https://gep.ui.ac.ir/?lang=en
Geography and Environmental Planning
E-ISSN: 2252- 0910
Document Type: Research Paper
Vol. 34, Issue 4, No.92, Winter 2024, pp. 1- 2
Received: 15/05/2023 Accepted: 17/07/2023

Comparing the Performance of Spatial and Non-Spatial Self-Organizing Neural Networks in Clustering Socio-Economic Data of Isfahan Census Blocks

Hadi Tarigholizadeh[1], Babak Mirbagheri *[2], Ali Akbar Matkan[3]
1- Ph.D. Student of Remote Sensing and GIS, Center for Remote Sensing and GIS Research, Faculty of Earth Sciences, Shahid Beheshti University, Tehran, Iran
gholizadeh.hadi20@gmail.com

2- Assistant Professor, Center for Remote Sensing and GIS Research, Faculty of Earth Sciences, Shahid Beheshti University, Tehran, Iran
b_mirbagheri@sbu.ac.ir

3- Professor, Center for Remote Sensing and GIS Research, Faculty of Earth Sciences, Shahid Beheshti University, Tehran, Iran
a-matkan@sbu.ac.ir

Abstract

The increasing volume and dimensions of spatial data have made self-organizing neural networks a prominent tool for analyzing large and multi-dimensional datasets. Clustering, an approach for extracting knowledge from big data, aims to group similar data into clusters. This research focused on clustering socio-economic data of census blocks associated with urban sustainable development using self-organizing neural networks with and without spatial parameters referred to as SOM and Geo-SOM, respectively. Both algorithms employ the same clustering process but differ in the inclusion of spatial parameters, specifically the geographic coordinates of block centroids, in the Geo-SOM algorithm. The SOM and Geo-SOM algorithms were trained and applied to cluster the data. The resulting clusters exhibited distinct dissimilarities, demonstrating that clustering census block data solely based on non-spatial attributes leads to heterogeneous and incongruent clusters, whereas incorporating spatial parameters yields homogeneous and congruent clusters. Evaluation of the results using Silhouette coefficient indicated that Geo-SOM outperformed SOM in clustering the data with average Silhouette coefficients of -0.02 and 0.27 for SOM and Geo-SOM, respectively. Comparison of the outcomes highlighted the positive impact of incorporating spatial parameters on clustering socio-economic data.

Keywords: Self-organizing neural networks, SOM, Geo-SOM, Clustering, Spatial data, Census blocks, Isfahan City

*Corresponding Author

Tarigholizadeh, H., Mirbagheri, B., & Matkan, A. A. (2023). Comparing the performance of spatial and non-spatial self-organizing neural networks in the clustering of socio-economic data of Isfahan census blocks. *Geography and Environmental Planning*, 34 (4), 1 - 2 .

2252-0910 © University of Isfahan
This is an open access article under the CC BY-NC 4.0 License (https://creativecommons.org/licenses/by-nc/4.0).

22108/GEP.2023.137707.1582 20.1001.1.20085362.1402.34.4.4.5

Introduction

In recent years, there has been a significant increase in the volume of spatial data available. To gain a comprehensive understanding of spatial data, it is crucial to extract meaningful knowledge from it by considering its unique characteristics. This process, known as "Knowledge Discovery in Databases" (KDDs), utilizes methods, such as Artificial Neural Networks (ANNs) to extract useful information and knowledge. Clustering, a widely used technique for large-scale data analysis, aims to group similar data into clusters while reducing the data size and ensuring that the internal differences within clusters are significantly smaller than the differences between clusters. However, clustering algorithms for spatial data differ from those for non-spatial data. Therefore, this research focused on using self-organizing neural networks to cluster socio-economic data related to sustainable urban development while comparing the performance of two algorithms, SOM and Geo-SOM, for this task.

Methodology

This research utilized socio-economic data of Isfahan census blocks (2015) obtained from Iranian Statistics Center. The dataset consisted of 13,362 statistical blocks, each characterized by socio-economic variables. The SOM and Geo-SOM algorithms were employed to cluster the socio-economic data of census blocks and their results were compared. The self-organizing map is a single-layer feed-forward neural network that organizes output neurons in a low-dimensional topological structure. The SOM effectively maps high-dimensional data onto two-dimensional feature maps, forming networks of units or neurons. On the other hand, the Geo-SOM algorithm is an adaptation of the self-organizing map that incorporates a competitive learning process, focusing more on spatial aspects by restricting the search for the Best Matching Unit (BMU) to the geographical proximity of input patterns. Visualization tools, such as the U-matrix, Component Plane, and Hits Maps are employed to analyze the output of SOM and Geo-SOM algorithms. The network topology in this study was set to 7 x 7 after assessing multiple U-matrix sizes to identify clusters, resulting in a network with 49 neurons. Evaluation of the clustering results was performed using Silhouette coefficient, which measured similarity of the data within clusters and their dissimilarity between clusters.

Research Findings

The results obtained from this study revealed a fundamental distinction between the SOM and Geo-SOM algorithms, primarily stemming from their respective approaches to mapping input vectors onto network neurons, which in turn led to disparate clustering outcomes. Analysis of the component planes for both algorithms demonstrated notable variations in the relative weight assigned to each variable during the mapping process. Specifically, within the SOM algorithm, higher values of the variable of married population ratio tended to be associated with neurons corresponding to Cluster No. 4, while lower values exhibited a preference for neurons corresponding to Cluster No. 1. In contrast, the Geo-SOM algorithm assigned higher values of this variable to Clusters No. 8 and 9, with lower values predominantly allocated to Clusters No. 2 and 10.

The SOM algorithm yielded intertwined, heterogeneous, and scattered clusters when applied to the census block data. In contrast, the Geo-SOM algorithm achieved the classification of census block data into homogeneous and congruous clusters through the utilization of spatial parameters. Notably, the Geo-SOM algorithm exhibited a higher clustering quality as indicated by its average Silhouette coefficient of 0.27, surpassing the SOM algorithm's average Silhouette coefficient of -0.02.

Discussion of Results & Conclusion

This research employed two self-organizing map algorithms, SOM and Geo-SOM, to cluster census block data with the incorporation of spatial parameters. By integrating the geographic coordinates of census block centroids with other non-spatial inputs, we observed the positive impact of spatial parameters in the clustering process within the Geo-SOM algorithm. Comparatively, Geo-SOM exhibited superior performance, yielding cohesive and meaningful clusters of statistical blocks based on their similarities and spatial characteristics. The findings underscored the potential of the Geo-SOM algorithm in delineating homogeneous regions and identifying similar blocks within real-world datasets. We can conclude that utilization of this algorithm holds practical value for facilitating urban planning endeavors as it enables identification of homogeneous areas aligned with sustainable urban development principles and selected variables.

جغرافیا و برنامه‌ریزی محیطی
سال ۳٤، پیاپی ۹۲، شماره ٤، زمستان ۱٤۰۲، ص ۱۳۲- ۱۱۱
وصـول: ۱٤۰۲/۲/۲۵ پـذیرش: ۱٤۰۲/٤/۲٦

مقاله پژوهشی

مقایسهٔ عملکرد شبکه‌های عصبی خودسازماندهٔ مکانی و غیرمکانی در خوشه‌بندی داده‌های اجتماعی-اقتصادی بلوک‌های آماری شهر اصفهان

هادی تاریقلی زاده، دانشجوی دکتری سنجش از دور و GIS، مرکز مطالعات سنجش از دور و GIS، دانشکدهٔ علوم زمین، دانشگاه شهید بهشتی، تهران، ایران
gholizadeh.hadi20@gmail.com
بابک میرباقری *، استادیار مرکز مطالعات سنجش از دور و GIS، دانشکدهٔ علوم زمین، دانشگاه شهید بهشتی، تهران، ایران
b_mirbagheri@sbu.ac.ir
علی‌اکبر متکان، استاد مرکز مطالعات سنجش از دور و GIS، دانشکدهٔ علوم زمین، دانشگاه شهید بهشتی، تهران، ایران
a-matkan@sbu.ac.ir

چکیده

امروزه با افزایش حجم و ابعاد داده‌های مکانی و نیاز به درک کامل داده‌ها، شبکه‌های عصبی خودسازمانده به ابزاری استاندارد برای کار با داده‌های بزرگ و چندبُعدی تبدیل شده‌اند که می‌توانند در خوشه‌بندی، بصری‌سازی و انتقال داده‌های چندبُعدی در فضایی با ابعاد کمتر استفاده شوند. هدف از پژوهش حاضر، خوشه‌بندی داده‌های بلوک‌های آماری (شامل ده متغیر منتخب اجتماعی-اقتصادی مرتبط با رویکرد توسعهٔ پایدار شهری) با شبکه‌های عصبی خودسازمانده بدون استفاده از پارامترهای مکانی و به‌کارگیری مختصات جغرافیایی بلوک‌های آماری به‌عنوان پارامتر مکانی در روند خوشه‌بندی و مقایسهٔ نتایج حاصل‌شده، است. الگوریتم SOM رایج‌ترین شبکهٔ عصبی خودسازمانده و الگوریتم Geo-SOMمکانی‌شدهٔ الگوریتم SOM است. روند خوشه‌بندی هر دو الگوریتم یکسان است و تنها تفاوت این دو الگوریتم به‌کارگیری پارامترهای مکانی در روند اجرای الگوریتم Geo-SOM است. در پژوهش حاضر داده‌ها با الگوریتم SOM و Geo-SOM خوشه‌بندی شده است. نتایج نشان داد که خوشه‌های حاصل از دو الگوریتم به‌طور کامل، متفاوت است. خوشه‌بندی بلوک‌های آماری بدون توجه به خصوصیات مکانی و تنها با استفاده از معیار شباهت، منجر به خوشه‌های ناهمگن می‌شود و برعکس. با اعمال پارامترهای مکانی نه‌تنها از معیار شباهت، از ویژگی‌های مکانی داده‌ها نیز در فرآیند خوشه‌بندی استفاده می‌شود که این مسئله منجر به تولید خوشه‌های همگن می‌شود. ارزیابی نتایج با استفاده از ضریب سیلهوته بیانگر خوشه‌بندی مناسب‌تر الگوریتم Geo-SOM است؛ به‌طوری که میانگین ضریب سیلهونه برای الگوریتم SOM برابر ۰/۰۲- و برای الگوریتم Geo-SOM برابر ۰/۲۷ است. مقایسهٔ نتایج نشان‌دهندهٔ تأثیر مثبت پارامترهای مکانی در خوشه‌بندی داده‌های اجتماعی و اقتصادی است.

واژه‌های کلیدی: شبکه‌های عصبی خودسازمانده، Geo-SOM، خوشه‌بندی، داده‌های مکانی، بلوک‌های آماری، شهر اصفهان.

