

Unsupervised competitive learning NN [2] :

یادگیری بدون ناظر رقابتی شبکه عصبی یک رویه ی نورون برنده است که هر نورون برای پاسخ صحیح برای یک زیرمجموعه از داده های ورودی مبارزه می کند. این رویه برای حذف افزونگی از داده های بدون ساختار است. دو روش اصلی یادگیری بدون ناظر رقابتی شبکه عصبی عبارتند از نگاشت خود سازمان و نظریه تشدید تطبیق شبکه عصبی می باشند.

SOM: نگاشت خودسازمان که بعنوان نگاشت کوهنن نیز شناخته می شود یک کلاس از شبکه های عصبی است که از مفهوم یادگیری رقابتی استفاده می کند به گونه ای که نرون های خروجی با یکدیگر رقابت می کنند تا به مقادیر واقعی خروجی دست پیدا کنند، نتیجه فقط یک نرون یا گروهی از نرون ها بنام نرون های برنده را داراست. که با استفاده از اتصالات بازدارنده/مهار افقی[[1]](#footnote-1)( مسیرهای بازخورد منفی) بین نرون ها محقق می شود. در این راستا شبکه نرون های برنده را در چندین چرخه مشخص می کند: متعاقباً ، مجبور می شود براساس توزیع داده های ورودی ، سازماندهی مجدد خود را انجام دهد (از این رو به آنها نقشه های خود سازمان دهی می گویند آنها در ابتدا توسط مغز انسان الهام گرفته شده بودند ، که دارای مناطق تخصصی است که در آن ورودی های حسی مختلف از طریق نگاشت محاسباتی که از نظر توپولوژیکی مرتباً تجدید می شوند / پردازش می شوند. در SOM ، نورونها بر روی رئوسهای یک شبکه (معمولاً یک یا دو بعد) تنظیم می شوند. این شبکه مجبور است داد های با ابعاد بالا را با حفظ ویژگی های توپولوژیکی داده ورودی با استفاده از تابع همسایگی، در ابعاد پایین تر نشان دهد، تابع همسایگی که ورودی ها را به فضای توپولوژیکی ای تبدیل می کند که موقعیت های نرون در فضا نشان دهنده ی نمایش ویژگی های آماری ذاتی هستند که در مورد SOM های ذاتا غیرخطی به ما می گویند.

آموزش یک شبکه شامل SOM اساسا یک پروسه سه مرحله ای بعد از مقداردهی اولیه تصادفی اتصالات وزن دار است. این سه گام شامل:1. Competition : هر نرون در شبکه مقدار خود را با استفاده از یک تابع جداساز(discriminant function) محاسبه می کند که پایه ی رقابت میان نرون هاست. نرون با بزرگترین مقدار گسسته در گروه رقابتی، برنده اعلام می شود.

2. Cooperation : نرون برنده سپس مرکز همسایگی توپولوژیکال نرون های برانگیخته را در مرحله قبل قرار می دهد و پایه ای برای همکاری بین نرون های همسایه برانگیخته فراهم می کند.

3. Adaption : نرون های برانگیخته در همسایگی، مقادیر فردی تابع جداساز شان را با در نظر گرفتن توزیع داده ورودی از طریق تنظیمات دقیق به گونه ای که پاسخ نورون برنده برای ورودی بعدی مشابه افزایش می یابد، کاهش/ افزایش می دهند. مرحله Adaption در دو مرحله فرعی قابل تشخیص است:

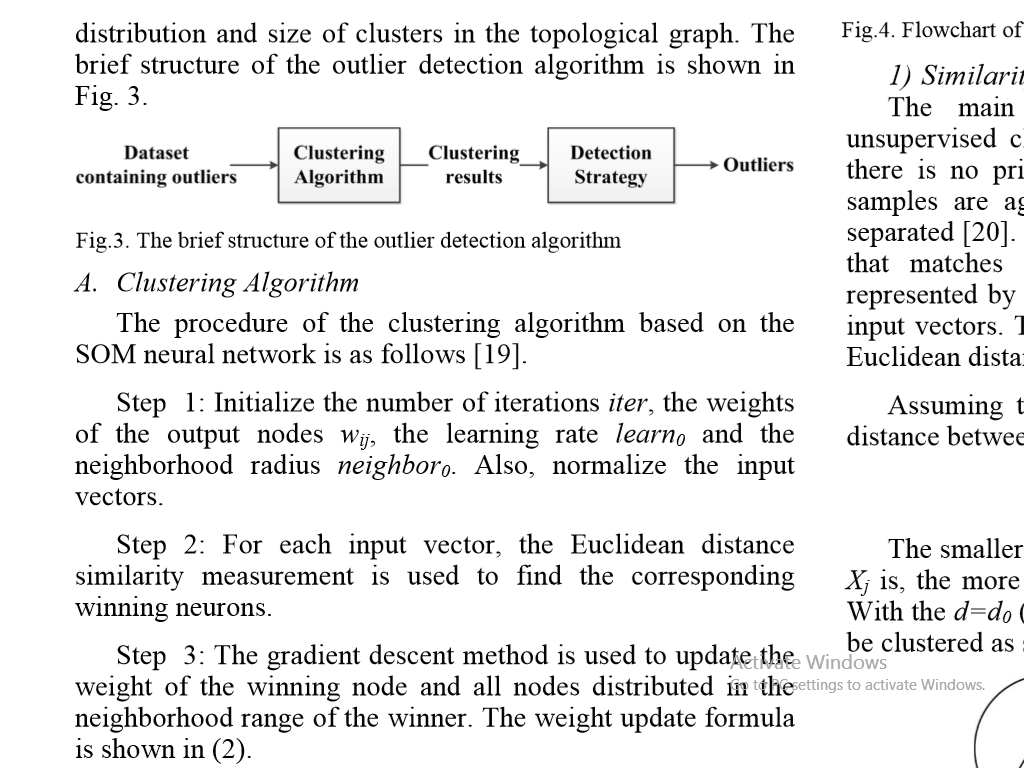
(1) مرحله مرتب سازی یا خود سازماندهی ، که در آن بردارهای وزن با توجه به فضای توپولوژیکی دوباره مرتب می شوند

(2) گام همگرایی که نگاشت بخوبی تنظیم شده و برای ارایه توصیف آماری فضای ورودی صحیح اعلام می شود. این مرحله ای است که نگاشت همگرا اعلام شده و در نتیجه آموزش داده شده است.

یک نیاز اساسی در یادگیری یک SOM افزونگی داده ورودی است تا درباره ساختار اصلی الگوهای فعالسازی نرون یاد بگیرد. بعلاوه مقدار مناسب داده برای ایجاد خوشه های قابل تشخیص نیاز است. با وجود داده های کافی برای مشکل طبقه بندی ، مشکل منطقه خاکستری بین خوشه ها و ایجاد خوشه های بی نهایت کوچک وجود دارد که داده های ورودی دارای الگوهای حداقل هستند.

پایان [2]

[3]



Outlier detection algorithm based on SOM neural network:

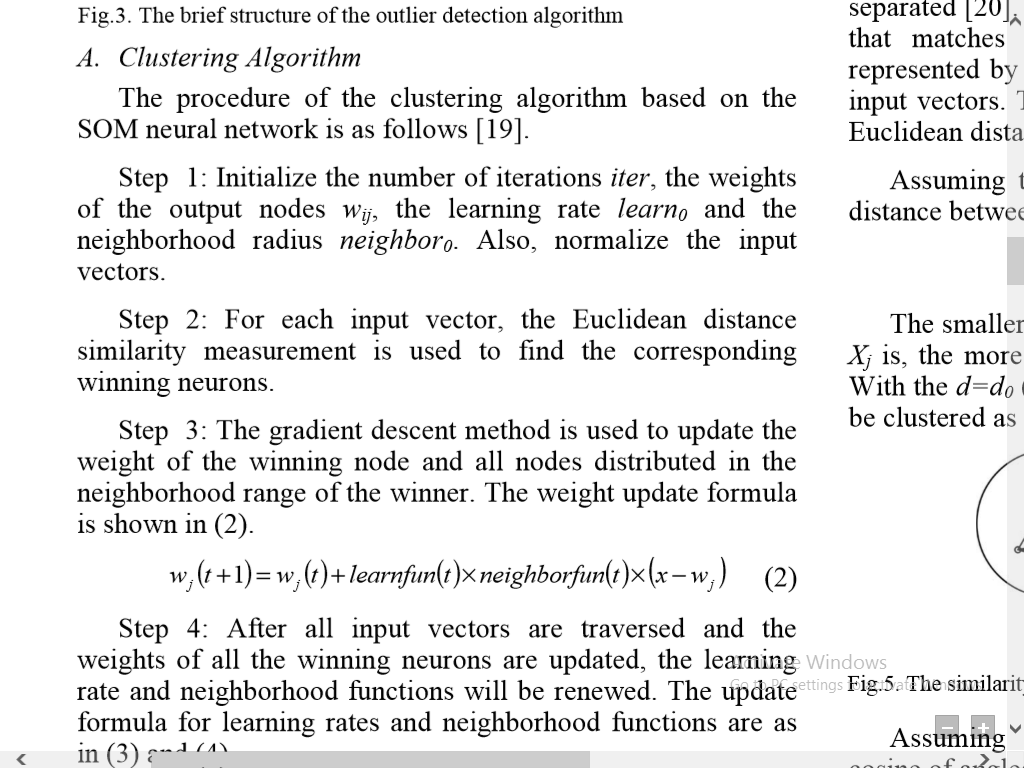
روش تشخیص داده پرت بر اساس شبکه عصبی SOM شامل دو بخش است: الگوریتم خوشه بندی و استراتژی تشخیص. الگوریتم خوشه بندی می تواند داده های پرت را بر اساس توزیع و سایز خوشه ها در گراف توپولوژیکی بیابد. ساختار مختصری از الگوریتم تشخیص داده پرت در شکل بالا نشان داده شده است.

1. الگوریتم خوشه بندی: فرایند الگوریتم خوشه بندی بر اساس شبکه عصبی SOM به صورت زیر است:

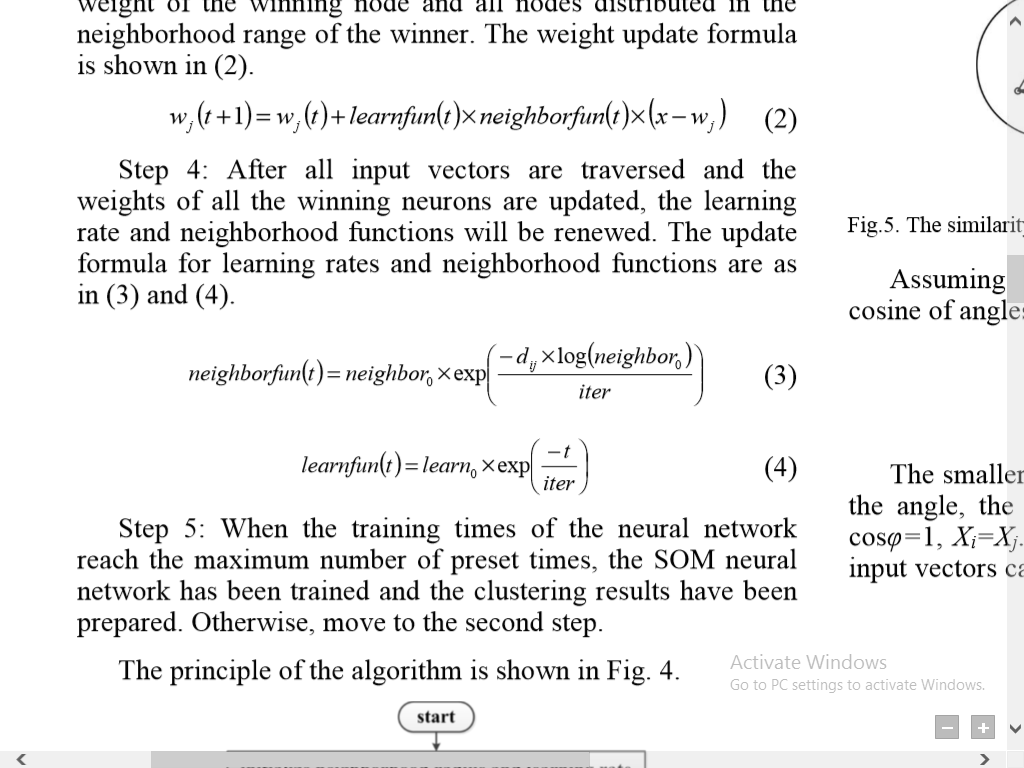
گام 1) مقداردهی اولیه تعداد چرخه iter ، وزن نودهای خروجیwij ، نرخ یادگیری learn0، وشعاع همسایگی neighbor0 و همچنین استانداردسازی بردارهای ورودی.

گام 2) برای هر بردار ورودی، معیار شباهت فاصله اقلیدسی برای یافتن نرون های برنده متناظر به کار می رود.

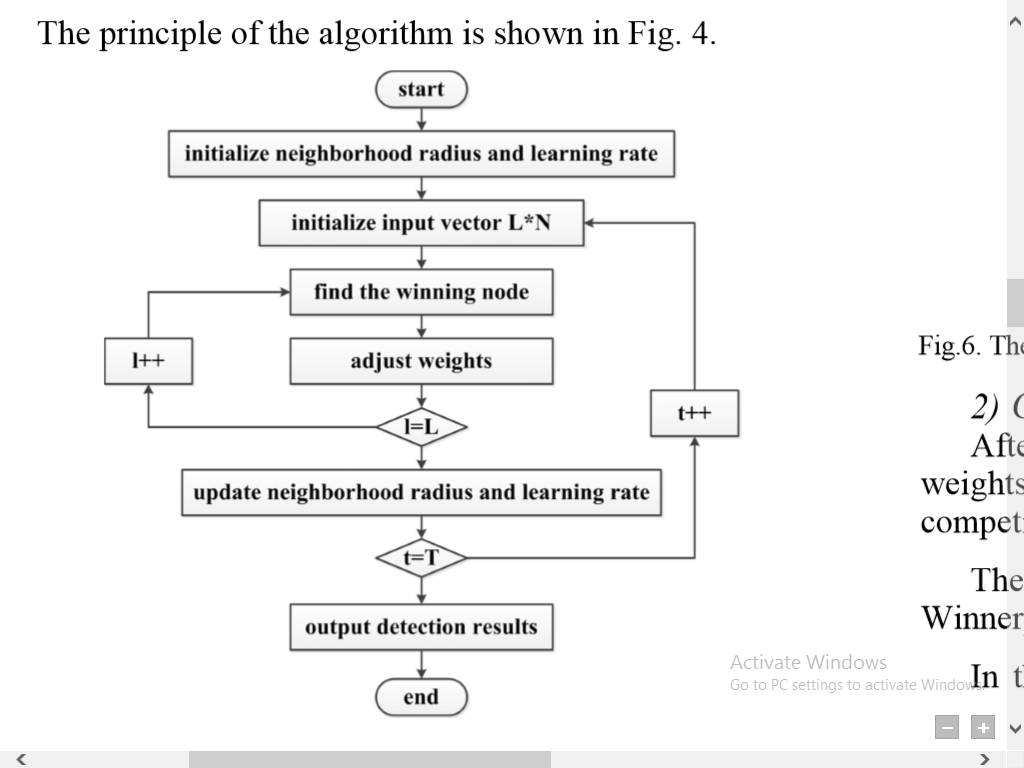
گام 3) روش گرادیان کاهشی به منظور به روز رسانی وزن نودهای برنده و توزیع همه نودهای همسایه ی برنده. فرمول به روز رسانی نودها به صورت زیر است:



گام 4) بعد از پیمایش همه ی بردارهای ورودی و به روز رسانی وزن همه نرون های برنده، نرخ یادگیری و تابع همسایگی تجدید می شوند. فرمول به روز رسانی نرخ یادگیری و تابع همسایگی به صورت زیر است:



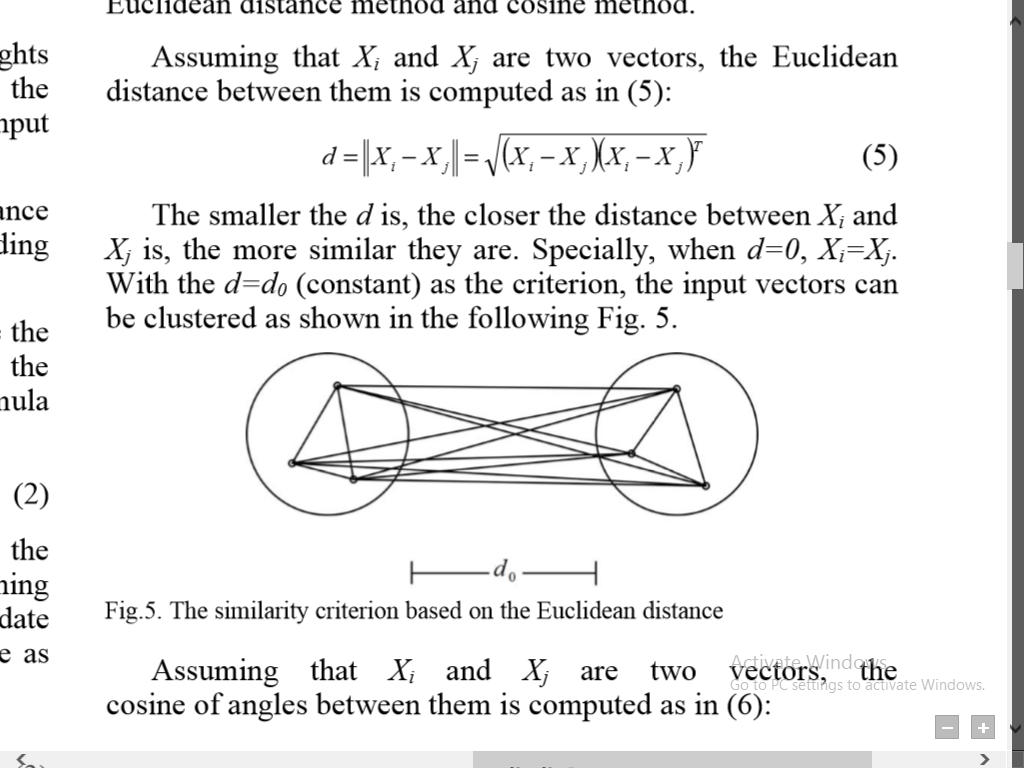
گام 5) وقتی تعداد آموزش شبکه عصبی به بیشینه تعداد فعلی برسد، شبکه عصبی SOM آموزش داده شده است و نتایج خوشه بندی آماده شده است. درغیر اینصورت به گام 2 برو.



1. شرط تشابه:

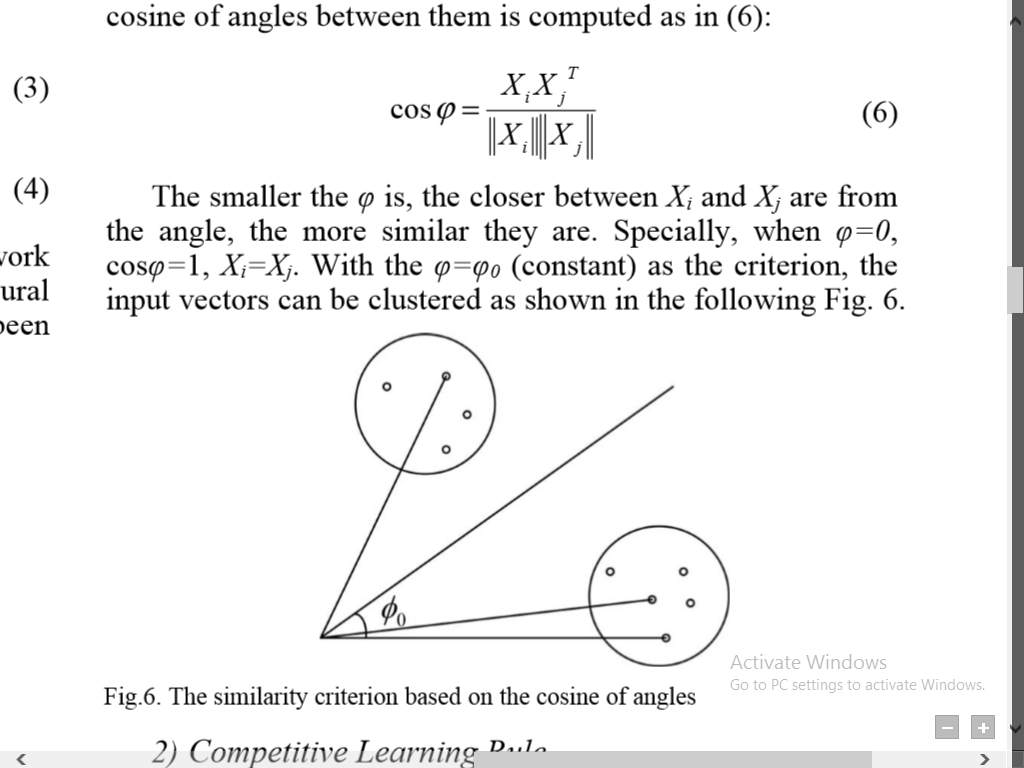
کاربرد اصلی شبکه عصبی SOM خوشه بندی بدون ناظر است.خوشه بندی بدون ناظر بدین معنی است که هیچ اطلاعات از پیش تعیین شده ای درباره مجموعه داده آزمون وجود ندارد. نمونه های مشابه با یکدیگر تجمیع شده و نمونه های مختلف از هم جدا می شوند. برای هر بردار ورودی نیازمند یافتن نودهایی که بیشترین تطبیق را دارند هستیم. به عبارت دیگر یافتن نرون هایی که با بردارهای وزن نشان داده می شوند و بیشترین شباهت به بردار ورودی را دارند. دو شرط برای اندازه گیری مشابهت وجود دارد: روش فاصله اقلیدسی و روش cosine.

فرض کنید که Xi و Xj دو بردار هستند و فاصله اقلیدسی بین آن ها بصورت زیر محاسبه می شود:



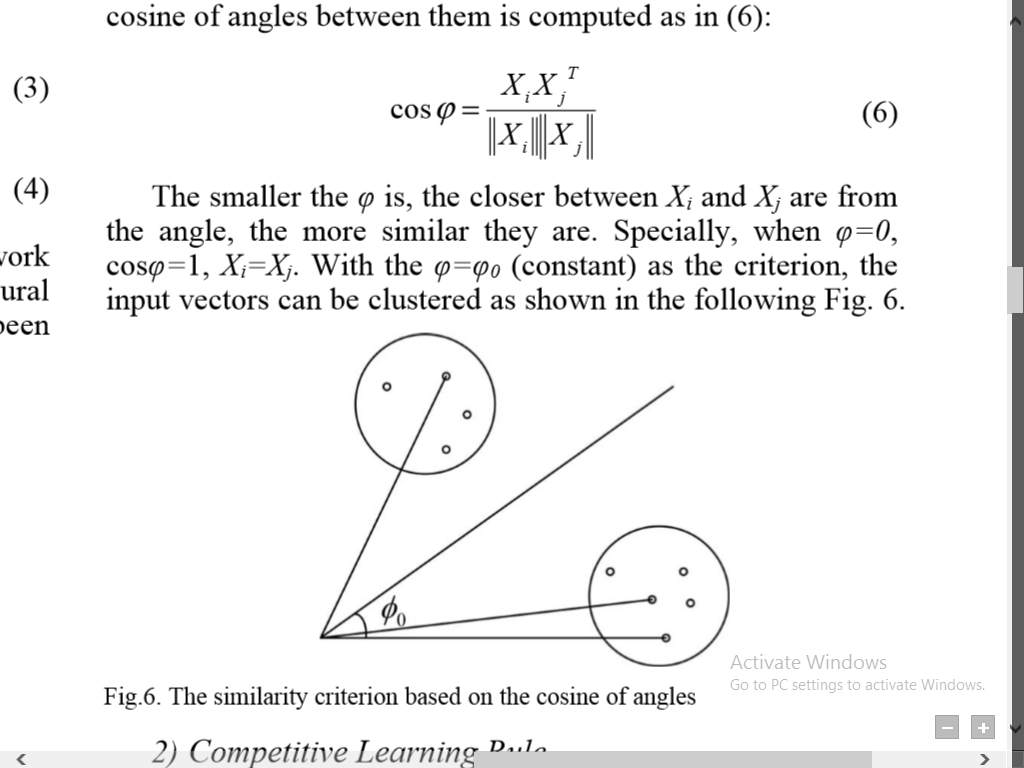
مقدارd کوچکتر بیانگر فاصله نزدیکتر بین Xi و Xj  و درنتیجه شباهت بیشتر میان آن دو است.

