چکیده

خوشه‌بندی مبتنی بر چگالی یکی از روش‌های مورد توجه در داده‌کاوی و الگوی‌بندی است. DBSCAN و افزایش سعی شده می‌کنند تا داده‌های مناسب و منحصربه‌فردی در سطح کاربرد بزرگ‌تر باشد. این روش اساساً در این روش مبتنی بر چگالی یکی از روش‌های مبتنی بر چگالی است. DBSCAN و ISDBSCAN هم‌اکنون در روش پیشنهادی همانند یک روش دو گروهی در این روش مبتنی بر چگالی یکی از روش‌های مبتنی بر چگالی است. DBSCAN و ISDBSCAN هم‌اکنون در روش پیشنهادی مبتنی بر چگالی است. DBSCAN و ISDBSCAN هم‌اکنون در روش پیشنهادی مبتنی بر چگالی است. DBSCAN و ISDBSCAN هم‌اکنون در روش پیشنهادی مبتنی بر چگالی است. DBSCAN و ISDBSCAN هم‌اکنون در روش پیشنهادی مبتنی بر چگالی است. DBSCAN و ISDBSCAN هم‌اکنون در روش پیشنهادی مبتنی بر چگالی است. DBSCAN و ISDBSCAN هم‌اکنون در روش پیشنهادی مبتنی بر چگالی است. DBSCAN و ISDBSCAN هم‌اکنون در روش پیشنهادی مبتنی بر چگالی است. DBSCAN و ISDBSCAN هم‌اکنون در روش پیشنهادی مبتنی بر چگالی است. DBSCAN و ISDBSCAN هم‌اکنون در روش پیشنهادی مبتنی بر چگالی است. DBSCAN و ISDBSCAN هم‌اکنون در روش پیشنهادی مبتنی بر چگالی است. DBSCAN و ISDBSCAN هم‌اکنون در روش پیشنهادی مبتنی بر چگالی است. DBSCAN و ISDBSCAN هم‌اکنون در روش پیشنهادی مبتنی بر چگالی است. DBSCAN و ISDBSCAN هم‌اکنون در روش پیشنهادی مبتنی بر چگالی است. DBSCAN و ISDBSCAN هم‌اکنون در روش پیشنهادی مبتنی بر چگالی است. DBSCAN و ISDBSCAN هم‌اکنون در روش پیشنهادی مبتنی بر چگالی است. DBSCAN و ISDBSCAN هم‌اکنون در روش پیشنهادی مبتنی بر چگالی است. DBSCAN و ISDBSCAN هم‌اکنون در روش پیشنهادی مبتنی بر چگالی است. DBSCAN و ISDBSCAN هم‌اکنون در روش پیشنهادی مبتنی بر چگالی است. DBSCAN و ISDBSCAN هم‌اکنون در روش پیشنهادی مبتنی بر چگالی است. DBSCAN و ISDBSCAN هم‌اکنون در روش پیشنهادی مبتنی بر چگالی است. DBSCAN و ISDBSCAN هم‌اکنون در روش پیشنهادی مبتنی بر چگالی است. DBSCAN و ISDBSCAN هم‌اکنون در روش پیشنهادی مبتنی بر چگالی است. DBSCAN و ISDBSCAN هم‌اکنون در روش پیشنهادی مبتنی بر چگالی است. DBSCAN و ISDBSCAN هم‌اکنون در روش پیشنهادی مبتنی بر چگالی است. DBSCAN و ISDBSCAN هم‌اکنون در روش پیشنهادی مبتنی بر چگالی است. DBSCAN و ISDBSCAN هم‌اکنون در روش پیشنهادی مبتنی بر چگالی است. DBSCAN و ISDBSCAN هم‌اکنون در روش پیشنهادی مبتنی بر چگالی است. DBSCAN و ISDBSCAN هم‌اکنون در روش پیشنهادی مبتنی بر چگالی است. DBSCAN و ISDBSCAN هم‌اکنون در روش پیشنهادی مبتنی بر چگالی است. DBSCAN و ISDBSCAN هم‌اکنون در روش پیشنهادی مبتنی بر چگالی است. DBSCAN و ISDBSCAN هم‌اکنون در روش پیشنهادی مبتنی بر چگالی است. DBSCAN و ISDBSCAN H

Impovement of density-based clustering algorithm using modifying the density definitions and input parameter

Ali Reza Pahlevanzadeh & Aliakbar Niknafs
Department of Computer Engineering, Shahid Bahonar University of Kerman, Kerman, Iran

Abstract

Clustering is one of the main tasks in data mining, which means grouping similar samples. In general, there is a wide variety of clustering algorithms. One of these categories is density-based clustering. Various algorithms have been proposed for this method; one of the most widely used algorithms is DBSCAN. DBSCAN can identify clusters of different shapes in the dataset and automatically identify the number of clusters. There are advantages and disadvantages in this algorithm. It is difficult to determine the input parameters of this algorithm by the user. Also, this algorithm is unable to detect clusters with different densities in the dataset. ISDBSCAN algorithm is another example of a density-based algorithm that eliminates the disadvantages of the DBSCAN algorithm. ISDBSCAN algorithm reduces the input parameters of DBSCAN algorithm and uses an input parameter k as the nearest neighbor's number. This method is also able to identify different density clusters, but according to the definition of the new core point, it is not able to identify some clusters in a different data set.