*نویسنده مسؤول

تاریقلی زاده، هادی، میرباقری، بابک، متکان، علی اکبر. (۱٤۰۲). مقایسه عملکرد شبکه های عصبی خود سازمانده مکانی و غیرمکانی در خوشه بندی داده های اجتماعی- اقتصادی بلوک های آماری شهر اصفهان. *جغرافیا و برنامه ریزی محیطی*، ۳٤ (٤)، ۱۳۲- ۱۱۱ .

2252-0910 © University of Isfahan
This is an open access article under the CC BY-NC 4.0 License (https://creativecommons.org/licenses/by-nc/4.0).

10.22108/GEP.2023.137707.1582 20.1001.1.20085362.1402.34.4.4.5

مقدمه

در سال‌های اخیر میزان داده‌های مکانی دردسترس به‌شدت افزایش یافته است. برای درک کامل داده‌های مکانی لازم است دانشی ساده و عمومی از داده‌های مکانی با درنظرگرفتن خصوصیات ویژهٔ مکان استخراج شود. به فرآیند شناسایی الگوهای معتبر، جدید، مفید و قابل فهم در داده‌ها «کشف دانش در پایگاه داده» (KDD) (Knowledge Discovery in Database) گفته می‌شود (Fayyad, Piatetsky-Shapiro, & Smyth, 1996). کشف دانش جغرافیایی (GKD) (Geographic Knowledge Discovery) به‌معنای کشف دانش از روابط بین پدیده‌های جغرافیایی است که زیرمجموعهٔ KDD است. یک مؤلفهٔ اصلی در این فرآیند داده‌کاوی مکانی (SDM) (Spatial Data Mining) است که الگوها و روابط ناشناخته را در پایگاه داده‌های مکانی کشف می‌کند (Klösgen & Zytkow, 1996). باتوجه به اهمیت و نیاز گسترده به دانش استخراج‌شده از داده‌های مکانی می‌توان گفت که ۱- داده‌های با ابعاد بالا می‌تواند برای تفسیر و تجزیه‌وتحلیل چالش‌برانگیز باشد؛ ۲- میزان محاسبات با ابعاد داده‌ها در ارتباط و تجزیه‌وتحلیل داده‌های با ابعاد بالا ازنظر محاسباتی پرهزینه است؛ ۳- روابط این داده‌های با ابعاد بالا به‌راحتی قابل تجسم نیست. خوشه‌بندی یکی از روش‌های استاندارد برای کار با مجموعه داده‌های بزرگ است؛ زیرا حجم داده‌ها را با یکی‌کردن داده‌های مشابه در گروه‌ها (خوشه‌ها) کاهش می‌دهد؛ به‌طوری که تفاوت داخلی داده‌ها به‌طور چشمگیری، کمتر از تفاوت‌های بین خوشه‌ای می‌شود (Andrienko et al., 2010). خوشه‌بندی ازطرفی، یکی از اساسی‌ترین مسائل یادگیریِ ماشینهٔ نظارت‌نشده و از طرف دیگر، روشی مفید در استخراج دانش از داده‌های مکانی است که هدف اصلی آن تفکیک داده‌ها در خوشه‌هایی با ویژگی‌های مشابه است (Jain, Murty, & Flynn, 1999). ساختار داده‌ها با خوشه‌بندی سازماندهی و تجزیه‌وتحلیل و اکتشاف داده‌ها تسهیل می‌شود.

داده‌های مکانی اغلب خودهمبستگی مکانی (Spatial Autocorrelation) دارند که یک مفهوم اساسی در علوم مکانی است (Sui, 2004) و بدین معناست که پدیده‌های نزدیک شباهت‌های بیشتری نسبت به پدیده‌هایی دارند که از یکدیگر فاصله دارند (Tobler, 1970). خوشه‌بندی مکانی، سازماندهی اشیا مکانی در خوشه‌هاست؛ به‌گونه‌ای که اشیا داخل یک خوشه نه‌تنها ازنظر ویژگی‌های اشیا، ازنظر مکانی نیز به یکدیگر مشابه هستند و با اشیا داخل خوشه‌های دیگر تفاوت دارند (Miller, 2010). الگوریتم خوشه‌بندی داده‌های مکانی، دراساس با داده‌های غیرمکانی متفاوت است (Grubesic, Wei, & Murray, 2014). یک ویژگی اساسی داده‌های مکانی این است که مشاهده‌ها بیشتر وابسته به مکان است (Sui, 2004). اگر داده‌های مکانی با نادیده‌گرفتن وابستگی مکانی خوشه‌بندی شوند، ممکن است منجر به خوشه‌بندی نامناسب داده‌ها شوند (Openshaw, 1999). محققان به الگوریتم‌های خوشه‌بندی مکانی در طی سالیان توجه کرده‌اند؛ زیرا تعداد الگوریتم‌هایی که وابستگی مکانی را در روند داده‌ها لحاظ کرده‌اند، محدود است.

داده‌کاوی مکانی با به‌کارگیری شبکه‌های عصبی مصنوعی (ANN) (Artificial Neural Networks) و پایگاه داده‌های مکانی (Spatial Database)، الگوها و روابط ناشناخته در داده‌ها را کشف می‌کند تا اطلاعات را استخراج و درنهایت، آنها را به دانش جدید و بالقوه مفید تبدیل کند (Yuan et al., 2004). مفهوم ANN از زمان معرفی اولیه در

دههٔ 1980 میلادی گسترش و تنوع زیادی یافته است؛ به‌طوری که امروزه به یکی از حوزه‌های بسیار فعّال پژوهشی تبدیل شده است. با افزایش حجم و ابعاد داده‌ها و نیاز به توان محاسباتی بالا و دقیق، شبکه‌های عصبی مصنوعی به ابزاری استاندارد برای حل مسائل خوشه‌بندی و طبقه‌بندی تبدیل شده‌اند.

یکی از انواع شبکه‌های عصبی مصنوعی، نقشهٔ خودسازمانده (Self-Organizing Map) (SOM) است که کوهونن در آغاز دههٔ 1980 میلادی پیشنهاد کرد(Kohonen, 1995). الگوریتم SOM می‌تواند برای خوشه‌بندی، بصری‌سازی و انتقال داده‌های چندبُعدی در فضایی با ابعاد کمتر استفاده شود (Stefanovic & Kurasova, 2018). نقشهٔ خودسامانده (SOM) متداول‌ترین و رایج‌ترین الگوریتم برای خوشه‌بندی داده‌ها با ابعاد بالا و یا چندبُعدی است (Galutira, Fajardo, & Medina, 2019). الگوریتم SOM یک شبکهٔ عصبی بدون نظارت است که داده‌های با ابعاد بالا یا چندبُعدی را در یک فضای تک‌بُعدی، دوبُعدی و در مواردی خاص سه‌بُعدی به نام نقشه نگاشت می‌کند (Wankhede, 2014). عنصر اصلی الگوریتم SOM رقابت انعطاف‌پذیر گره‌ها در لایهٔ خروجی است که در هر تکرار، گره برنده نه‌تنها به‌روز می‌شود، همسایگان آن نیز تعدیل می‌شود. شبکهٔ SOM می‌تواند به‌گونه‌ای تنظیم شود که داده‌های گروهی و کلاس‌های به‌نسبت مشابه با یکدیگر را شناسایی کند (Sheela & Deepa, 2012). باکائو، لوبو و پاینهو الگوریتم Geo-SOM را در سال 2005 معرفی کردند که انطباق‌یافتهٔ الگوریتم SOM و مبتنی بر فرآیند یادگیری رقابتی است و جنبه مکانی دارد؛ زیرا جست‌وجو را به مجاورت جغرافیایی الگوهای ورودی محدود می‌کند (Bação, Lobo, & Painho, 2004; Hagenauer & Helbich, 2016). مهم‌ترین مزیت دو الگوریتم SOM و Geo-SOM درمقایسه با سایر روش‌های داده‌کاوی این است که این الگوریتم‌ها نه‌تنها نتیجه را به‌صورت برآوردهای عددی (بیشتر روش‌های داده‌کاوی) به دست می‌آورند، نتیجه را به‌صورت بصری نیز ارائه می‌کنند (Stefanovic & Kurasova, 2011). بصری‌سازی به محققان اجازه می‌دهد تا خوشه‌ها و روابط بین داده‌ها را ببینند و نتایج خوشه‌بندی را تحلیل و تفسیر کنند (Zhang & Fang, 2012).

پیشینهٔ پژوهش

پژوهش‌های متعدّدی در موضوعات مختلف برای خوشه‌بندی با استفاده از الگوریتم‌های SOM انجام شده است. (Hagenauer & Helbich, 2016)پژوهشی با عنوان «SPAWNN ابزاری برای تحلیل فضایی با شبکه‌های عصبی خودسازمانده» انجام داده‌اند. محققان با به‌کارگیری الگوریتم‌های CNG (Contextual Neural Gas) و Geo-SOM داده‌های اجتماعی–اقتصادی بلوک‌های سرشماری سال 2010 میلادی را در ایالات متحده با متغیرهای قومیت، سن، مسکن و خانوارهای شهر شیکاگو خوشه‌بندی کردند. نتایج پژوهش بینش‌های مهمی را دربارهٔ ویژگی‌های اجتماعی–اقتصادی شهر شیکاگو نشان داد که این شهر به‌دلیل تنوع اجتماعی–اقتصادی با تفکیک شدید قومیتی در سراسر مناطق شهری مواجه است.

لی و همکاران در پژوهشی با عنوان «استفاده از نقشهٔ خودسازماندهی برای طبقه‌بندی کیفیت آب ساحلی: به‌سوی درک بهتر الگوها و فرآیندها» الگوریتم SOM را برای طبقه‌بندی کیفیت آب‌های ساحلی به کار بردند. در این پژوهش

برای بررسی ویژگی‌های فضایی کیفیت آب در ناحیهٔ ساحلی فوجیان، ۱۹ متغیر کیفی آب در ۹٤ محل نمونه‌برداری‌شده بررسی شد. در این پژوهش برای نشان‌دادن الگوهای متغیرهای کیفیت آب از صفحات مؤلفهٔ SOM (Component Planes) استفاده شد که در آن الگوهای مشابهی برای متغیرهایی که با یکدیگر در ارتباط بودند، مشاهده شد که نشان‌دهندهٔ یک منبع مشترک بود. نتایج نشان داد که الگوریتم SOM یک ابزار مؤثر برای درک بهتر الگوها و فرآیندهای محرک کیفیت آب است(Li et al., 2018).

