در این نقطه لازم است مراکز جدید مجدداً محاسبه شوند، چرا که حضور نقاط جدید می تواند منجر به تغییر در مراکز خوشه گردد. به محض یافتن k مرکز جدید، پیوند جدیدی بین نقاط داده و نزدیک ترین مرکز جدید ایجاد و بدین طریق یک حلقه تشکیل می شود. به خاطر این حلقه، k مرکز موقعیت خود را به صورت مرحله به مرحله تغییر می دهند. بالاخره، موقعیتی حاصل خواهد شد که مراکز اصلاً حرکت نمی کنند. این مسئله به معیار همگرایی برای خوشه بندی اشاره می کند. شبه برنامه (شبه رمز) برای الگوریتم خوشه بندی میانگین k به شکل الگوریتم 1 لیست شده است.



توجه!

این فایل تنها قسمتی از ترجمه می باشد.

برای تهیه مقاله ترجمه شده کامل با فرمت ورد (قابل ویرایش) همراه با نسخه انگلیسی مقاله، [اینجا](#) کلیک نمایید.

همچنین برای مشاهده سایر مقالات این رشته [اینجا](#) کلیک نمایید.