This paper presents a method for improving ISDBSCAN algorithm. A proposed approach, such as ISDBSCAN, uses an input parameter k as the number of nearest neighbors and provides a new definition for core point. This method performs clustering in three steps, with the difference that, unlike ISDBSCAN algorithm, it can create a new cluster in the final stage. In the proposed method, a new criterion, such as the number of dataset dimensions used to detect noise in the used data set. Since the determination of the k parameter in the proposed method may be difficult for the user, a new method with genetic algorithm is also

*Corresponding author
موجود در همسایگی آن دسته‌بندی می‌شود.

روش خوشه‌بندی DBSCAN

روش خوشه‌بندی DBSCAN یکی از مهم‌ترین استادی‌های روش‌های درون‌مایه‌ای دیگر خوشه‌بندی‌ها است. روش خوشه‌بندی DBSCAN می‌تواند بتواند این دو چیز را انجام دهد: ۱) این روش می‌تواند روش‌های درون‌مایه‌ای را با توجه به ابعاد گروه‌برداری‌های درون‌مایه‌ای را انجام دهد. ۲) این روش می‌تواند روش‌های درون‌مایه‌ای را با توجه به ابعاد گروه‌برداری‌های درون‌مایه‌ای را انجام دهد.

۱-مقدمه

خوشه‌بندی یکی از مهم‌ترین استادی‌های روش‌های درون‌مایه‌ای است. در کاربردهای برون‌مرزی، روش‌های درون‌مایه‌ای به طور خودکار از داده‌ها برای شناسایی جهت‌های تقویمی استفاده می‌کنند. در اینجا، می‌توان به دو گروه از روش‌های درون‌مایه‌ای، که در جهت‌های تقویمی استفاده می‌شوند، اشاره کرد: ۱) روش‌های درون‌مایه‌ای با توجه به ابعاد گروه‌برداری‌های درون‌مایه‌ای را انجام دهد. ۲) روش‌های درون‌مایه‌ای با توجه به ابعاد گروه‌برداری‌های درون‌مایه‌ای را انجام دهد.

روش‌های خوشه‌بندی معمول همچنین می‌توانند به صورت خودکار اجرا شوند. این روش‌ها برای یافتن نقاط اصلی در داده‌ها، از طرف دیگر، به صورت خودکار اجرا شوند. این روش‌ها برای یافتن نقاط اصلی در داده‌ها، از طرف دیگر، به صورت خودکار اجرا شوند.

روش‌های خوشه‌بندی DBSCAN

روش خوشه‌بندی DBSCAN یکی از مهم‌ترین استادی‌های روش‌های درون‌مایه‌ای است. روش خوشه‌بندی DBSCAN می‌تواند بتواند این دو چیز را انجام دهد: ۱) این روش می‌تواند روش‌های درون‌مایه‌ای را با توجه به ابعاد گروه‌برداری‌های درون‌مایه‌ای را انجام دهد. ۲) این روش می‌تواند روش‌های درون‌مایه‌ای را با توجه به ابعاد گروه‌برداری‌های درون‌مایه‌ای را انجام دهد.

۱۲-ISB-DBSCAN

ISB-DBSCAN می‌تواند به صورت خودکار اجرا شود. این روش می‌تواند با توجه به ابعاد گروه‌برداری‌های درون‌مایه‌ای، این روش می‌تواند با توجه به ابعاد گروه‌برداری‌های درون‌مایه‌ای، این روش می‌تواند با توجه به ابعاد گروه‌برداری‌های درون‌مایه‌ای، این روش می‌تواند با توجه به ابعاد گروه‌برداری‌های درون‌مایه‌ای، این روش می‌تواند با توجه به ابعاد گروه‌برداری‌های درون‌مایه‌ای، این روش می‌تواند با توجه به ابعاد گروه‌برداری‌های درون‌مایه‌ای، این روش می‌تواند با توجه به ابعاد گروه‌برداری‌های درون‌مایه‌ای، این روش می‌تواند با توجه به ابعاد گروه‌برداری‌های درون‌مایه‌ای، این روش می‌تواند با توجه به ابعاد گروه‌برداری‌های درون‌مایه‌ای، این روش می‌تواند با توجه به ابعاد گروه‌برداری‌های درون‌مایه‌ای، این روش می‌تواند با توجه به ابعاد گروه‌برداری‌های درون‌مایه‌ای، این روش می‌تواند با توجه به ابعاد گروه‌برداری‌های درون‌مایه‌ای، این روش می‌تواند با توجه به ابعاد گروه‌برداری‌های درون‌مایه‌ای، این روش می‌تواند با توجه به ابعاد گروه‌برداری‌های درون‌مایه‌ای، این روش می‌تواند با توجه به ابعاد گروه‌برداری‌های درون‌مایه‌ای، این روش می‌تواند با توجه به ابعاد گروه‌برداری‌های درون‌مایه‌ای، این روش می‌تواند با توجه به ابعاد گروه‌برداری‌های درون‌مایه‌ای، این روش می‌تواند با توجه به ابعاد گروه‌برداری‌های درون‌مایه‌ای، این روش می‌تواند با توجه به ابعاد گروه‌برداری‌های درون‌مایه‌ای، این روش می‌تواند با توجه به ابعاد گروه‌برداری‌های درون‌مایه‌ای، این روش می‌تواند با توجه به ابعاد گروه‌برداری‌های درون‌مایه‌ای، این روش می‌تواند با توجه به ابعاد گروگردان این روش یازدهمین سال ۱۴۹۸ شماره ۱۰ مه ۱۳۹۸
که تعداد پایین‌تریک که انتخاب پرآمرت پم DBSCAN به‌طور معمول خیلی ساده‌تر از شرایط هم‌سایگی است، به‌طور دموی رنگ‌ربندی کاربردی کاربردی می‌داند که نیوکتیتی نسبت به داده‌ها بیشتر است.