لیائو وهمکاران در پژوهشی با عنوان «کاوش پایگاه داده یک بررسی محیطی خاک با استفاده از یک نقشهٔ جغرافیایی خودسازماندهی: یک مطالعهٔ آزمایشی» الگوریتم SOM را برای خوشه‌بندی پایگاه دادهٔ زیست‌محیطی خاک به کار بردند. مجموعه داده‌ها با روش نمونه‌برداری منظم شبکه‌ایی از ۵ فلز سنگین و نمونه‌های خاک نیز از ۵ عنصر متنوع با چولگی بالا تشکیل شده بود. As و Pb الگوی توزیع سطحی و صفحات مؤلفهٔ مشابهی داشتند که نشان‌دهنده همبستگی بالای این دو عنصر است؛ به این معنا که آنها در محیط خاک نیز رفتار مشابهی دارند (Liao et al., 2019).

بررسی مطالعات پیشین نشان‌دهندهٔ عملکرد مناسب شبکه‌های عصبی خودسازمانده در خوشه‌بندی داده‌هاست. در این مطالعات محققان باتوجه به اهمیت داده‌کاوی و دانش به‌دست‌آمده از تحلیل داده‌ها و توانایی‌های الگوریتم‌های شبکه‌های عصبی در داده‌کاوی، جنبه‌های مختلف الگوریتم‌های شبکه‌های عصبی خودسازمانده را بررسی کرده‌اند.

کمیسیون جهانی محیط‌زیست و توسعه (World Commission on Environment and Development) (WCED) مفهوم توسعهٔ پایدار را به‌عنوان توسعه‌ای که نیازهای نسل حاضر را بدون به خطر انداختن توانایی نسل‌های آینده برای برآوردن نیازهای خود مرتفع می‌کند، تعریف کرده است (Aldegheishem, 2014). پایداری شهری شامل ابعاد فیزیکی، اقتصادی و فرهنگی است (Rodrigues & Franco, 2020). هدف اصلی توسعهٔ پایدار، تأمین نیازهای اساسی، بهبود و ارتقای سطح زندگی برای همه، حفظ و ادارهٔ بهتر اکوسیستم‌ها و آینده‌ای امن‌تر و سعادتمند است. توسعهٔ پایدار شهری فرآیندی است که طی آن گردش انرژی در شهر در شرایط حداقلی عملکردی، بیشترین کارایی را دارد و توزیع یکسانی را در به حرکت در آوردن عناصر به‌عنوان یک مجموعهٔ یکپارچه و متصل ایجاد می‌کند (تقوایی و صفرآبادی، ۱۳۹۲). باوجود مزایای فراوان به‌کارگیری شبکهٔ عصبی خودسازمانده در خوشه‌بندی داده‌های با ابعاد بالا و قدرت تفسیرپذیری آن، بررسی پیشینهٔ پژوهش نشان می‌دهد که در پژوهش‌های داخلی به استفاده از الگوریتم‌های فوق در خوشه‌بندی داده‌های اجتماعی–اقتصادی مرتبط با توسعهٔ پایدار شهری توجه نشده است؛ بنابراین در پژوهش حاضر سعی بر آن است که علاوه‌بر استفاده از شبکه‌های عصبی خودسازمانده برای خوشه‌بندی داده‌های اجتماعی–اقتصادی مرتبط با توسعهٔ پایدار شهری، عملکرد دو الگوریتم SOM و Geo-SOM در خوشه‌بندی داده‌های فوق نیز مقایسه شود.

منطقهٔ مطالعه‌شده

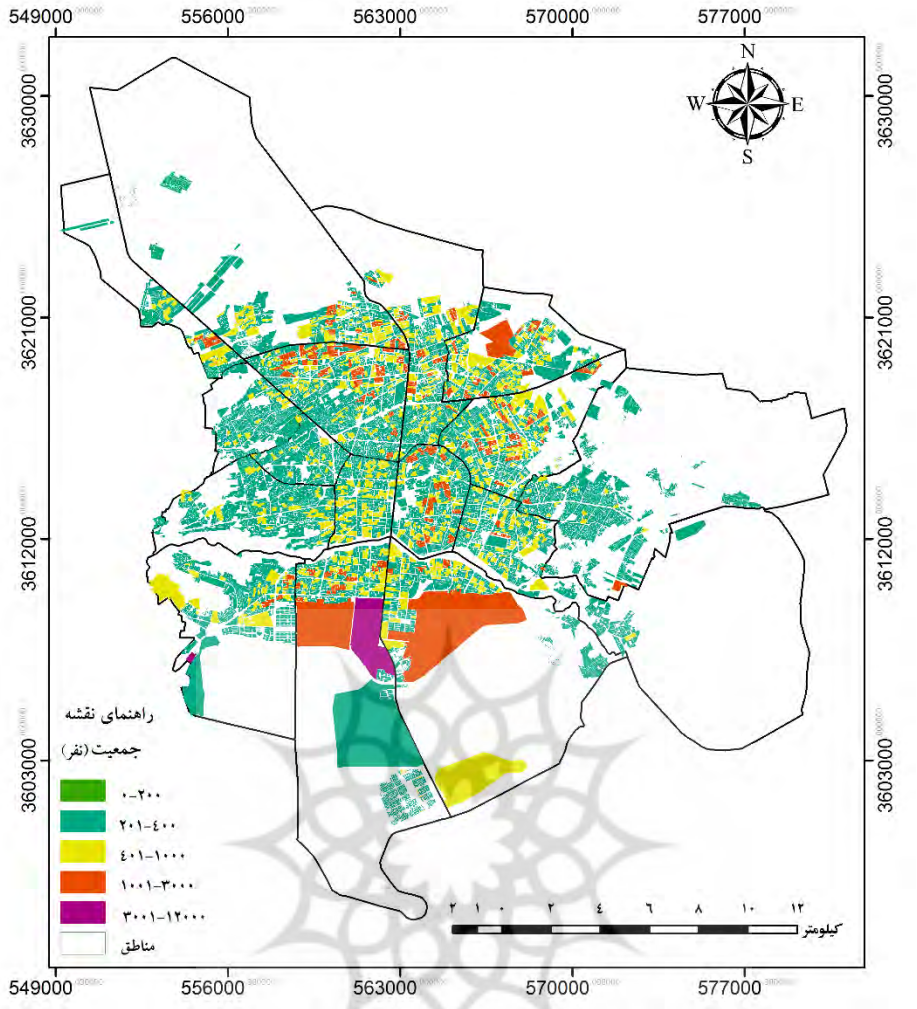
کلانشهر اصفهان سومین شهر پرجمعیت ایران با طول جغرافیایی ٥٤ درجه و ۳۹ دقیقه و ٤۰ ثانیهٔ شرقی و عرض جغرافیایی ۳۵ درجه و ۳۸ دقیقه و ۳۰ ثانیهٔ شمالی است. این کلانشهر ۱۵ منطقهٔ شهرداری دارد و بیشترین تراکم

جمعیت آن به‌ترتیب متعلق به مناطق ۸ و ۱۰ است. این مناطق در سال‌های گذشته رشد فراوانی داشته که باعث افزایش تراکم آنها شده است. همچنین، در مناطق ۱ و ۲ که درواقع، بخش مرکزی شهر اصفهان را تشکیل می‌دهند، به‌علت قدمت بافت و تخریب فضای مسکونی، تراکم جمعیت آنها نسبت به متوسط تراکم کل مناطق بسیار کمتر شده است. شکل ۱ مناطق پانزده‌گانهٔ شهر اصفهان را به همراه بلوک‌های آماری و آمار جمعیتی آنها نمایش می‌دهد.

داده‌های پژوهش

داده‌های استفاده‌شده در پژوهش حاضر، داده‌های بلوک‌های آماری سرشماری جمعیت سال ۱۳۹۵ کلانشهر اصفهان است که مرکز آمار ایران ارائه کرده است. تعداد کل بلوک‌های آماری با جمعیت، برابر ۱۳۳۶۲ عدد است که متغیرهای متعدّد اجتماعی-اقتصادی دارد. در این متغیرها بلوک‌های آماری براساس ویژگی‌هایی چون جمعیت کل، سطح تحصیلات، فعالیت‌های اجتماعی-اقتصادی و ویژگی ساختمان‌ها توصیف می‌شود. همان‌طور که پیشتر نیز اشاره شد، هدف از پژوهش حاضر خوشه‌بندی بلوک‌های جمعیتی براساس متغیرهای اقتصادی-اجتماعی مرتبط با توسعهٔ پایدار شهری است. براساس مرور پژوهش‌های صورت‌گرفته دربارهٔ متغیرهای مرتبط با توسعهٔ پایدار شهری، ۱۰ متغیر برای خوشه‌بندی بلوک‌های آماری انتخاب شد که عبارت است از: تراکم جمعیت در فضا (سپهوند و عارف‌نژاد، ۱۳۹۲)، نسبت جمعیت متأهل (موسوی، ۱۳۹۷)، تعداد مهاجران واردشده (احمدی، حسینی‌فر و نصیری هندخاله، ۱۳۹۵)، میزان اشتغال (توده فلاح و همکاران، ۱۳۹۷)، نسبت جمعیت جوان (Lez'er et al., 2019)، متوسط مساحت واحد مسکونی (نصیری دارانی، ۱۴۰۱)، تعداد واحد مسکونی آپارتمانی (محمدزاده، ۱۳۹۴)، میزان استحکام بنا (Patel & Patel, 2021)، نسبت باسوادی (برقی، ۱۳۹۷) و نسبت محصلان (نصیری دارانی، ۱۴۰۱).

بر این اساس، ستون‌های مرتبط با متغیرهای ده‌گانهٔ فوق از داده‌های بلوک‌های آماری انتخاب و درصورت لزوم نسبت آنها به جمعیت کل برای هر بلوک جمعیتی محاسبه شده است. درنهایت، متغیرهای محاسبه‌شده برای هم‌مقیاس‌سازی و ورود به الگوریتم‌های لازم براساس میانگین و انحراف معیار استانداردسازی شد. در پژوهش حاضر علاوه‌بر متغیرهای فوق از مختصات جغرافیایی مراکز هندسی بلوک‌های آماری نیز به‌عنوان پارامتر مکانی استفاده شده است.



شکل ۱: نقشهٔ شهر اصفهان به تفکیک مناطق و بلوک‌های آماری (منبع: نویسندگان ۱٤۰۲)

Figure 1: Regions and Census Blocks of Isfahan City

روش‌شناسی پژوهش

در پژوهش حاضر از دو الگوریتم SOM و Geo-SOM برای خوشه‌بندی داده‌های اجتماعی-اقتصادی بلوک‌های آماری استفاده و نتایج حاصل با یکدیگر مقایسه شده است.