فرض کنید که Xi و Xj دو بردار هستند و cosineبین زاویه ها بصورت زیر محاسبه می شود:



مقدار کوچکتربیانگر نزدیکی Xi و Xj تحت زاویه و شباهت بیشتر میان ان هاست و وقتی است

cos و Xi = Xj است.

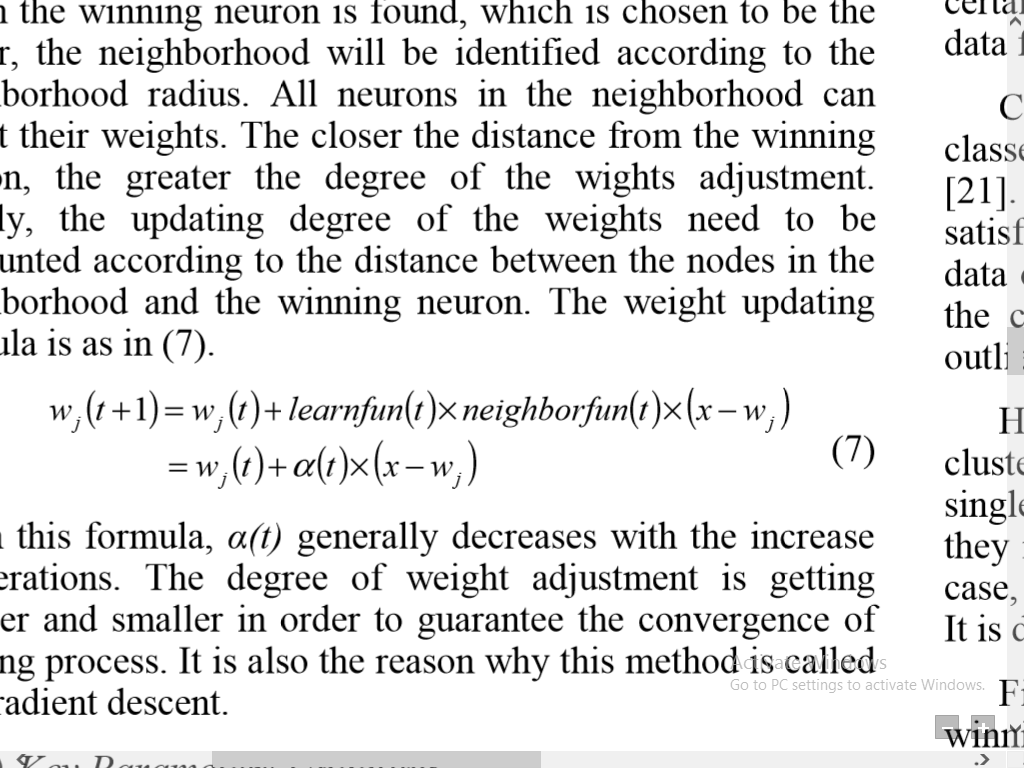


1. قانون یادگیری رقابتی:

بعد از یافتن نرون های برنده، نیاز است که وزن ها را تنظیم کنیم به گونه ای که برای رقابت مطلوب تر باشد. در اینجا دو روش برای تنظیم وزن ها هست: الگوریتم Winner Take All (WTA) و الگوریتم Kohonen .

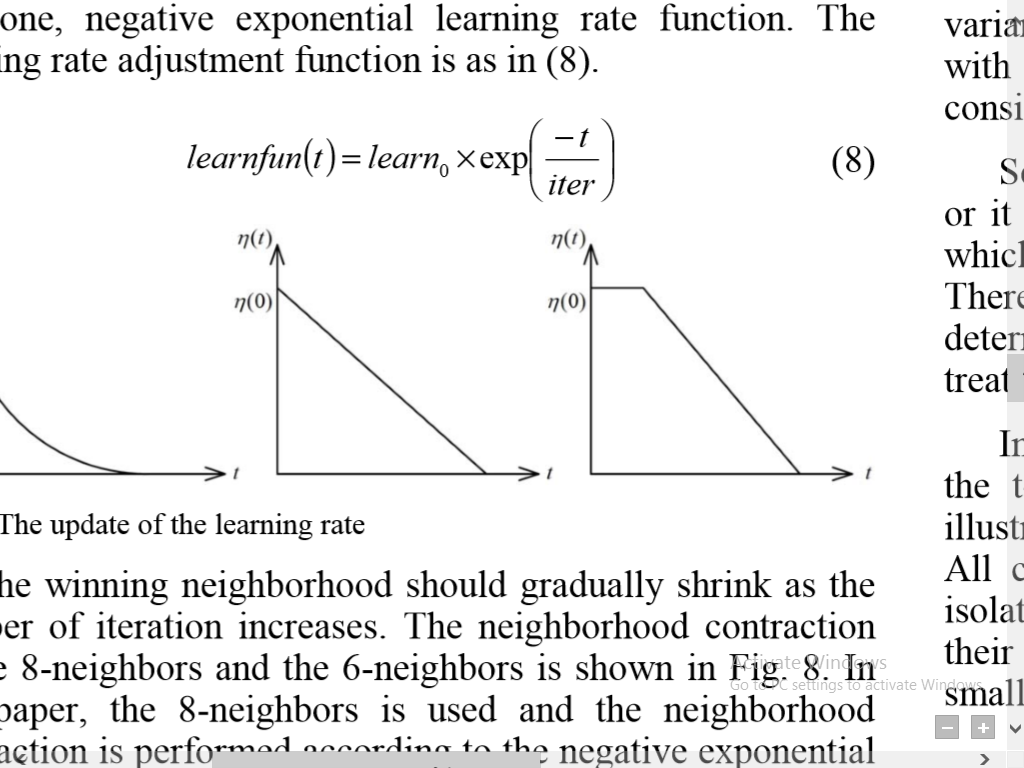
در الگوریتم WTA فقط نرون های برنده حق دارند که وزن های خود را تنظیم کنند. در این روش در تمام فرایند آموزش، تعداد کمی از نرون ها می توانند برنده شوند و وزن خود را تنظیم کنند و بنابراین تعداد زیادی نرون مرده در شبکه عصبی یافت می شوند.

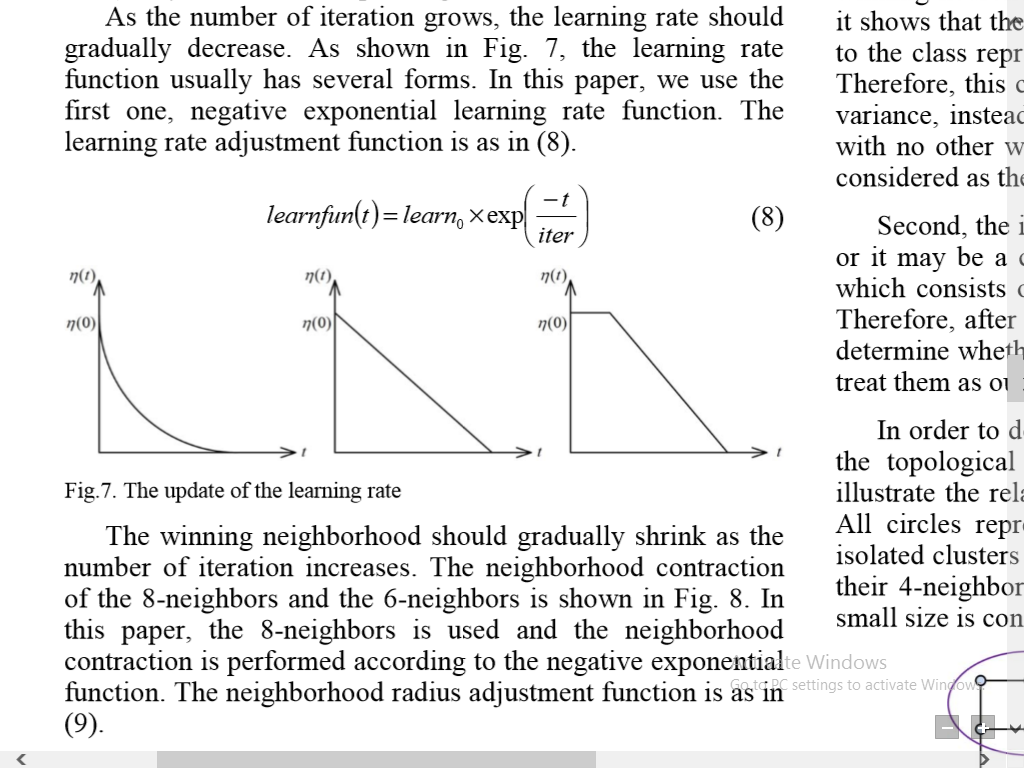
مهار افقی (lateral inhibition) الگوریتم Kohonen ناحیه ای است. وقتی نرون برنده ای یافت می شود که منتخب مرکزیت است، همسایگی براساس شعاع همسایگی تعیین می شود. تمام نرون ها در همسایگی می توانند وزن خود را تنظیم کنند. فاصله نزدیک تر به نرون برنده ، درجه تنظیم وزن بیشتری داراست. به طور خلاصه درجه به روز رسانی وزن ها باید براساس فاصله بین نودها در همسایگی و نرون برنده، کاهش یابد. فرمول به روز رسانی وزن ها به صورت زیر است:



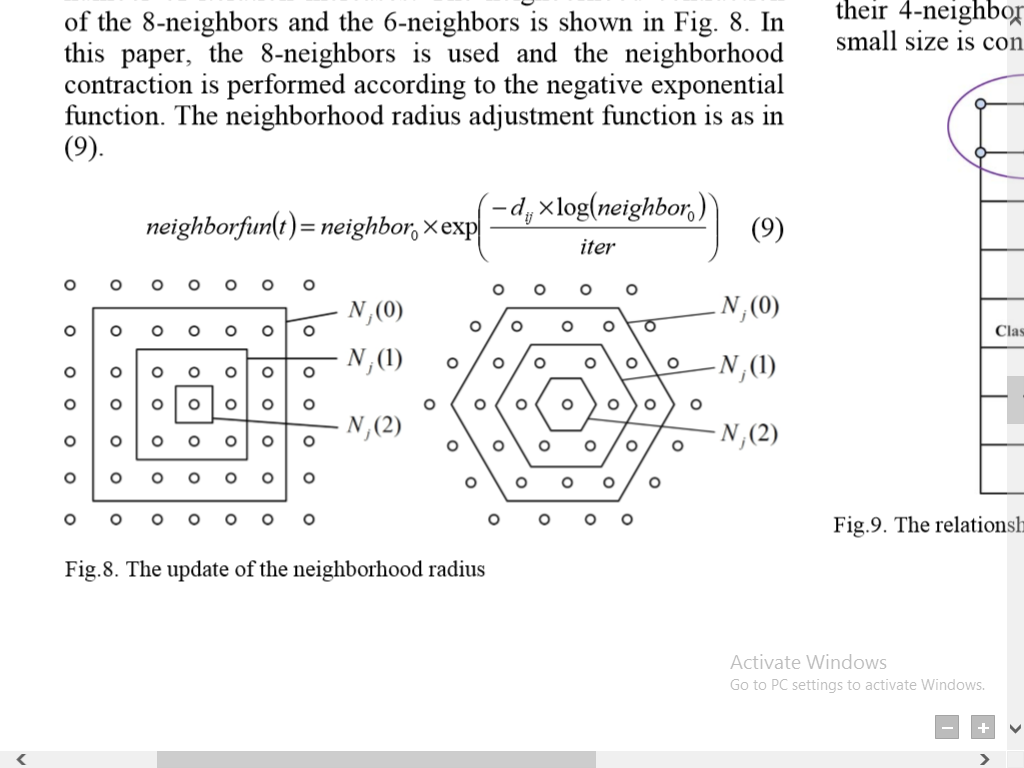
در این فرمول (t) α با افزایش تعداد چرخه ها کاهش می یابد. درجه تنظیم وزن ها کمتر و کمتر می شود تا همگرایی فرایند یادگیری را تضمین کند. هم چنین دلیلی است بر اینکه چرا نام این روش گرادیان کاهشی است.

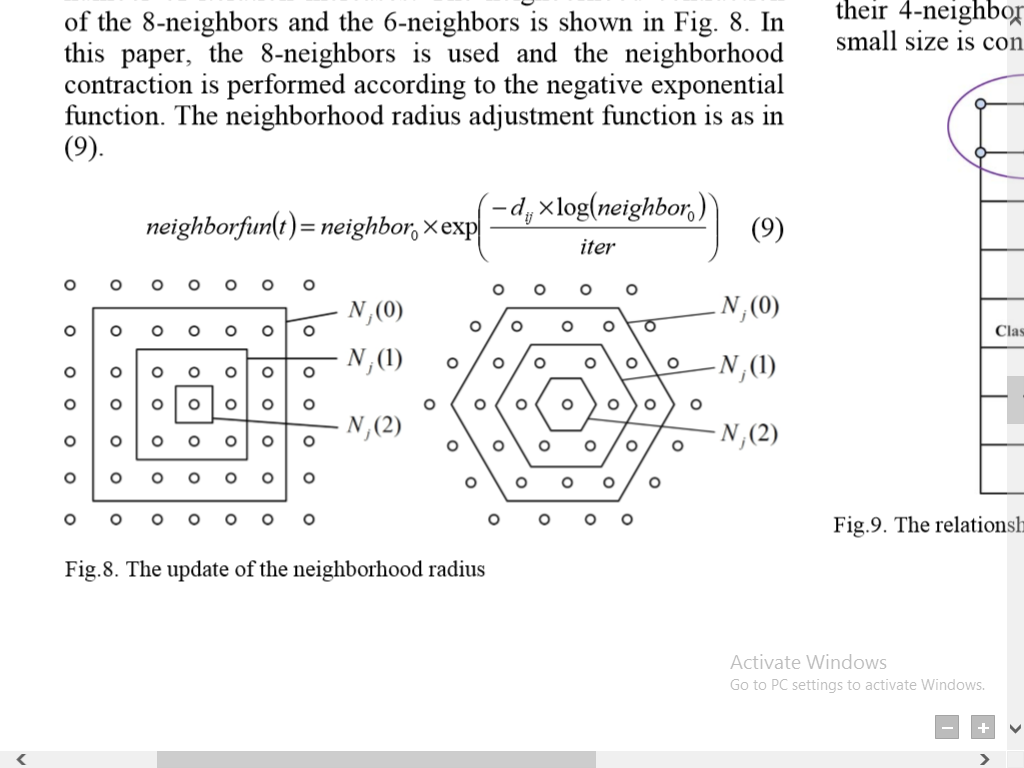
1. به روز رسانی پارامترهای کلیدی: با زیاد شدن تعداد چرخه ها، نرخ یادگیری باید کاهش یابد. همانطور که در شکل 7 میبینید تابع نرخ یادگیری معمولا فرم های مختلفی دارد. در این تحقیق ما از اولی، تابع نرخ یادگیری نمایی منفی استفاده می کنیم. تابع تنظیم نرخ یادگیری به صورت زیر است:





همسایگی برنده باید به تدریج با افزایش تعداد چرخه ها ، تقسیم شود. در شکل زیر فرم 8 همسایگی و 6 همسایگی نشان داده شده است. در این تحقیق طبق تابع نمایی منفی ما از 8 همسایگی استفاده کردیم . تابع تنظیم شعاع همسایگی نیز به صورت زیر است:



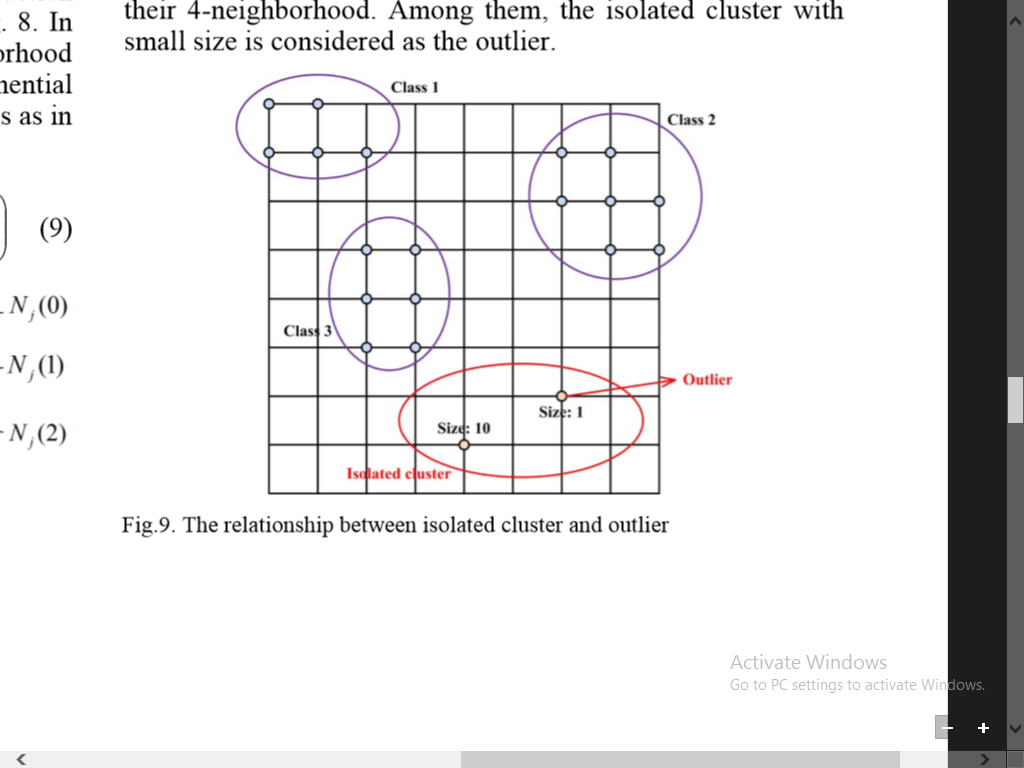


1. استراتژی تشخیص:

داده های پرت می توانند به دو نوع مختلف تقسیم شوند: داده های پرت تنها، داده های پرت مجتمع[[2]](#footnote-2) . داده پرت ایزوله به تغییر تصادفی داده در زمانی خاص اشاره دارد. داده های پرت مجتمع، طیف کوچکی از تغییرات داده هستند. خوشه بندی فرایند تقسیم یک مجموعه به چندین کلاس شامل اشیا با ارتباطات با شباهت بالا می باشد. نقاط داده که بعنوان داده پرت قضاوت می شوند باید یک یا دو شرطی را که در ادامه می آید را براورده کنند. اول اینکه نقاط داده به هیچ کلاسی تعلق ندارند و دیگر اینکه سایز کلاس بطور واضحی کوچکتر از سایر کلاس هاست. داده های پرت معمولا مکمل نتایج خوشه بندی هستند. اگرچه در عمل مشخص شده است که برخی خوشه ها با واریانس میان کلاسی بزرگتر، تنها با یک نرون برنده در توپولوژی شبکه عصبی SOM نشان داده نمی شوند بلکه به چند نرون نسبتا نزدیک نگاشت می شوند. در این مورد، تشخیص داده پرت فقط به یافتن کوچکترین کلاس ها مربوط نمی شود و به دو مرحله زیر تقسیم می گردد:

اولا ما به دنبال خوشه های نمونه بدون نرون های برنده دیگر در 4 همسایگی هستیم. اگر نرون های برنده دیگری در 4 همسایگی یک نرون برنده باشند ، نشان می دهد که نقاط داده دیگری وجود دارند که بسیار شبیه به کلاس نرون برنده در دیتاست هستند. بنابراین این کلاس باید بعنوان یک خوشه با واریانس بزرگتر، به جای یک خوشه ایزوله درنظر گرفته شود.

دوما خوشه بندی ایزوله الزاما داده پرت نیست و ممکن است کلاسی با واریانس بین کلاسی بسیار کوچک باشد که شامل تعدادی از نقاط داده بسیار شببیه به هم باشد. بنابراین پس از یافتن خوشه های ایزوله، ضروری است که مشخص شود آیا سایز این خوشه ها به اندازه کافی کوچک هست که بعنوان داده پرت درنظر گرفته شوند یا خیر. برای توصیف استراتژی تشخیص بطور جزیی تر، از گراف توپولوژیکی شبکه عصبی در شکل زیر استفاده می کنیم تا ارتباط میان خوشه های ایزوله و داده های پرت را نشان دهیم. دایره ها بیانگر نرون های برنده اند. دو دایره قرمز خوشه های ایزوله هستند زیرا نرون برنده دیگری در 4 همسایگی انها وجود ندارد. از میان آنها خوشه های ایزوله با سایز کوچک بعنوان داده پرت در نظر گرفته می شوند.

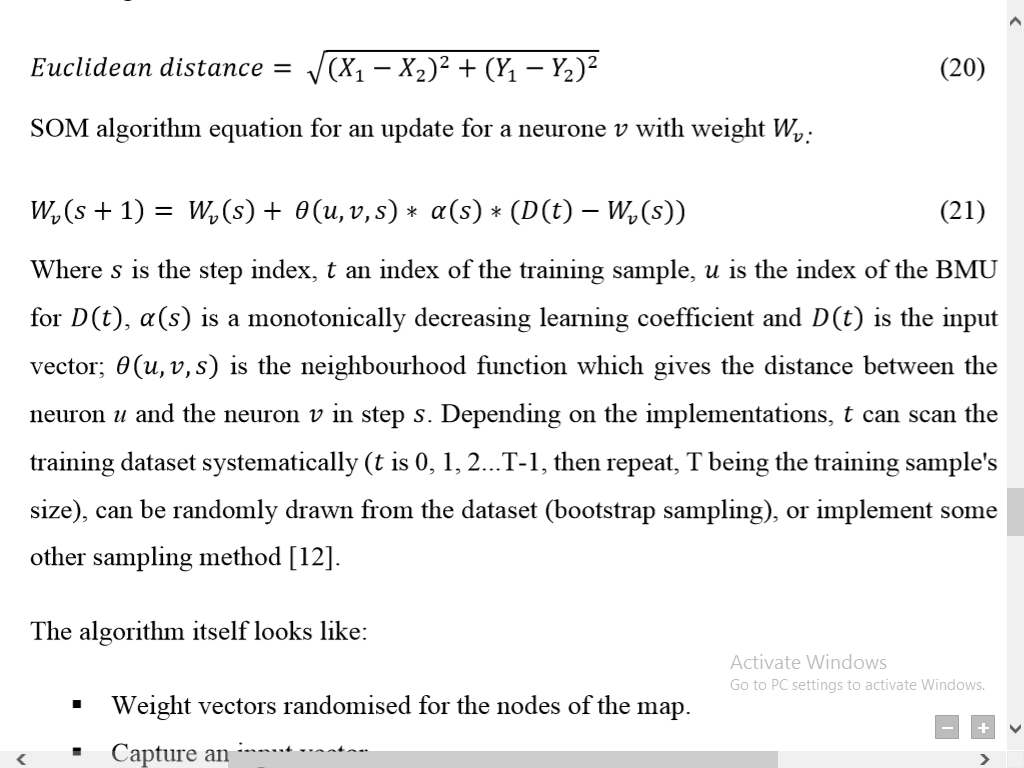


در مجموع داده های پرت یافت شده توسط الگوریتم خوشه بندی شبکه عصبی SOM باید خوشه های ایزوله ای باشند که نرون برنده ای در 4 همسایگی شان وجود ندارد و سایز کلاسی کوچکی داشته باشند.