در [2] یک گروهی از خوشه‌بندی بدون پارامتر پیشنهاد شده است که که DBSCAN شرایط چگالی‌های محلی مجموعه‌داده با ایجاد یک فهرست چگالی نمایش داده استفاده می‌کند. بعد از این تحقیق، نرمال از پارامترهای تراکم در فهرست چگالی نمایش داده شده است. همچنین، ایجاد یک مجموعه جواب‌های خوشه‌بندی ایجاد شدگان، به‌طور ترتیب نتایج خوشه‌بندی محلی و دارای پارامترها مختلف از فهرست نرمال چگالی به‌دست می‌آید، مثلاً، سیس یک قانون بر خودش در پارامترها محلی و نتایج به‌دست‌آمده از اجرای DBSCAN توصیه می‌شود. در [8] یک کار، دست‌آمده از آزمایش‌هایی که می‌توانند به استفاده از DBSCAN بیش‌ترین مزیت را دارند.

در [15] یک گروهی جدید خوشه‌بندی معرفی شده است که داده‌ها را با تغییر چگالی‌های متفاوت در محاوره گروه‌بندی می‌کند. این گروهی جدید خوشه‌بندی می‌کند. این اعداد نداشته‌ای قابل تغییر در طول زمان و داده‌های نهایی به‌دست می‌آید. در [4] یک نظریه نویسی گروهی جدید DBSCAN است. در این کار، آن خوشه‌بندی با اینکه برای اولین بار، Eps یک مقدار متفاوت از DBSCAN می‌شود. به‌طور تزیین نیاز می‌دارند. 

در [16] توصیه‌هایی از DBSCAN در بیش‌ترین مزیت بر روی یک از مشکلات عمده در انتخاب پرآمرت تشخیص هم‌سایگی ایجاد شده است. به‌طور دموی رنگ‌ربندی کاربردی کاربردی می‌داند که نیوکتیتی نسبت به داده‌ها بیشتر است.

در [8] یک روش برای پیشنهاد گروهی ایجاد شده است. برای تعیین مقدار خوشه‌بندی هر‌کدام از DBSCAN می‌شود.

[1] Core point
[2] Border point
[3] Noise point
تعداد داده‌هایی که یک گراف خوشه‌بندی شوند، به چندین تجزیه تقسیم می‌شوند و \( \text{Eps} \) و \( \text{MinPts} \) می‌شود. بعد از آن \( \text{Eps} \) و \( \text{MinPts} \) شده و بنا برای آن داده‌ها خوشه‌بندی می‌شوند.

### ISB-DBSCAN برugar algoritmden

در این بخش از مقاله، مفاهیم اولیه در زمینه الگوریتم‌های خوشه‌بندی مبتنی بر چگالی، تعریف و توضیحی در خصوص الگوریتم DBSCAN، \( \text{MinPts} \) و همچنین درخت پوشاک کمیته‌ی ایران همکاری داده‌ها است.

#### DBSCAN

در الگوریتم DBSCAN، خوشه‌بندی مبتنی بر چگالی، خوشه‌بندی با استفاده از تراکم داده‌ها در جمع‌آوری داده‌های موجود در همسایگی \( \text{MinPts} \) می‌شود. الگوریتم DBSCAN دارد و یکی از تست‌های مناسب‌تر است. انرژی DBSCAN برای تعریف می‌شود.

1. \( \text{MinPts} \): کمینه تعداد داده‌های موجود در اطراف یک داده را مشخص می‌کند.
2. \( \text{Eps} \): شعاع همسایگی را مشخص می‌کند.