الگوریتم نقشه‌های خودسازمانده (SOM)

نقشهٔ خودسازمانده یک شبکهٔ پیشخور تک‌لایه است که در آن نورون‌های خروجی در یک ساختار توپولوژیکی با ابعاد کم مرتب شده است. با الگوریتم SOM داده‌های با ابعاد بالا به نقشه‌های ویژگی دوبُعدی شبکهٔ نگاشت می‌شوند (شکل۲). نقشه‌های خودسازمانده سعی می‌کند که در مرحلهٔ نگاشت، روابط توپولوژیکی را حفظ کند؛ یعنی الگوهایی را که در فضای ورودی نزدیک به یکدیگر هستند به واحدهایی که در فضای خروجی نزدیک هستند، نگاشت می‌کند و برعکس. برای نگاشت به هر نورون یک بردار وزنی برابر با ابعاد فضای ورودی وجود دارد؛ یعنی هر واحد i یک

بردار وزنی متناظر با بُعد d $w_i=\{w_{i1}, w_{i2},\ldots..,w_{id}\}$ دارد. ابعاد فضـای ورودی به‌طور معمول، از ابعاد شـبکهٔ خروجی بیشتر است. بیشتر الگوریتم‌های SOM در یک شبکهٔ مستطیلی از واحدها پیاده‌سازی می‌شوند. ایدهٔ اصلی این است که درطول آموزش، هر داده با تمام واحدها مقایسـه و شـبیه‌ترین آنها که به‌عنوان بهترین واحد تطبیق (Best Matching Unit) (BMU) شـناخته می‌شود، انتخاب شود. سپس BMU و همسایگان آن در شبکه به‌روزرسانی شوند تا به آن دادهٔ خاص نزدیک‌تر شـوند. درحین نگاشـت تنها یک نورون برنده وجود خواهد داشت و آن نورونی است که بردار وزنی آن نزدیک‌ترین فاصـله را با بردار ورودی داشـته باشـد که با محاسـبهٔ فاصلهٔ اقلیدسی بین بردار ورودی و بردار وزنی تعیین می‌شود.

روند اصلی آموزش شبکه‌های عصبی خودسازمانده به‌صورت زیر است (Kohonen, 1995).

۱. در وزن‌های اولیهٔ مقداردهی شـده بیشـترین مقدار برای شعاع تابع همسایگی R و میزان یادگیری α تعیین و تا زمانی که شرایط توقف برقرار نیست، مراحل ۲ تا ۸ انجام می‌شود؛

۲. فاصلهٔ اقلیدسی برای هر نورون j محاسبه می‌شود؛

$$D(j)=\sqrt{\sum_{i=1}^{n}(x_i-w_{ij})^2} \qquad \text{رابطهٔ (۱)}$$

x_i **مؤلفهٔ** ام جزء بردار ورودی x است و w_{ij} پیوند وزنی x_i در نورون واقع در (ij) است.

۳. نزدیک‌ترین نورون به‌عنوان نورون برنده انتخاب می‌شود؛

$$D(j)=min \qquad \text{رابطهٔ (۲)}$$

فاصلهٔ اقلیدسی نورون کمترین مقدار باشد.

٤. هر نرون j باتوجه به همسایگی آن با نورون i به‌روزرسانی می‌شود؛

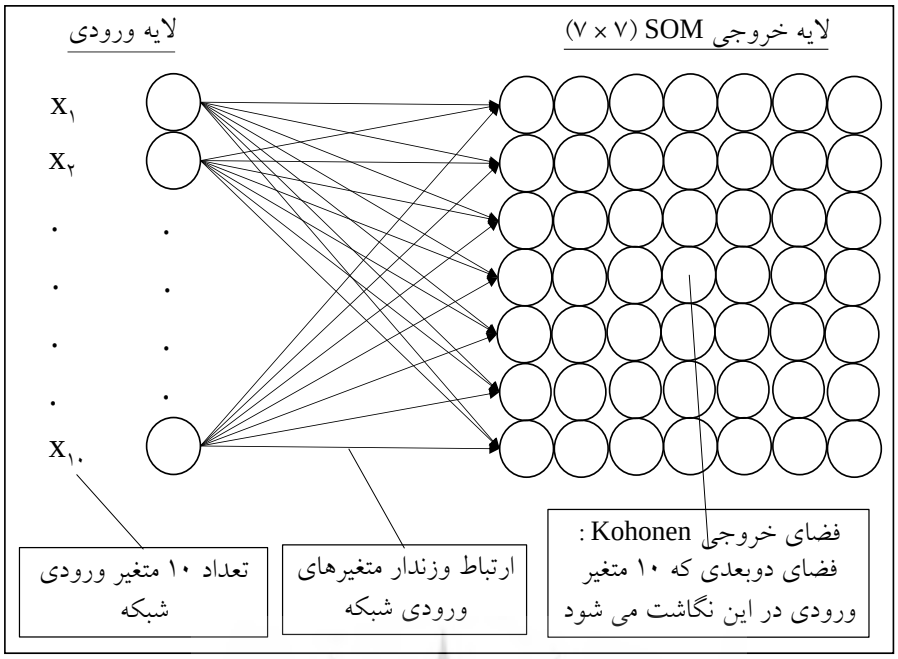
$$w_{ij}(new)=w_{ij}(old)+\alpha(x_i-w_{ij}(old)) \qquad \text{رابطهٔ (۳)}$$

۵. میزان یادگیری α به‌روزرسانی می‌شود؛

٦. شعاع عملکرد تابع همسایگی توپولوژیک کاهش پیدا می‌کند؛

۷. بررسـی شـرایط توقف تا زمانی تکرار می‌شود که معیار توقف به دست آید. به‌طور معمول، معیار توقف، تعداد ثابتی از تکرار را دارد.

برای هم‌گرایی و ثبات نقشـه، میزان یادگیری و شـعاع همسایگی در هر تکرار کاهش می‌یابد؛ بنابراین هم‌گرایی به سمت صفر میل خواهد کرد.

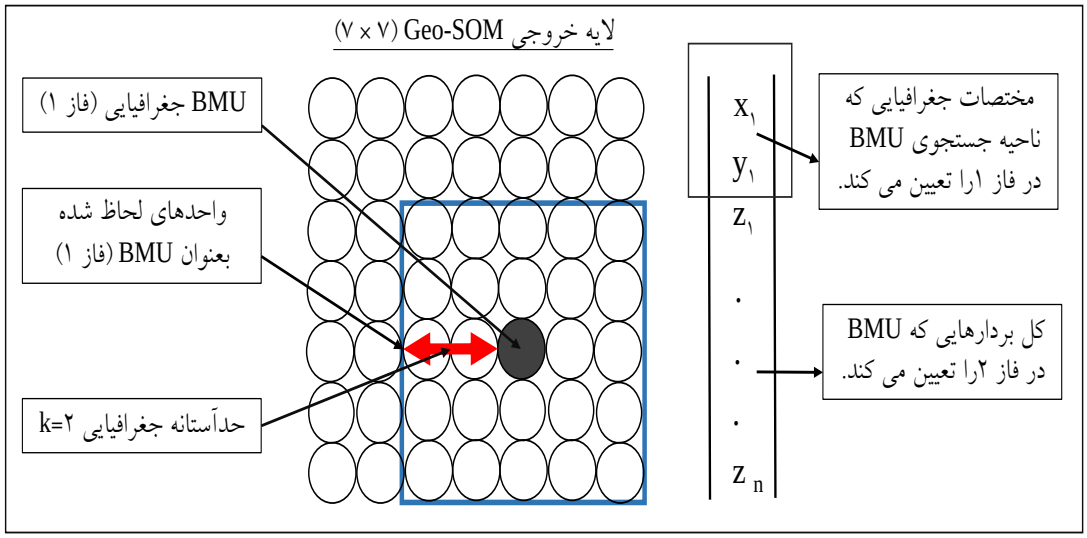


شکل ۲: دیاگرام الگوریتم SOM (منبع: نویسندگان ۱٤۰۲)

Figure 2: Diagram of Self-Organizing Map Algorithm

الگوریتم Geo-SOM

الگوریتم Geo-SOM توزیع واحدهای SOM را نه‌تنها باتوجه به شباهت بین همسایگان، باتوجه به تراکم جغرافیایی بردارهای ورودی نیز ایجاد می‌کند (Bação et al., 2004). در الگوریتم آموزشی SOM، مهم‌ترین مرحله در تعیین اینکه کدام الگوها با یکدیگر خوشه‌بندی شوند، مرحلهٔ انتخاب بهترین واحد تطبیق (BMU) است؛ بنابراین با تغییر نحوهٔ انتخاب BMU می‌توان به مختصات جغرافیایی اهمیت بیشتری داد. در الگوریتم Geo-SOM برای قراردادن BMU در مجاورت جغرافیایی الگوی ورودی، فرآیند جست‌وجوی انتخاب بهترین واحد تطبیق (BMU) در دو مرحله انجام می‌شود. در مرحلهٔ اول، همسایگی جغرافیایی که امکان جست‌وجو برای BMU در آن وجود دارد، انتخاب و در مرحلهٔ دوم با استفاده از سایر متغیرها جست‌وجوی نهایی انجام می‌شود. این همسایگی جغرافیایی با پارامتر k (حد آستانهٔ جغرافیایی: Geographic Tolerance) که در فضای خروجی نقشهٔ SOM تعریف شده است، کنترل می‌شود (شکل ۳). در حالت $k = 0$، الگوریتم نزدیک‌ترین واحد را ازلحاظ جغرافیایی به‌عنوان BMU انتخاب می‌کند. با تنظیم k برابر با اندازهٔ نقشهٔ SOM، مختصات جغرافیایی نادیده گرفته می‌شود. وقتی $k = 0$ است مکان‌های نهایی بردارها در فضای ورودی به‌طور تقریبی، متناسب با مکان‌های جغرافیایی بردارهای آموزشی خواهد بود (Bação et al., 2004).



شکل ۳: دیاگرام الگوریتم Geo-SOM (منبع: نویسندگان ۱۴۰۲)

Figure 3: Diagram of Geo-SOM Algorithm

ابزارهای بصری‌سازی در الگوریتم SOM و Geo-SOM

برای تجزیه‌وتحلیل خروجی الگوریتم‌های SOM و Geo-SOM از ابزارهای بصری‌سازی (Visualization) متعدّدی استفاده می‌شود که عبارت است از:

ماتریس U (U-Matrix)

یک راه به‌نسبت ساده برای شناسایی خوشه‌های SOM و Geo-SOM استفاده از ماتریس U (ماتریس فاصلهٔ یکپارچه) (Unified Distance Matrix) است. توپولوژی پیش‌فرض نقشه‌های خودسازمانده، شش‌ضلعی است. این شش‌ضلعی‌ها، مکان نورون‌ها و روابط همسایگی آنها را با نورون‌های دیگر در توپولوژی شبکه نشان می‌دهد که ماتریس U نامیده می‌شود. فاصلهٔ بین گره‌های همسایه در ماتریس U باتوجه به فاصلهٔ اقلیدسی وزن‌دار می‌شوند و با رنگ روی نقشه ترسیم می‌شوند (Löhr et al., 2010). رنگ‌های سردتر (آبی) نورن‌های همسایه را که نزدیک هستند (شباهت دارند) نشان می‌دهد و رنگ‌های داغ‌تر (قرمز) نورن‌های همسایه را که فاصلهٔ زیاد دارند (متفاوت هستند)، نشان می‌دهد. سرانجام، می‌توان گفت در ماتریس U گره‌های با فاصلهٔ کم به یک خوشه تعلق دارند و گره‌های با فاصلهٔ زیاد مرزهای خوشه‌های مختلف را نشان می‌دهند (Park et al., 2018). بنابراین با استفاده از ماتریس U می‌توان نتایج خوشه‌بندی شبکه را تحلیل و ارزیابی کرد.