پایان [3]

Self-organizing map[1]

داده چند بعدی را درفضای ابعاد کوچکتری نمایش می دهد که می تواند هرچند بعدی باشد اما معمولا یک یا دو بعدی است. یک نمایش از فضای ورودی نمونه های آموزشی که نمونه برداری شده یک map گفته می شود. به طور کلی روشKohonen یک شبکه تولید می کند که اطلاعات را به گونه ای ذخیره می کند که الگوریتم هر ارتباط توپولوژیکی در مجموعه آموزشی را نگه می دارد. یک نوع شبکه عصبی هوشمند است که با استفاده از روش های یادگیری بدون ناظر اموزش داده می شود.SOM از سایر شبکه های عصبی متفاوت است زیرا یادگیری رقابتی را برخلاف یادگیری تصحیح خطا( مانند انتشار به عقب با کرادیان کاهشی) اعمال می کنند و از یک تابع همسایگی برای حفظ ویژگی های توپولوژیکی فضای ورودی استفاده می کنند. SOM در دو مرحله آموزش و نگاشت[[3]](#footnote-3) عمل می کند. فرایند یادگیری map را با استفاده از مثال های ورودی، اندازه گیری بردار می سازد.نگاشت بطور خودکار یک بردار ورودی جدید را طبقه بندی می کند.SOM اجزایی که نود نام دارند یا بعنوان نرون نیز شناخته می شوند را ترکیب می کند که در ان بردار وزن مربوط به هر نرون با همان اندازه بعنوان بردار داده ورودی و موقعیتی در فضای نگاشت است. آرایش صحیح نرون ها یک طرح دوبعدی قراردادی است در یک شبکه شش ضلعی یا مستطیلی است.SOM یک نگاشت از ورودی با ابعاد بالا به ابعاد کمتر است. فرایند برای قرار دادن یک بردار از داده روی نقشه برای یافتن گره با کمترین بردار فاصله استاندارد از بردار فضای داده است.در مرحله آموزش مثال ها در شبکه قرار می گیرند.این الگوریتم فاصله اقلیدسی از همه بردارهای وزن را محاسبه می کند. نرون با بیشترین شباهت بردار وزن به ورودی، BMU نامیده می شود.این روش وزن های BMU و نزدیکترین نرون ها در شبکه SOM به بردار ورودی را تنظیم می کند. الگوریتم SOM برای به روز رسانی نرون v با وزن Wv به صورت معادله زیر عمل می کند.



که s بیانگر شاخص گام است، t نمونه ی آموزشی است،D(t) بردار ورودی، u بیانگر BMU برای D(t) ، (s)α یک ضریب یادگیری بطور یکنواخت در حال کاهش است و تابع همسایگی است که فاصله بین نرون u و v با گام s را می دهد. بسته به پیاده سازی،t (0,1,2,…,T-1) می تواند بطور سیستماتیک مجموعه داده آموزشی را بپیماید و بعد تکرار شود(T سایز نمونه های آموزشی است)، می تواند بطور تصادفی از مجموعه داده گرفته شود(نمونه برداریbootstrap) و یا سایر روش های نمونه برداری را پیاده سازی کند. خود الگوریتم به شرح زیر است:

* بردارهای وزن برای نودهای نقشه تصادفی می شوند
* بردار ورودی را بگیر
* به هر گره (نرون) روی نقشه برو و مشابهت بین بردار ورودی و بردار وزن نقشه را پیدا کن. نود با بهترین عدد را پیدا کن
* از فرمول بالا برای به روز رسانی نودها در همسایگیBMU استفاده کن
* S را افزایش بده و فرایند را از مرحله 2 تکرار کن تا S کمتر از حد مجاز تکرار شود

یک مساله شناخته شده برای SOM انتخاب مناسب مقادیر اولیه است، با استفاده از انتخاب مقادیر اولیه تصادفی وزن های som برای این روش.

Fig : plot of the training progress

Fig :WSS(Within Groups Sum of Squares) with some potential clusters(Index)

Table :Cross Validation results of a Self-organized map model

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| MSE | ERR | BER | Acc. | Prec. | FNR | FPR | TNR | TPR | Num |
| 1.001 | 0.006 | 0.452 | 0.994 | 0.040 | 0.901 | 0.004 | 0.996 | 0.099 | CV1 |
|  |  |  |  |  |  |  |  |  | CV2 |
|  |  |  |  |  |  |  |  |  | CV3 |
|  |  |  |  |  |  |  |  |  | CV4 |
|  |  |  |  |  |  |  |  |  | CV5 |
|  |  |  |  |  |  |  |  |  | Mean |

Fig :Self-organized map CV AUC and ROC results

CV1

CV2

CV3

CV4

Test1

Test2

Test3

Table :confusion matrix of a self-organized map

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | TP | FP | TN | FN |
| Test1 |  |  |  |  |
| Test2 |  |  |  |  |
| Test3 |  |  |  |  |

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| AUC | MSE | ERR | BER | Acc. | Prec. | FNR | FPR | TNR | TPR |  |
|  |  |  |  |  |  |  |  |  |  | T1 |
|  |  |  |  |  |  |  |  |  |  | T2 |
|  |  |  |  |  |  |  |  |  |  | T3 |
|  |  |  |  |  |  |  |  |  |  | Mean |

Table : Test result of Self-organized map

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Specificity | Sensivity | AUC | MSE | ERR | BER | Acc. | Prec. | FNR | FPR | TNR | TPR | Algorithm |
| - | - | 0.800 | 1.001 | 0.006 | 0.452 | 0.994 | 0.040 | 0.901 | 0.004 | 0.996 | 0.099 | SOM |
| 0.676 | 0.536 | 0.629 |  |  |  |  |  |  |  |  |  | LOF40 |
| 0.679 | 0.497 | 0.613 |  |  |  |  |  |  |  |  |  | KNN1 |
| 0.645 | 0.527 | 0.603 |  |  |  |  |  |  |  |  |  | URF100 |
| 0.650 | 0.463 | 0.555 |  |  |  |  |  |  |  |  |  | AE50\_Tanh |
| 0.436 | 0.712 | 0.554 |  |  |  |  |  |  |  |  |  | IF100 |

Table : مقایسه نتایج الگوریتم های مختلف

Fig :AUC curve –SOM10,20,40,80

Fig :AUC curve –All models

[1] 144117IAPM, Igor Anohhin. “DATA MINING AND MACHINE LEARNING FOR FRAUD DETECTION.” (2017).

[2] Usama, Muhammad, et al. "Unsupervised machine learning for networking: Techniques, applications and research challenges." IEEE Access 7 (2019): 65579-65615.

[3] Y. Liu and H. Lu, "Outlier detection algorithm based on SOM neural network for spatial series dataset," 2018 Tenth International Conference on Advanced Computational Intelligence (ICACI), Xiamen, 2018, pp. 162-168.

[4] Stefanovič, P., and O. Kurasova. "Outlier detection in self-organizing maps and their quality estimation." Neural Network World 28.2 (2018): 105-117.

[5] Anomaly detection using a self-organizing map and particle swarm optimization- sharif university

SOM [5] شناخته شده ترین روش شبکه های عصبی بدون ناظر برای کلاس بندی است. معماری SOM یک شبکه عصبی feed- forward با یک لایه از نورون های ساختار یافته در یک آرایه rectangular است. وقتی یک الگوی ورودی به SOM داده می شود هر نورون محاسبه می کند که هر ورودی چقدر به وزن هایش شبیه است. نورون هایی که وزنشان بیشترین شباهت را دارند، برنده ی رقابت برای الگوی ورودی هستند و وزن نورون های برنده تقویت می شود تا بر خروجی اثر بگذارد. نورون های برنده در هر گام بیشترین میزان یادگیری را دارند، با همسایگانی که کمترین میزان یادگیری را دارند. [34]

مزایا

کار با مجموعه داده های با ابعاد بالا دشوار است. SOM اطلاعات را کاهش می دهد در حالیکه مهم ترین ارتباطات توپولوژیکی اجزای داده در دو بعد را فراهم می کند.

SOM ها با استفاده از یادگیری بدون ناظر آموزش داده می شوند، هیچ فرض قبلی یا اطلاعاتی از پیش در مورد اجزای داده در دسترس نیستند.

الگوریتم SOM در کار با داده های عظیم بسیار موثرعمل می کند حتی وقتی که داده ها نویزی باشند.

معایب

تعداد کلاسترها نیاز است که تعیین شوند. کلاسترینگ یک فرایند دو بخشی است: تعیین تعداد کلاسترها و کلاسترینگ داده ها. از انجایی که مشخصات مجموعه داده از پیش مشخص نیست، تعیین تعداد کلاسترها کار پیش پا افتاده ای نیست. این می تواند با اجرای الگوریتم با مقادیر مختلف کلاستر و انتخاب مناسب ترین نتیجه کلاسترینگ بر اساس گروهی از شایستگی ها باشد.

کاربر می تواند با بررسی دستی یا اعمال الگوریتم های سنتی مانند وراثت یا partitive مرزهای کلاستر را پیدا کند.