برای یک‌پارامتری روش‌های خوشه‌بندی مبتنی بر چگالی لازم است تا این اصلاحات تعریف شوند. تعریف 1: \( r \) نقطه‌ی هسته یک همسایگی و \( q \) شعاع همسایگی برای نقطه‌ی \( r \) در نظر گرفته شود. اگر همسایگی به شعاع \( r \) برای نقطه‌ی \( r \) هم‌مرکز باشد، اگر \( r \) باشد (شکل 1).

\[ N = q \text{ in data set } D \text{ such that dist}(p,q) < \epsilon \]

(شکل 1): The density of local points at a point \( p \) to radius \( \epsilon \).

تعاریف 2: داده‌ها را در دسترس مستقیم چگالی \( q \) را درون یک همسایگی به شعاع \( r \) با هسته‌ی \( q \) پاشد (شکل 2).

سر برای (3) EDBSCAN مبتنی بر جوشنده و \( \text{Eps} \) و \( \text{MinPts} \) می‌شود. این روش در این روش می‌تواند به‌طور خودکار تعریف می‌شود.

---

1. Histogram equalization
2. Dominant sets
3. Binary Differential Evolution DBSCAN
4. Efficient density-based clustering algorithm
ISB-DBSCAN

۲-۳- الگوریتم DBSCAN
همانطور که در节یل نیز بیان شد، الگوریتم DBSCAN نیاز به دو پارامتر ورودی برای خوش‌نمی‌دارد که این آنها توسط کاربر مشکل است. الگوریتم DBSCAN نیز نیاز به یک پارامتر ورودی به عنوان k ندیده‌ترین همسایه‌دار این روش می‌شود. یک گوگلی است. یا رازیه‌دار DBSCAN از مراپی برخورد است. تعبیه یک پارامتر k به‌طورعموم ساده‌تر از پارامترهای همچنین این روش در تعبیه خودکاری با چگالی‌های مختلف از نوع مجموعه داده بهتر از DBSCAN عمل می‌کند. یکی از دلایل این روش در مباسه‌ای با تشخیص داده‌ای به داده نهایی است. برای اینکه این روش نیاز است این اصولاتی تعریف شوند:

تعریف ۵: به ارایه هر k، فاصله‌ای \(x, y \in D\) یافت k شماره داده می‌شود.

\[ k = \text{NN}(p) \]

تعریف ۷: مجموعه‌ای همسایه‌دار p در مجموعه داده D به‌صورت \(k\)-ندیده‌ترین همسایه‌دار می‌شود. مجموعه \(k\)-ندیده‌ترین همسایه‌دار \(p\) به‌صورت \(p, x \in D \mid k \leq \text{NN}(p) \leq k\) می‌شود.

\[ \text{NN}(p) = \{Q \in D \mid q \leq \text{NN}(p) \} \]

تعریف ۸: بی‌صرفی زیر تعریف ۹ همسایگی نقطه p به‌صورت زیر تعریف می‌شود:

\[ q \in D \cap \text{IS}(p) \]

تعریف ۱۰: اگر عضوی داده در این بزرگتر از 2\(k/3\) باشد، آن‌گاه آن داده به خوبی داده شده‌است. می‌شود.

\[ q \geq k \]

تعریف ۱۱: داده در مجموعه داده را در دسترس مستقیم چگالی داده q قرار داده.

\[ S(q) \]

تعریف ۱۲: اگر داده p با توجه به k در دسترس چگالی داده q قرار داده، اگر مجموعه‌ای از زنجیره داده‌ها مانند \(p_1, p_2, \ldots, p_n\) وجود داشته باشد که، \(p_1, p_2, \ldots, p_n\) به‌طور که از دسترس مستقیم چگالی p با توجه به k بوده و همچنین

\[ q \geq \text{NN}(p) \]

کاربرد الگوریتم DBSCAN در کشف خوش‌نمی‌دار

به‌منظور در مجموعه داده‌هایی با داده‌های نوپاورد است. پس از بررسی هسته‌های داده‌های خودکاری، نمونه‌های رابطه آنتگن‌شده شده و یک خوشه‌ای جدید تشکیل می‌دهند. داده‌های مرزی به خوشه‌ای تعلق می‌گیرد که به نمایندگی آن ندیده‌تر باشند و داده‌های پرت نیز حذف خواهند شد.

1. Influence space

سال ۱۳۹۸ شماره ۲ پاییز ۴۰
الگوریتم DBSCAN

ISDBSCAN یکی از الگوریتم‌های جدیدی است که برای یافتن گروه‌های داده‌های کوچک و اتفاقی با استفاده از مقدار اثرات نزدیک به منظور انتخاب گروه‌های مستقل و تجزیه و تحلیل آنها استفاده می‌شود. این الگوریتم با استفاده از مقدار اثرات نزدیک به منظور انتخاب گروه‌های مستقل و تجزیه و تحلیل آنها استفاده می‌شود.