صفحات مؤلفه‌ها (Component Planes)

صفحهٔ مؤلفه، وزن نسبی هریک از مؤلفه‌های (متغیرهای) بردارهای ورودی را به‌طور جداگانه، نشان می‌دهد (Kohonen, 1995). تعداد صفحات مؤلفه‌ها در هر الگوریتم SOM و Geo-SOM برابر با اندازهٔ ابعاد دادهٔ ورودی است؛ بنابراین برای هر متغیر از بردار ورودی یک صفحهٔ مؤلفه وجود دارد (در این پژوهش ۱۰ صفحه برای الگوریتم

SOM و ۱۲ صفحه برای الگوریتم Geo-SOM ایجاد می‌شود). وزن‌های هریک از بردارهای ورودی را می‌توان با استفاده از صفحهٔ مؤلفه بصری‌سازی کرد. در این صفحات، رنگ‌های روشن‌تر و تیره‌تر به‌ترتیب نشان‌دهندهٔ وزن‌های بزرگ‌تر (نگاشت‌شدن متغیرهایی با مقدارهای بیشتر) و کوچک‌تر (نگاشت‌شدن متغیرهایی با مقدارهای کمتر) است. با مقایسهٔ صفحات مؤلفه می‌توان الگوهای اتصال و ارتباط مؤلفه‌ها را با یکدیگر بررسی کرد. اگر صفحات مؤلفه‌ها با یکدیگر مشابه باشد، بردارهای ورودی همبستگی بالایی دارند و برعکس. صفحات مؤلفه را می‌توان براساس شاخص شباهت آنها مقایسه و متغیرهای مشابه و با همبستگی و غیرمشابهِ بدون همبستگی را از یکدیگر جدا کرد (Löhr et al., 2010).

Hits Map

Hits Maps نقشهٔ مفید دیگری است که نشان می‌دهد یک نورون شبکه به‌عنوان BMU چند بار انتخاب شده است و یا به عبارت دیگر، چند دادهٔ ورودی به هر نورون نگاشت شده است. همچنین، با این نقشه می‌توان توزیع بردارهای ورودی را درسطح نورون‌های شبکه بررسی کرد که بهتر است داده‌ها به‌طور یکنواخت، در سراسر نورون‌های شبکه توزیع شده باشد (Kassambara, 2017).

نقشهٔ خوشه‌ها (Clusters Map)

در الگوریتم SOM به‌طور معمول، تعداد نورون‌ها بسیار بیشتر از تعداد خوشه‌های واقعی است. در این الگوریتم تعداد تقریبی خوشه‌ها را می‌توان با استفاده از ماتریس U و براساس فاصله‌های میان نرون‌ها تعریف کرد. گفتنی است در اینجا منظور از نقشهٔ خوشه‌ها نقشه‌ای مبتنی بر ماتریس U است که در آن شمارهٔ خوشه برای هر نورون تعیین شده باشد. براساس این نقشه می‌توان نقشهٔ جغرافیایی بلوک‌های آماری را تولید کرد که در آن برای هر بلوک، شمارهٔ خوشهٔ مربوط به آن مشخص شده باشد.

اندازهٔ نقشه

اندازهٔ نقشه (تعداد نورون‌های خروجی) برای اجرای SOM حیاتی است؛ با این حال هیچ قانون دقیقی برای تعیین اندازهٔ نقشه وجود ندارد. الهونیمی و وسانتو در سال 2000 میلادی روشی پیشنهاد کردند که تعداد بهینهٔ نورون‌ها نزدیک به $5\sqrt{n}$ است (n تعداد نمونه‌های مطالعه‌شده است). اگر این فرض صحیح باشد، اندازهٔ بهینهٔ نقشه در مطالعهٔ حاضر باید برابر با ۵۷۸ نورن باشد که عدد بزرگی است. آزمون‌های متعدّد در پژوهش حاضر نشان داد اگر اندازه نقشهٔ خیلی کوچک باشد، خوشه‌ها تفکیک‌ناپذیر می‌شوند یا اگر خیلی بزرگ باشد، نه‌تنها خوشه‌ها تفکیک‌پذیر نخواهند بود، در بعضی از نورون‌ها نیز بردار ورودی نگاشت نمی‌شود؛ بنابراین اگر اندازهٔ نقشه باتوجه به ابعاد داده‌ها و تعداد متغیرهای مطالعه‌شده پس از بررسی اندازه‌های مختلف انتخاب شود، تفکیک خوشه‌ها دقیق‌تر و به‌تبع آن در همهٔ نورون‌ها بردار ورودی نگاشت خواهند شد. در پژوهش حاضر پس از بررسی اندازه‌های مختلف، نقشه‌ای به اندازهٔ ٤۹ نورون یا شبکه (۷*۷) انتخاب شد(Vesanto & Alhoniemi, 2000).

ارزیابی نتایج خوشه‌بندی

در پژوهش حاضر برای ارزیابی نتایج خوشه‌بندی از ضریب سیلهوته (Silhouette) استفاده شد. این ضریب برای یک خوشه، درجهٔ شباهت داده‌های داخل یک خوشه و شباهت‌نداشتن با خوشه‌های دیگر را اندازه‌گیری می‌کند (Kaufman & Rousseeuw, 2009). این ضریب برای هر نمونه، هر کلاس و مجموعه کل داده‌ها محاسبه می‌شود. ضریب سیلهوته براساس دوری و نزدیکی نمونه‌ها و خوشه‌ها به یکدیگر و با استفاده از رابطهٔ ٤ قابل محاسبه است.

رابطهٔ (٤)

$$S(i) = \frac{(b(i) - a(i))}{max[a(i), b(i)]}$$

در این رابطه $S(i)$ مقدار ضریب سیلهوته محاسبه‌شده برای نمونهٔ (i) است. در این رابطه $a(i)$ میانگین عدم تشابه (Dissimilarity) بین مشاهدهٔ i با سایر مشاهده‌ها در یک خوشهٔ مشابه و $b(i)$ کمینهٔ میانگین عدم تشابه مشاهدهٔ i نسبت به تمام مشاهده‌ها در خوشه‌های دیگر است. براساس رابطهٔ فوق مقدار $S(i)$ بین ۱− و ۱+ قرار دارد. اگر $S(i)$ به ۱+ نزدیک‌تر باشد، به این معناست که خوشهٔ پیشنهادشده برای نمونهٔ لازم مناسب و چنانچه $S(i)$ به ۱− نزدیک‌تر باشد، به این معناست که خوشهٔ پیشنهادشده برای نمونهٔ لازم نامناسب است.

کیفیت کلی خوشه‌بندی را می‌توان با استفاده از میانگین ضریب سیلهوته برای کل مجموعه داده اندازه‌گیری کرد که براساس رابطهٔ ۵ تعریف می‌شود.

رابطهٔ (۵)

$$SC = \frac{1}{n}\sum_{i=1}^{n} S(i)$$

در این رابطه n تعداد کل نمونه‌هاست و مقدار بالاتر SC نشان‌دهندهٔ خوشه‌بندی مناسب است (Hsu & Li, 2010).

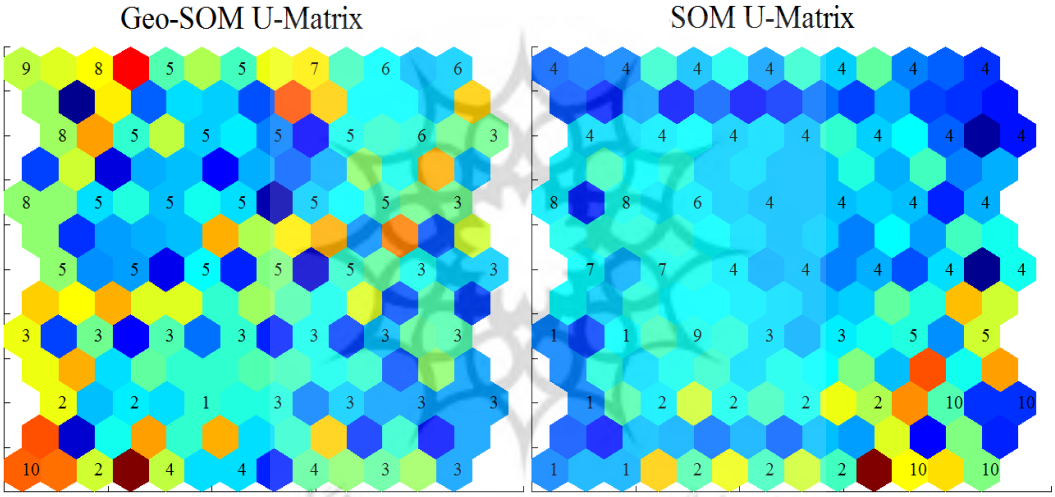
یافته‌های پژوهش و تجزیه‌وتحلیل

ماتریس U

شش‌ضلعی‌های حاوی اعداد در این ماتریس نشان‌دهندهٔ مراکز نورون‌های شبکه است. همچنین، این اعداد نشان می‌دهد که نورون‌های برنده‌شده متعلق به کدام خوشه است. همان‌طور که اشاره شد، توپولوژی شبکه در این مطالعه، پس از بررسی اندازه‌های متعدّد ماتریس U برای شناسایی خوشه‌ها، ۷ در ۷ انتخاب شده است؛ بنابراین ٤۹ نورون در شبکه وجود دارد. همان‌طور که اشاره شد، فاصلهٔ بین نورن‌های متعلق به یک خوشه کمتر از فاصلهٔ بین نورن‌های متعلق به خوشه‌های دیگر است. باتوجه به آنچه در شکل ٤ آمده است، فاصلهٔ بین نورن‌های متعلق به خوشهٔ ٤، در ماتریس U الگوریتم SOM با شش‌ضلعی‌های آبی پررنگ و فاصلهٔ این نورن‌ها با نورن‌های خوشه‌های دیگر با شش‌ضلعی‌های روشن‌تر زرد یا قرمز مشخص شده است. همچنین، فاصلهٔ نورن‌های متعلق به خوشهٔ ۱۰، نسبت به یکدیگر با شش‌ضلعی‌های آبی رنگ و با نورن‌های متعلق به خوشه‌های ۲ و ۵ با شش‌ضلعی زرد و نارنجی رنگ مشخص شده است.