اخیرا

Refrence:

Anomaly detection using a self-organizing map and particle swarm optimization- sharif university

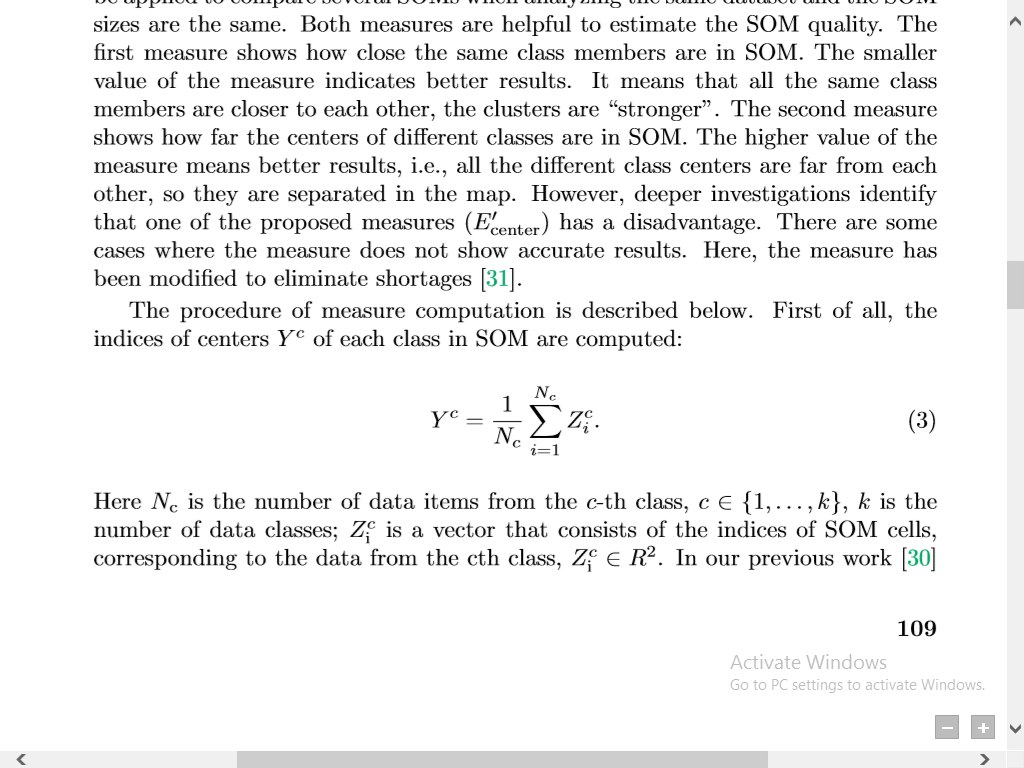
پایان [5]

[4]

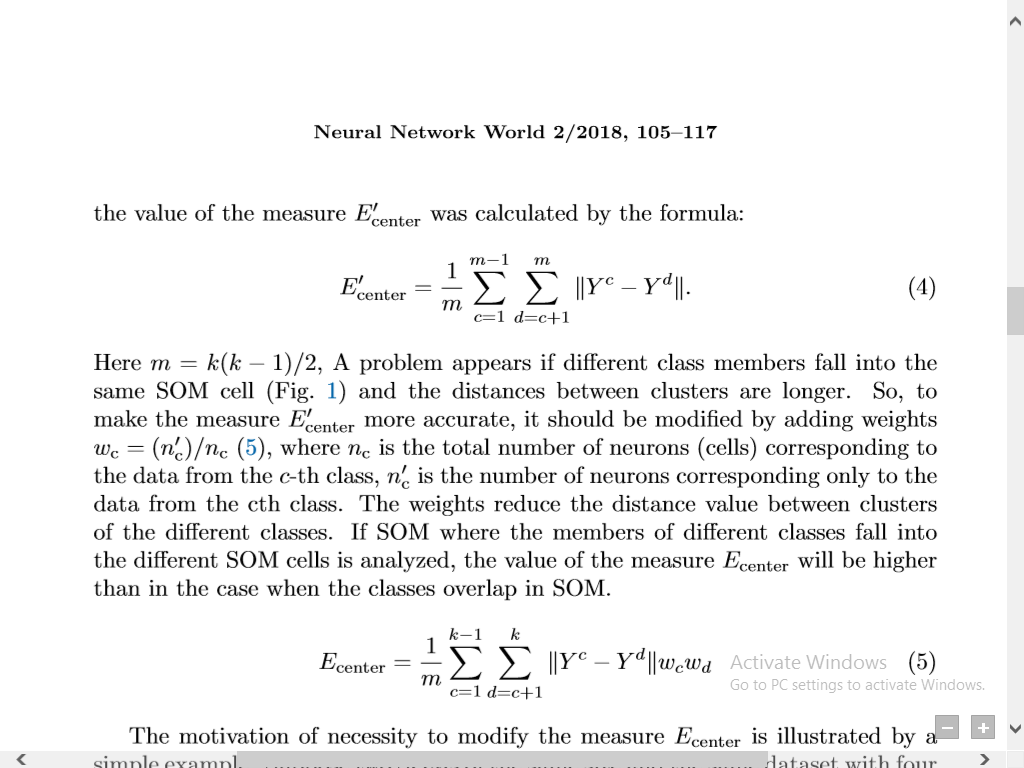
Estimation of the SOM quality

بعد از اموزش SOM معمولا خطاهای کمی و توپولوژیکی محاسبه می شوند. خطای کمی نشان می دهد که نرون های آموزش داده شده ی شبکه به چه میزان با بردارهای ورودی سازگار هستند. این یک فاصله میانگین بین بردارهای داده ورودی Xp, P=1,…,N و نرون های برنده است. خطای توپولوژیکی نشان می دهد که شبکه آموزش داده شده به چه میزان توپولوژی داده های تحلیل شده را نگه می دارد. وقتی داده های کلاس بندی شده توسط روش های خوشه بندی تحلیل می شوند، نیاز است انطباق بین کلاس های داده ها و خوشه های بدست آمده ارزیابی شود. این انطباق بیانگر این است که داده ها به کلاس مناسب نسبت داده شده اند. در یک مورد ناسازگار، محقق باید علل ناسازگاری را جستجو کند. یکی از دلایل ممکن این است که داده به کلاس نامناسبی نسبت داده شده است. اگرچه که نه خطای توپولوژیکی و نه خطای کمی نشان نمیدهد که آیا کلاس های داده ی تحلیل شده با خوشه های تشکیل شده در SOM منطبق هستند یا نه. خطاهای دیگری وجود دارند که به ارزیابی انطباق بین کلاس ها و خوشه های بدست آمده که از SOM بدست نیامده اند، کمک میکنند. اگرچه که در آن موارد داده باید بدون ابهام به یکی از خوشه ها نسبت داده شود. منحصربفرد بودن SOM در مقایسه با سایر روش های خوشه بندی این است که در نتایج SOM هیچ خوشه ی سختگیرانه بیان شده ای[[4]](#footnote-4) وجود ندارد. بعنوان مثال اینکه کدام ایتم داده به کدام خوشه نسبت داده شود، بیان نمی شود، یک محقق فقط می تواند خوشه ها را در نمایش بصری SOM مشاهده کند.

معیارها می توانند برای مقایسه چندSOM زمانی که یک دیتاست را تحلیل می کنند و سایز SOM ها یکی است، اعمال شوند. هردو معیار برای تخمین کیفیت SOM موثرند. معیار اول نشان می دهد که اعضای یک کلاس درSOM چقدر بهم نزدیکند. مقدار کمتر نشان دهنده نتایج بهتر است. این به این معنی است که اعضای یک کلاس بهم دیگر بسیار نزدیک هستند. معیار دوم نشان می دهد که مرکزیت کلاس های مختلف در SOM چقدر ازهم دورند. مقدار بیشتر نشان دهنده نتایج بهتر است. اگرچه که تحقیقات بیشتر نشان داده که یکی از معیارهای ارایه شده E’center معایبی دارد. مواردی وجود دارد که این معیار نتایج دقیق را نشان نمی دهد. در اینجا این معیار تغییراتی یافته تا نقایص را حذف کند. فرایند محاسبه این معیار در زیر توصیف شده است. ابتدا شاخص مرکزیت های هر کلاسYc در SOM محاسبه می شود:



در اینجا منظور از Nc تعداد داده ها از کلاس c ام c {1,…,k} است.k تعداد کلاس های داده است.Zic برداری است که شامل شاخص های سلول های SOM، متناظر با داده از کلاس c ام است.



پایان [4]

1. Lateral inhibition [↑](#footnote-ref-1)
2. Assembled outliers [↑](#footnote-ref-2)
3. Map [↑](#footnote-ref-3)
4. Strictly expressed clusters [↑](#footnote-ref-4)