 Require: D={(x₁,x₂,…,xₙ)}: the dataset.
 K: the number of neighbors.
 Ensure: C= {C₁,C₂,…,Cₘ}: set of clusters.
 1: function ISDBSCAN D,k
 2: ClusterID=1
 3: mark all points xₑD as "UNCLASSIFIED"
 4: calculate the influence space ISₗ(xₑ) of each point xₑD
 5: for all xₑD do
 6: if xₑ is marked as "UNCLASSIFIED" then
 7: expandCoreCluster xₑ, ClusterID then
 8: ClusterID++
 9: end if
 10: end if
 11: end for
 12: for all xₑD do
 13: if xₑ is marked as "UNCLASSIFIED" then
 14: search each point yₑISₗ(xₑ) in the influence space of xₑ
 15: if points in ISₗ(xₑ) are all marked as "NOISE" or there is no point in ISₗ(xₑ) then
 16: xₑ is labeled as "NOISE"
 17: else
 18: mark xₑ as ClusterID of the closest core point in ISₗ(xₑ)
 19: end if
 20: end if
 21: end for
 22: end function
 23: function ExpandCoreCluster (xₑ,ClusterID)
 24: SeedList= ISₗ(xₑ)
 25: if |SeedList| ≥ 2kₙ then
 26: xₑ is labeled as ClusterID
 27: else
 28: false
 29: end if
 30: for all yₑSeedList do
 31: if |ISₗ(yₑ)| > 2kₙ then
 32: yₑ is labeled as ClusterID
 33: for all zₑ|ISₗ(yₑ)| do
 34: if zₑ is labeled as "UNCLASSIFIED" or zₑ is labeled as "NOISE" then
 35: if zₑ is not in SeedList then
 36: add zₑ into SeedList
 37: end if
 38: end if
 39: end if
 40: end for
 41: end if
 42: end for
 43: true
 44: end function

(1)[ISDBSCAN (6): ISDBSCAN Clustering Algorithm [1]]

(2)[ISDBSCAN (5): ISDBSCAN Clustering Algorithm [2]]

لری 1398 شماره 2 یوپیم 1400


درخت پوشاک کمپنی 1 از نظر فراکسیون کمپنی در گراف‌های ارتباطی (ورژن‌های ساخته شده) [7] منظور از یک درخت پوشاک در دامنه است که توسط شرکت انتخاب شده است. گراف‌ها یا گروه‌های چندین دامنه هستند که به صورت غیرمستقیم و به صورت مستقیم با هم کنار هم قرار گرفته‌اند. این درخت پوشاک کمپنی در مرحله بعدی از درخت پوشاک کمپنی برای ترکیب دادن به‌طور مستقیم می‌تواند از استفاده در مقایسه با الگوریتم‌های متوازی استفاده کرده باشد. این الگوریتم‌ها به صورت خلاصه باید مورد استفاده قرار گیرند.

1. مینیمال سپینینگ تری [3]

2. مدل‌سازی دنیهی DBSCAN

3. الگوریتم پیشنهادی DBSCAN
محارب یک گل‌کار قادر داریم، به‌خوبی از یکدیگر چیدمان.

بعد از پیداکردن تمام داده‌هایی که در بعضی مواقع ممکن است، داده‌های جدید به‌دست آمده که هم‌چنان وضع هیچ خوش‌خوانی قرار نگرفته باشند. برای جلوگیری از این وضع، روایت‌های جدید و جدید‌تر در هر مورد، مطابق با محدوده‌های مربوط به محدوده‌های مربوط به محدوده‌های مربوط به محدوده‌های مربوط به محدوده‌های مربوط به محدوده‌های مربوط به محدوده‌های مربوط به محدوده‌های مربوط به

-۱- راهکاری: ابتدا داده‌های داده‌هایی که در این مجموعه موجود نبوده و به‌جا مانده، با هر دو مورد، بایستی تجزیه و تحلیل شوند. این روش در این حالت، ممکن است بهترین راه حل به‌توام مورد استفاده قرار گیرد.

-۲- راهکاری: ابتدا داده‌های موجود ذخیره گردد و این داده‌ها به‌طور جداگانه تجزیه و تحلیل شوند. این روش ممکن است بهترین راه حل به‌توام مورد استفاده قرار گیرد.

-۳- راهکاری: ابتدا داده‌های موجود ذخیره گردد و این داده‌ها به‌طور جداگانه تجزیه و تحلیل شوند. این روش ممکن است بهترین راه حل به‌توام مورد استفاده قرار گیرد.