تحلیل و بررسی این ماتریس‌ها نشان می‌دهد که نتایج حاصل از دو الگوریتم SOM و Geo–SOM به‌طور کامل،

متفاوت است. در ماتریس U الگوریتم SOM بیشتر نورون‌های شبکه (۲۳ نورون) در خوشهٔ ٤ و تعداد ۷ نورون در خوشهٔ ۲ و سایر نورون‌های شبکه در خوشه‌های دیگر دسته‌بندی شده است. در ماتریس U الگوریتم Geo-SOM به‌جزء خوشه‌های ۱، ۷، ۹ و ۱۰ که به هرکدام یک نورون تعلق دارد، ٤۵ نورون دیگر شبکه در ۷ خوشهٔ دیگر دسته‌بندی شده است. تفاوت اصلی این دو الگوریتم در نحوهٔ نگاشت بردارهای ورودی در نورون‌های شبکه است که این تفاوت نیز منجر به دسته‌بندی متفاوت نورون‌های شبکه در خوشه‌های هر دو الگوریتم می‌شود؛ برای مثال، بردارهای ورودی نگاشت‌شده در قسمت بالا و سمت چپ نورون‌های شبکهٔ SOM باید ویژگی‌های به‌طور تقریبی، مشابهی داشته باشند تا در خوشهٔ ٤ دسته‌بندی شوند؛ درحالی که در همین بخش از نورون‌های شبکهٔ Geo-SOM، بردارهایی نگاشت شده‌اند که ازطرفی، ویژگی‌های متفاوت دارند و از طرف دیگر، در چندین خوشهٔ متفاوت دسته‌بندی شده‌اند. این تفاوت الگوریتم‌ها در نگاشت بردارهای ورودی به نورون‌های شبکه به‌دلیل وزن نسبی متفاوت هریک از متغیرها در بردارهای ورودی الگوریتم‌هاست.



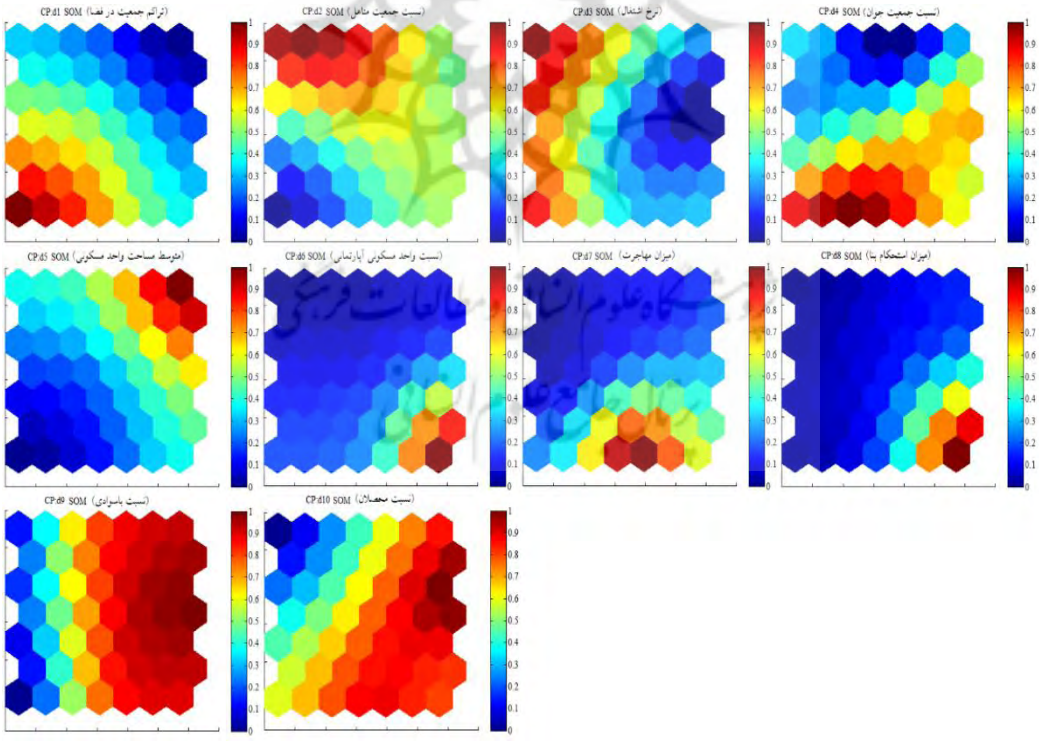
شکل ٤: ماتریس U و شمارهٔ خوشه‌های الگوریتم‌های SOM و Geo-SOM (منبع: نویسندگان ۱٤۰۲)

Figure 4: U-Mat Cluster Number of SOM and Geo-SOM Algorithms

صفحات مؤلفه‌ها

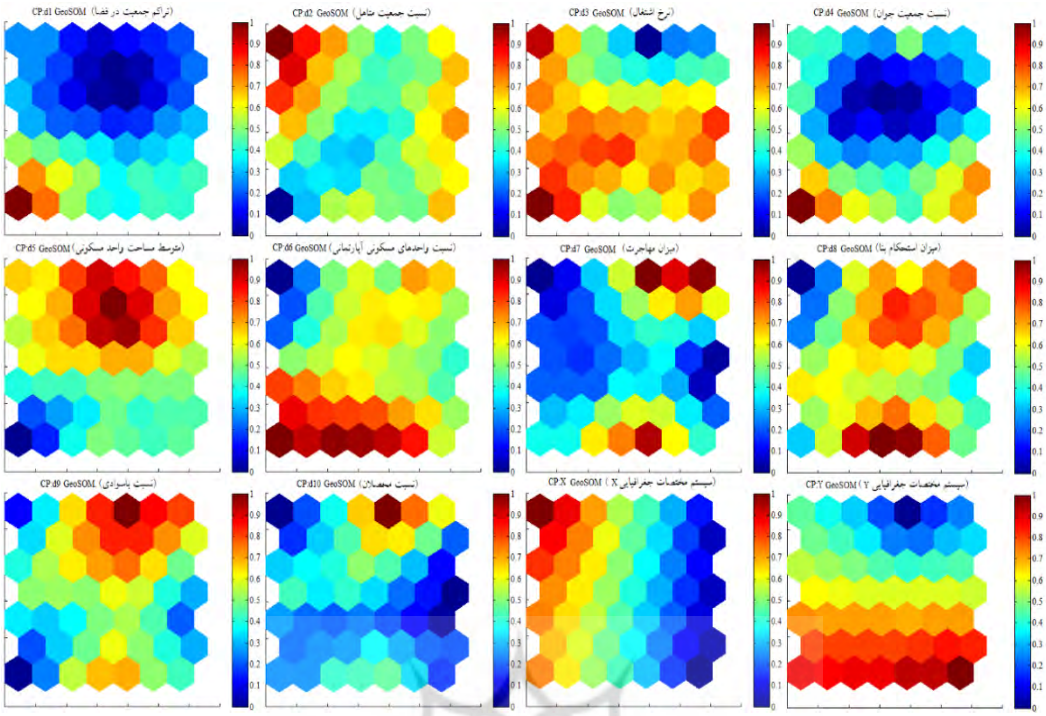
الگوریتم نقشه‌های خودسازمانده، صفحهٔ مؤلفهٔ همهٔ بردارهای ورودی را به‌طور جداگانه، ترسیم می‌کند. این صفحات نشان‌دهندهٔ وزنی است که بردار ورودی را به نورون شبکه نگاشت می‌کند. وزن هریک از بردارهای ورودی بین ۰ تا ۱ محاسبه شده است. با صفحات مؤلفه می‌توان وزن نسبی هریک از مؤلفه‌ها را در نگاشت‌شدن به نورون‌های شبکه برآورد و با وزن نسبی مؤلفه‌های دیگر مقایسه و همبستگی موجود بین بردارهای ورودی و نحوهٔ نگاشت‌شدن آنها را نیز تحلیل و بررسی کرد. صفحات عناصر بردارهای ورودی هر دو الگوریتم نشان می‌دهد که وزن نسبی هریک از متغیرها در نگاشت‌شدن به نورن‌های شبکه به‌طور کامل، متفاوت است؛ برای مثال، متغیر نسبت جمعیت متأهل، وزن نسبی متفاوتی را در نگاشت‌شدن به شبکه در الگوریتم‌های SOM و Geo-SOM دارد. در الگوریتم SOM مقدارهای

بیشتر این متغیر ورودی تمایل به نگاشت‌شدن در نورون‌های بالا سمت چپ شبکه را دارد (نورون‌های قرمز رنگ)؛ زیرا منطبق بر خوشهٔ شمارهٔ ٤ است و مقدارهای کمتر این متغیر تمایل به نگاشت‌شدن در نورون‌های پایین سمت چپ شبکه را دارد؛ زیرا انطباق زیادی با خوشهٔ شمارهٔ ۱ دارد (شکل ۵). در الگوریتم Geo-SOM، مقدارهای بالاتر این متغیر در نورون‌های سمت چپ شبکه و مقدارهای پایین‌تر آن در نورن‌های پایین سمت چپ شبکه نگاشته شده است؛ زیرا مقدارهای بالاتر بر خوشه‌های شمارهٔ ۸ و ۹ و مقدارهای پایین‌تر بر خوشه‌های شمارهٔ ۲ و ۱۰ منطبق است (شکل ٦). همچنین، در الگوریتم SOM مقدارهای زیاد متغیر نسبت باسوادی تمایل به نگاشت‌شدن در نورون‌های سمت راست شبکه را دارد که انتظار می‌رود در خوشه‌های شمارهٔ ٤، ۵ و ۱۰ دسته‌بندی شود. درصورتی که مقدارهای پایین این متغیر تمایل به نگاشت‌شدن در نورون‌های سمت چپ شبکهٔ منطبق بر خوشه‌های شمارهٔ ۱ و ۸ را دارد. در الگوریتم Geo-SOM مقدارهای بالای متغیر میزان باسوادی تمایل به نگاشت‌شدن در نورون‌های بالا سمت راست شبکهٔ منطبق بر خوشه‌های شمارهٔ ۷، ٦ و یا ۵ را دارد. درصورتی که مقدارهای پایین این متغیر تمایل به نگاشت‌شدن در نورون‌های پایین سمت چپ شبکهٔ منطبق بر خوشه‌های شمارهٔ ۱۰ و ۲ را دارد که نتایج حاصل در جدول‌های ۱ و ۲ بیانگر صحت عملکرد این الگوریتم‌ها با این متغیرهاست. همچنین، مقایسهٔ صفحات مؤلفهٔ بردارهای ورودی نشان می‌دهد که نحوهٔ نگاشت‌شدن و وزن نسبی مؤلفه‌ها در هر الگویتم به‌طور کامل، متفاوت است؛ درحالی که صفحهٔ مؤلفهٔ هریک از متغیرهای ورودی شبیه به یکدیگر نیست؛ بنابراین نحوهٔ نگاشت‌شدن این متغیرها به نورون‌های شبکه متفاوت است و این متغیرها با یکدیگر همبستگی ندارند.