-۴- راهکاری: ابتدا داده‌های موجود ذخیره گردد و این داده‌ها به‌طور جداگانه تجزیه و تحلیل شوند. این روش ممکن است بهترین راه حل به‌توام مورد استفاده قرار گیرد.
خرچنگی الگوریتم تعداد خوشه‌بندی هدست آمده‌ای توجه به عدد کا وردی است. آگر مقدار \( k \) بیشتری توسط کنترل انتخاب شود، این روش خوشه‌بندی می‌تواند در اولویت مجموعه‌ای وارد واقع شود و خوشه‌بندی را به راه حل صحیح از آن در به‌کار رفتن الگوریتم پیشنهادی ارائه شده است.

\[ \text{شکل (7): الگوریتم خوشه‌بندی } \]

(Figure-7): MDD-ISB-DBSCAN Clustering Algorithm

پیشنهادی

همان‌طور که در مبحث قبلاً نیز بیان شد، الگوریتم MDD-ISB-DBSCAN به به یک پارامتر \( k \) که یک مقدار مثبت دارد، پارامتر \( k \) می‌تواند به جای نام بایستنی ورودی DBSCAN استفاده شده و خوشه‌بندی دقیق‌تری را تولید کند. در واقع، با استفاده از این پارامتر و محاسبه مقدار IS\(_k\) می‌توان این کمک کرد که داده‌ها به دست این استخراج این پارامتر نسب به ورودی DBSCAN های الگوریتم مانندی باید از پارامترها است. مشکل اساسی زمانی است که کارکردهای ضمنی از مجموعه داده نشان خواهد داد که به آن هیچ دانش قبلی ندارد: پس انتخاب این پارامتر می‌تواند کار مشکلی باشد.

برای حل مشکل بالا یک روش پیشنهادی به نام PMDD-ISB-DBSCAN در این مسئله ارائه می‌شود که می‌تواند مقدار خوشه‌بندی را به اطمینان بیشتری تعمیم کند. برای این کار از الگوریتم تنظیم استفاده می‌شود تا با تایب

\[ \text{Require: } d = \{x_1, x_2, \ldots, x_n\} : \text{the dataset.} \]

\[ K: \text{the number of neighbors.} \]

\[ \text{Ensure: } C = \{C_1, C_2, \ldots, C_k\} : \text{set of clusters.} \]

1. function MDD-ISB-DBSCAN(d,k)
2. ClusterID=1
3. mark all points \( x_i \in D \) as -1
4. calculate the influence space \( IS_k(x_i) \) of each point \( x_i \in D \)
5. for all \( x_i \in D \) do
6. if \( x_i \) is marked as -1 then
7. expandCoreCluster(d,ClusterID,IS\(_k\))
8. ClusterID++
9. end
10. end
11. if there was a point with -1 label then
12. p=Sort points with IS\(_k\) length in dataset
13. for all \( p \) in dataset
14. select the largest length of IS\(_k\); \( p \)
15. BorderNoiseOrNewCluster(d,ClusterID,IS\(_k\),l);
16. end
17. end
18. end
19. function BorderNoiseOrNewCluster(d,ClusterID,IS\(_k\),l)
20. seedList=IS\(_k\);
21. if there was a label in the neighborhood of the seedList for the selected point (other than noise and -1) then
22. If length(seedList)> Dataset Dimensions then
23. d(i)=neighborhood label;
24. else
25. d(i)=noise;
26. end
27. return d,false;
28. end
29. if there was more than one labels in the neighborhood of the seedList for the selected point (other than noise and -1) then
30. If length(seedList)> Dataset Dimensions then
31. d(i)=label of nearest core point;
32. else
33. d(i)=noise;
34. end
35. return d,false;
36. end
37. if there is no label (label -1) then
38. expandNewCluster(d,ClusterID,1) then
39. ClusterID=ClusterID+1;
40. end
41. end
42. return d,ClusterID;
43. end
44. function expandNewCluster(d,ClusterID,1)
45. seedList=IS\(_k\);
46. if length(seedList)> Dataset Dimensions then
47. result=true;
48. d(i)=ClusterID;
برازش پیشنهادی بهینه‌ترین مقدار $k$ را تخمین بزنید. پارامترهای ورودی الگوریتم زنتیک طبق جدول (1) تعیین می‌شوند.

<table>
<thead>
<tr>
<th>جدول (1): پارامترهای ورودی الگوریتم زنتیک</th>
</tr>
</thead>
<tbody>
<tr>
<td>پارامتر ورودی</td>
</tr>
<tr>
<td>تعداد جمعیت اولیه</td>
</tr>
<tr>
<td>تغییر ترکیب</td>
</tr>
<tr>
<td>تغییر جهش</td>
</tr>
<tr>
<td>تعداد بازگشت</td>
</tr>
<tr>
<td>تعداد متغیرهای کروماتوم</td>
</tr>
</tbody>
</table>

برای تعیین خوشه‌کار پارامتر ورودی $k$ در شکل (8) رسم شده است.

$$f(x) = \sum_{i=0}^{M} (\text{mst}(C_i)) - N, \quad M = 1$$

$mst(C_i)$ تعداد خوشه، $C_i$ داده‌های آماده خوشه، $N$ تعداد داده‌های نهایی در طبق فرمول (2) تعیین می‌شود.

$$mst(C_i) \leq B(C_i) \quad \text{و} \quad B(C_i) = \text{مجموعه داده و}$$

$$M = \begin{cases} 0, & 2 \times (\text{Avg}(C_i) + SD(C_i)) \leq B(C_i) \\ 1, & 2 \times (\text{Avg}(C_i) + SD(C_i)) > B(C_i) \end{cases}$$

$\text{Avg}(C_i)$ میانگین پایین‌ترین $C_i$ و $\text{SD}(C_i)$ انحراف معیار پایین‌ترین $C_i$。

برازش $k$ برای اینکه $M$ را بهینه‌ترین مقدار $k$ بنا به پارامتر پیشنهادی مصرف می‌شود.

با تعداد داده‌های جمعیت داده تخمین بر (2) انجام می‌شود.

یکی از مهم‌ترین قسمت‌ها در الگوریتم زنتیک تعريف برای پروژه بهره‌مند سیستم مورد نظر است. استفاده از میانگین ارزیابی خوشه‌بندی بعنوان فاز پرواز، می‌تواند به‌خوبی در خوشه‌های کروی جواب مورد نظر را پدیدار کند در خوشه‌بندی مبتنی بر چگالی نزدیک آن خوشه‌ها بانگ‌سلام کروی سیکل تیز است. لذا در پیش‌باز همین مقدار استفاده باشد تا میزان تکیه از میانگین دقیق نباید.

پردازش پیشنهادی بهینه‌ترین مقدار $k$ را تخمین بزنید. پارامترهای ورودی الگوریتم زنتیک طبق جدول (1) تعیین می‌شوند.
در این بخش، الگوریتم DBSCAN تخمین پارامتر کلیک کار از پیش‌بینی می‌شود. الگوریتم DBSCAN می‌تواند از مجموعه داده‌های متنوعی استفاده کند و در اینجا به‌عنوان یک مثال به‌عنوان یک شرکت در مجموعه متنوعی استفاده کرده است. الگوریتم DBSCAN مجموعه‌های بزرگی از این مشکل وجود یک چکیده می‌کند تعداد خوشه‌ها را تعیین کند. در اینجا، الگوریتم PMDD-IBS-DBSCAN فرض بر این گرفته شده است که تعداد خوشه‌ها در مجموعه داده‌ها ۲ باشد.

روش ارائه‌شده از نظر زمان اجرای یک بررسی قرار گرفته است. در این مورد، الگوریتم DBSCAN از عناصر خوشه‌بندی به‌عنوان یک بررسی کیفی خوشه‌بندی است. در این مورد، الگوریتم DBSCAN از یک بررسی مستقیم برای مقایسه امکان‌پذیری اقتدار الگوریتم و مقایسه زمان حساب نمایان یا دقیق‌تر و در نهایت نتیجه‌گیری می‌شود و سال‌ها مورد استفاده قرار می‌گیرد.

نتایج ارزیابی

در این بخش، الگوریتم PMDD-IBS-DBSCAN به‌عنوان یک پیش‌بینی می‌شود. در این مورد، الگوریتم PMDD-IBS-DBSCAN از میان نتایج تایپ شده است. هر دو الگوریتم در نرم‌افزار مربوط به‌عنوان پیش‌بینی می‌شود. در این مورد، الگوریتم PMDD-IBS-DBSCAN می‌تواند از عناصر خوشه‌بندی به‌عنوان یک بررسی کیفی خوشه‌بندی است. در این مورد، الگوریتم PMDD-IBS-DBSCAN می‌تواند از یک بررسی مستقیم برای مقایسه امکان‌پذیری اقتدار الگوریتم و مقایسه زمان حساب نمایان یا دقیق‌تر و در نهایت نتیجه‌گیری می‌شود و سال‌ها مورد استفاده قرار می‌گیرد.

پنج‌تایی ارزیابی

در این بخش، الگوریتم PMDD-IBS-DBSCAN به‌عنوان یک پیش‌بینی می‌شود. در این مورد، الگوریتم PMDD-IBS-DBSCAN از میان نتایج تایپ شده است. هر دو الگوریتم در نرم‌افزار مربوط به‌عنوان پیش‌بینی می‌شود. در این مورد، الگوریتم PMDD-IBS-DBSCAN می‌تواند از عناصر خوشه‌بندی به‌عنوان یک بررسی کیفی خوشه‌بندی است. در این مورد، الگوریتم PMDD-IBS-DBSCAN می‌تواند از یک بررسی مستقیم برای مقایسه امکان‌پذیری اقتدار الگوریتم و مقایسه زمان حساب نمایان یا دقیق‌تر و در نهایت نتیجه‌گیری می‌شود و سال‌ها مورد استفاده قرار می‌گیرد.
پیشرفت که های ورودی نویس کاربر، تمام داده‌ها را شامل می‌شود پس قدر به تجزیه صحیح این داده‌ها نیست. در مجموعه‌های جهانی این مورد بطور ممکن بهتر عمل می‌کند.