شکل ۵: صفحات مؤلفه‌های الگوریتم SOM (منبع: نویسندگان ۱٤۰۲)

Figure 5: Component Planes of SOM Algorithm

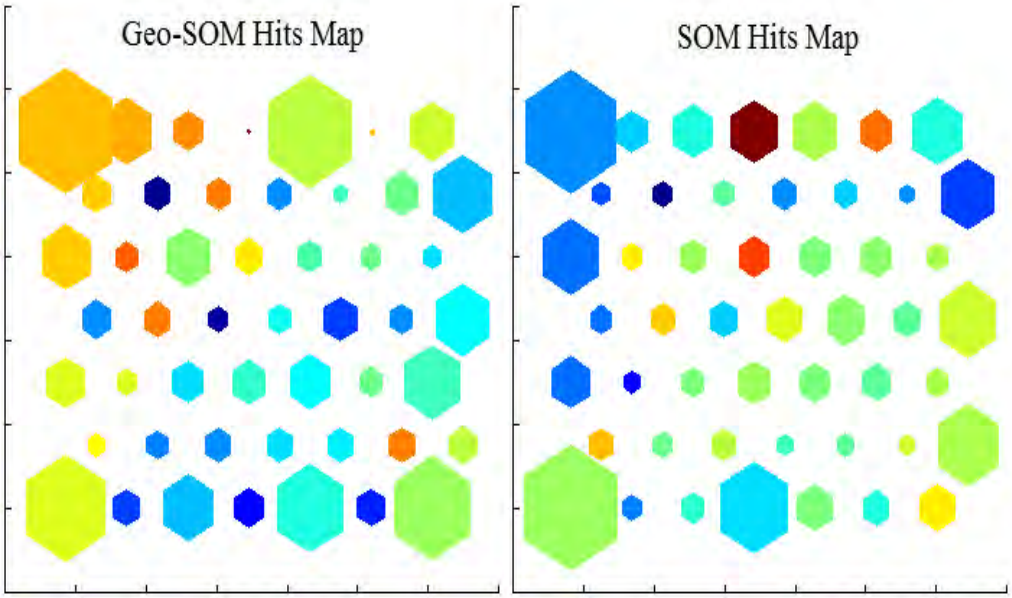


شکل ٦: صفحات مؤلفه‌های الگوریتم Geo-SOM (منبع: نویسندگان ۱٤۰۲)

Figure 6: Component Planes of Geo-SOM Algorithm

Hits Maps

الگوریتم نقشه‌های خودسازمانده، نقشهٔ Hits Maps را پس از اتمام نگاشت بردارهای ورودی ترسیم می‌کند و تعداد نسبی برنده‌شدن نورون‌های شبکه را نشان می‌دهد. هرچه اندازهٔ شش‌ضلعی‌ها بزرگ‌تر باشد، نشان‌دهندهٔ این است که این نورون، دفعات بیشتری برنده شده و تعداد بیشتری بردار ورودی در آن نورون نگاشت شده است و برعکس. تحلیل و بررسی Hits map الگوریتم‌های SOM و Geo-SOM نشان می‌دهد که توزیع بردارهای ورودی در نورون‌ها برای این دو الگوریتم به‌طور کامل، متفاوت است (شکل٥). در الگوریتم SOM دو نورون شبکه بیشتر از بقیه نورون‌ها برنده شده است؛ بنابراین تعداد بردارهای ورودی بیشتری در این دو نورون نگاشت شده است؛ بدین معنا که تنها چند نورون شبکه، بیشترین شانس برنده‌شدن را داشتند و بقیهٔ نورن‌های شبکه شانس کمتری داشتند. به همین خاطر، بردارهای ورودی کمتری نیز در آنها نگاشت شده است. این درحالی است که بردارهای ورودی الگوریتم Geo-SOM به‌طور تقریبی، در نورون‌های بیشتری نگاشت شده است؛ به‌طوری که بیشتر نورن‌های شبکه درمقایسه با الگوریتم SOM شانس بیشتری برای برنده‌شدن دارند و تعداد بیشتری بردار ورودی نیز در آنها نگاشت شده است. تأثیر متغیرهای ورودی در برنده‌شدن نورون‌ها و نگاشت‌شدن بردارهای ورودی در این نورون‌ها نشان می‌دهد که درنظرگرفتن بُعد مکانی، تأثیری مثبت در روند نگاشت بردارهای ورودی در نورون‌های برنده داشته است.

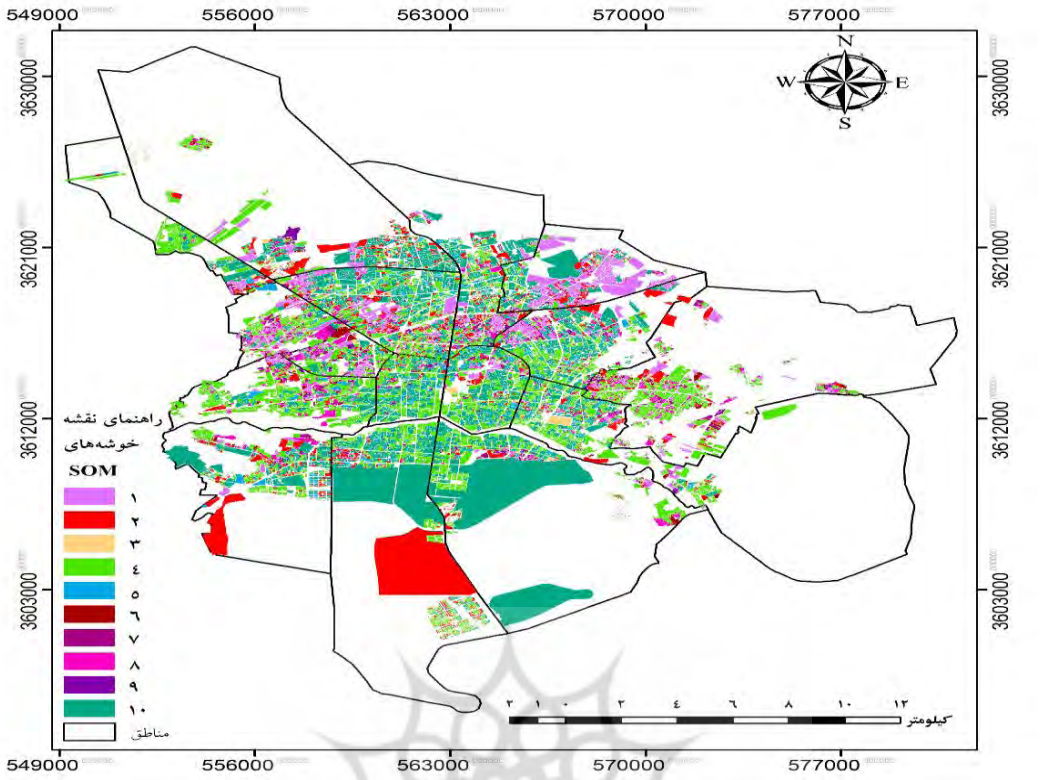


شکل ۷: Hits Map الگوریتم‌های SOM و Geo-SOM (منبع: نویسندگان ۱۴۰۲)

Figure 7: Hits Map of SOM and Geo-SOM Algorithms

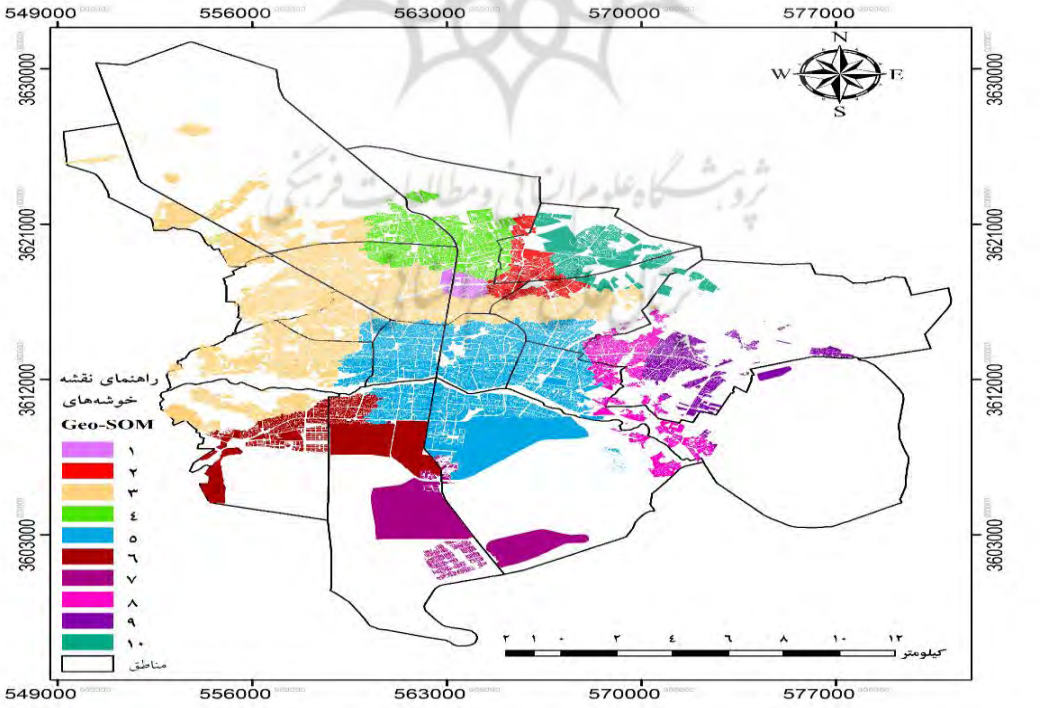
نقشهٔ بلوک‌های آماری خوشه‌بندی‌شده

شکل‌های ۸ و ۹ نشان‌دهندهٔ بلوک‌های آماری خوشه‌بندی‌شدهٔ کلانشهر اصفهان با استفاده از دو الگوریتم SOM و Geo-SOM است. همان‌طور که مشاهده می‌شود نتایج حاصل از خوشه‌بندی هر دو الگوریتم به‌طور کامل، متفاوت است. الگوریتم SOM داده‌های بلوک آماری را در خوشه‌هایی درهم‌تنیده، ناهمگن و پراکنده خوشه‌بندی کرده است (شکل ۸). متغیرهای مطالعه‌شده (داده‌های بلوک آماری) وابسته به مکان هستند؛ بنابراین بلوک‌هایی که ازلحاظ مکانی نزدیک به یکدیگر هستند، تمایل به ویژگی‌های مشابه دارند. پس اگر چنین داده‌هایی خوشه‌بندی شوند و ویژگی‌های مکانی آنها لحاظ نشده باشد، ممکن است به خوشه‌بندی نامناسب منجر شوند که به‌طور کامل، در نتایج خوشه‌بندی این الگوریتم‌ها مشهود است. الگوریتم Geo-SOM با به‌کارگیری ویژگی‌های مکانی و استفاده از معیار شباهت داده‌های بلوک آماری در روند آموزش الگوریتم، داده‌های بلوک آماری را در خوشه‌هایی متجانس و همگن دسته‌بندی کرده است (شکل ۹). همان‌طور که انتظار می‌رفت، هر خوشه در یک محدودهٔ جغرافیایی توزیع شده است. از آنجایی که روند خوشه‌بندی و متغیرهای بردارهای ورودی هر دو الگوریتم یکسان و تنها تفاوت این دو الگوریتم به‌کارگیری مکانی در روند خوشه‌بندی الگوریتم Geo-SOM است، تحلیل ماتریس‌های U، صفحات مؤلفه‌ها و Hits Map هر دو الگوریتم نشان داد که تأثیر پارامترهای مکانی بر نتایج خوشه‌بندی داده‌های بلوک آماری درخور توجه است.