روش‌های موجود در DBSCAN در مجموعه‌های داده MDD-ISB-DBSCAN

که دارای خوشه‌های تجزیه‌بندی به همبود و همبود داده‌ها در اکثر فرآیند بایان می‌شود، عملیات خوشه‌بندی را برای تخصیص داده‌ها در مجموعه‌های داده DBSCAN است. مجموعه‌های داده‌ها در مجموعه‌های مختلف می‌توانند در میان DBSCAN و MDD-ISB-DBSCAN تغییری در میان DBSCAN و MDD-ISB-DBSCAN داشته باشند. روش DBSCAN در مجموعه‌های داده‌ها می‌تواند در مقایسه با دیگر

روش‌های موجود نیز بهتر عمل کرده است.

(جدول 3) نرخ صحیح بررسی داده‌ها در MDD-ISB-DBSCAN

<table>
<thead>
<tr>
<th>MDD-ISB-DBSCAN</th>
<th>IS-DBSCAN</th>
<th>ISB-DBSCAN</th>
<th>MDD-ISB-DBSCAN and the MDD-ISB-DBSCAN method</th>
</tr>
</thead>
<tbody>
<tr>
<td>99.58</td>
<td>63.33</td>
<td>64.58</td>
<td>Flame</td>
</tr>
<tr>
<td>100</td>
<td>100</td>
<td>100</td>
<td>Jain</td>
</tr>
<tr>
<td>99.66</td>
<td>99</td>
<td>99.66</td>
<td>R15</td>
</tr>
<tr>
<td>96.93</td>
<td>78</td>
<td>80</td>
<td>D31</td>
</tr>
<tr>
<td>100</td>
<td>100</td>
<td>100</td>
<td>Spiral</td>
</tr>
<tr>
<td>94.33</td>
<td>83</td>
<td>83.43</td>
<td>Pathbiased</td>
</tr>
<tr>
<td>94.98</td>
<td>89.47</td>
<td>94.98</td>
<td>Compound</td>
</tr>
<tr>
<td>99.87</td>
<td>95.68</td>
<td>95.30</td>
<td>Aggregation</td>
</tr>
<tr>
<td>92.66</td>
<td>66.66</td>
<td>99.78</td>
<td>Iris</td>
</tr>
<tr>
<td>69.66</td>
<td>55.05</td>
<td>55.61</td>
<td>Wine</td>
</tr>
<tr>
<td>61.48</td>
<td>55.55</td>
<td>55.55</td>
<td>Heart</td>
</tr>
<tr>
<td>91.74</td>
<td>80.74</td>
<td>82.45</td>
<td>Mianakин</td>
</tr>
</tbody>
</table>
Figure 9: Comparison of different methods for the correct identification of clusters.
(Figure 10): Convergence diagrams of different dataset in the genetic algorithm
6-نتیجه‌گیری

در این مقاله یک روش خوشه‌بندی مبتنی بر ISBD-DBSCAN در اثر منعطفه‌گرایی و سطح خوردنی بود. این روش مبتنی بر اعمال همبستگی در سطح مرحله‌بندی می‌باشد. در مرحله نخست برای تمام داده‌ها فضای نهایی یا ISKX محاسبه می‌شود. در مرحله دوم تمامی داده‌های استثنایی که خارجی اصلی را شکل دهند، ایجاد می‌شوند. در مرحله سوم داده‌ها که بر پایه 1-دارند براساس فهرست همبستگی در ISKX خوشه‌بندی می‌شود. یکی از تقارنهای این روش با الگوریتم ISBD-DBSCAN خوشه جدید در مرحله سوم همچنین تعیین جدیدی از داده‌های است. در سمت دوم دو روش بیشنهاید برای تنظیم خودکار پرامتر ورودی ارائه شد. این روش می‌تواند به طور خودگر با رأسم مجموعه داده بهترین پرامتر ورودی را با الگوریتم ZNN و تابع پراش بیشینه تشخیص دهد. یکی از مسئله‌ها است که از آنجا که ممکن است، کاربر در تعیین پرامتر تغییر مشکل شود. می‌تواند مورد واقع شود. یکی از دانست به روش خوشه‌بندی مورد قبولی را ارائه کند.

7- منابع

عالیضا بهلولزاده، فارغ التحصیل کارشناسی ارشد رشته هواشناسی دانشگاه شهید باهنر کرمان است. زمینه پژوهشی مورد علاقه ایشان داده کاوی است.
نامی رایانه ایشان عبارت است از:
pahlevanzadeh@eng.uk.ac.ir

علي اکبر نیک نیک، عضو هیأت علمی و دانشیار بخش مهندسی کامپیوتر دانشگاه شهید باهنر کرمان، تحصیلات خود را در دانشگاه شیراز، تربیت مدرس و شهری باهنر کرمان به پایان رسانده و در سال‌های اخیر در حوزه‌های پژوهشی مرتبط با داده‌کاوی و هوش تجاری فعالیت می‌کند.
نامی رایانه ایشان عبارت است از:
nikaafs@uk.ac.ir