شکل ۸: نقشهٔ بلوک‌های آماری خوشه‌بندی‌شدهٔ شهر اصفهان حاصل از الگوریتم SOM (منبع: نویسندگان ۱٤۰۲)

Figure 8: Map of Clustered Census Blocks of Isfahan City Produced by SOM Algorithm



شکل ۹: نقشهٔ بلوک‌های آماری خوشه‌بندی‌شدهٔ شهر اصفهان حاصل از الگوریتم Geo-SOM (منبع: نویسندگان ۱٤۰۲)

Figure 9: Map of Clustered Census Blocks of Isfahan City Produced by Geo-SOM Algorithm

در الگوریتم‌های SOM و Geo-SOM ۱۳۳۶۲ بلوک آماری در ۱۰ خوشه طبق جدول‌های ۱ و ۲ خوشه‌بندی شده است. در جدول۱ که مربوط به نتایج حاصل از الگوریتم SOM است، خوشهٔ ٤ با ٤٤/٦٦ درصد از کل بلوک‌ها بیشترین تعداد و خوشهٔ ۹ با ۱/۹ درصد کمترین تعداد بلوک را دارد که در ماتریس U الگوریتم SOM نیز مشاهده می‌شود که ۲۳ نورون از ٤۹ نورون در خوشهٔ ٤ و تنها ۱ نورون در خوشهٔ ۹ دسته‌بندی شده است. در نتایج مربوط به الگوریتم Geo-SOM (جدول ۲) خوشهٔ ۳ با ٤٥۰۸ بلوک آماری خوشه‌بندی‌شده و با ۳۳/۷۳ درصد کل بلوک‌ها بیشترین تعداد و خوشهٔ ۱ با ۱/۳٦ درصد کمترین تعداد بلوک را دارد که در ماتریس U الگوریتم Geo-SOM نیز مشهود است که ۱۷ نورون از ٤۹ نورون در خوشهٔ ۳ و تنها ۱ نورون در خوشهٔ ۱ دسته‌بندی شده است. باتوجه به نتایج الگوریتم Geo-SOM که در آن بلوک‌های دسته‌بندی‌شده در یک محدودهٔ جغرافیایی خاص محدود شده‌اند، می‌توان ویژگی‌های هر خوشه را به‌صورت «منطقه‌ای» تحلیل کرد؛ برای مثال، در خوشهٔ ٥ واقع در مرکز و جنوب شهر اصفهان، میانگین مساحت واحد مسکونی برابر با ۱۲۹/۷ متر متربع است که نسبت به میانگین کل شهر که برابر با ۱۱٥/٦ متر مربع است، بالاتر است که این خود نشان‌دهندهٔ بالاتربودن این شاخص ازمنظر توسعهٔ پایدار شهری در این منطقه است. همچنین، خوشهٔ ۱۰ با میانگین ۸۷/۹ متر مربع واقع در شمال شرقی شهر، واحدهای مسکونی کوچک‌تری نسبت به میانگین کل شهر دارد. میانگین میزان باسوادی در کلانشهر اصفهان ۹۱ درصد است؛ درحالی که خوشهٔ ۷ با میانگین ۹۸ درصد واقع در جنوبی‌ترین منطقهٔ شهر، درصد باسوادی بیشتری را نسبت به میانگین کل شهر و خوشهٔ ٦ با میانگین ٥۰ درصد واقع در جنوب شرقی شهر، میزان باسوادی کمتری نسبت به میانگین کل شهر دارد که ازمنظر توسعهٔ پایدار شهری نشان‌دهندهٔ بالاتربودن این شاخص در خوشهٔ ۷ و و پایین‌تربودن آن در خوشهٔ ٦ از میانگین کل شهر است. به‌طور کلی باتوجه به پراکنش خوشه‌ها در یک منطقهٔ خاص و بررسی میانگین متغیرهای هر خوشه می‌توان به نتایج مهمی دست یافت؛ برای مثال، خوشهٔ ۹ واقع در مناطق جنوبی شهر تراکم جمعیت کمتر، نسبت جمعیت متأهل بیشتر، میزان اشتغال بالاتر، نسبت جمعیت جوان متوسط، مساحت مسکونی بالاتر، تعداد مهاجران کمتر و نسبت باسوادی بالاتری از میانگین کل شهر دارد که نشان می‌دهد مناطق جنوبی ازلحاظ شاخص‌های توسعهٔ پایدار و کیفیت زندگی در وضعیت مناسبی قرار دارد. همچنین، خوشهٔ ٤ واقع در مناطق شمالی شهر به‌دلیل داشتن تراکم جمعیت بیشتر، نسبت جمعیت متأهل متوسط، میزان اشتغال متوسط، نسبت جمعیت جوان بالاتر، مساحت مسکونی کمتر، تعداد مهاجران بیشتر و نسبت باسوادی کمتر نسبت به مناطق جنوبی ازلحاظ شاخص‌های توسعهٔ پایدار درسطح پایین‌تری قرار دارد.

از آنجایی که توسعهٔ پایدار شهری روندی در راستای بهبود شرایط فیزیکی، اقتصادی، اجتماعی، فرهنگی است، مدیریت این روند نیازمند شناسایی وضعیت موجود براساس معیارهای مرتبط است. شناسایی مناطق همگن درقالب خوشه‌های مجزا با خصوصیات معین، خود عاملی مؤثر و تسهیل‌کننده برای برنامه‌ریزی‌های محیطیِ متمرکز کوتاه‌مدت و بلندمدت است؛ درنتیجه به‌کارگیری الگوریتم‌های خوشه‌بندی مبتنی بر مکان (Geo)، خوشه‌بندی مناطق شهری را فراهم می‌کند که مبتنی بر شاخص‌های توسعهٔ پایدار و استخراج مناطق همگن برای تسهیل در مدیریت و برنامه‌ریزی شهری است.

جدول ۱: میانگین متغیرهای اجتماعی-اقتصادی در خوشه‌های حاصل از الگوریتم SOM

Table 1: The Average of Socio-Economic Variables in the Clusters Produced by SOM Algorithm

خوشه‌های SOM	تراکم جمعیت در فضا	نسبت جمعیت متأهل	میزان اشتغال	نسبت جمعیت جوان	مساحت واحد مسکونی	تعداد واحد مسکونی آپارتمانی	مهاجران واردشده	میزان استحکام بنا	نسبت باسوادی	نسبت محصلان	تعداد بلوک‌های هر خوشه	درصد بلوک‌های هر خوشه
۱	۰/۰٤	۰/٤۹	۰/۵۰	۰/۲٤	۸۹	۲۹	۰/۰۲	۱۲/۵	۰/۸۱	۰/۲۰	۱۸۱٦	۱۳/۵
۲	۰/۰۳	۰/۵۳	۰/٤۵	۰/۲۵	۹۸/۸	۲۷	۰/۱۰	۲۲/۲۸	۰/۹٤	۰/۲۵	۱۷۹۷	۱۳/٤
۳	۰/۰۳	۰/۵٦	۰/٤۳	۰/۲٤	۱۰۱	۲۳/۹	۰/۰٤	۲۲/۷٤	۰/۹٤	۰/۲۵	۳۵۹	۲/٦۸
٤	۰/۰۲	۰/۵۸	۰/٤٦	۰/۲۱	۱۳۱/۷	۱٦/٦	۰/۰۲	۱۹/۲	۰/۹٦	۰/۲۱	٦۲۳٦	٤٦/٦
۵	۰/۰۲	۰/۵٦	۰/٤٤	۰/۲۲	۱۲۱	۵۲/۵	۰/۰۳	٦۱/۵	۰/۹۷	۰/۲۳	۳۷۹	۲/۹
٦	۰/۰۳	۰/۵۸	۰/٤۷	۰/۲۲	۱۰۲/۵	۱۹/۸	۰/۰۲	۱۸/۱	۰/۹۲	۰/۱۹	۱۹٦	۱,٤٦
۷	۰/۰۳	۰/۵٤	۰/٤۷	۰/۲۰	۹۹/٦	۲۳/۵	۰/۰۱	۱۸	۰/۸۸	۰/۱۹	٤۹۰	۳,٦٦
۸	۰/۰۲	۰/۵٦	۰/۵۰	۰/۲۰	۱۰۵/۸	۱۵/٦	۰/۰۱	۱۲/۵	۰/۸۳	۰/۱٦	۵۸۰	٤/۳٤
۹	۰/۰۳	۰/۵۳	۰/٤٤	۰/۲۳	۹٤	۲۷	۰/۰۱	۱۸/٦	۰/۹۲	۰/۲۲	۲۵٦	۱/۹
۱۰	۰/۰۲	۰/۵٦	۰/٤٦	۰/۲۲	۱۱۷/۹	۱۲٦/٦	۰/۰٤	۱۳۳/٤	۰/۹٤	۰/۲۳	۱۲۵۳	۹/۳۷
میانگین کل	۰/۰۲	۰/۵۵	۰/٤٦	۰/۲۲	۱۱۵/٦	۳۱/۷	۰/۰۳	۳۰/۳	۰/۹۱	۰/۲۲	۱۳۳٦۲	۱۰۰

منبع: نویسندگان ۱٤۰۲

جدول ۲: میانگین متغیرهای اجتماعی-اقتصادی در خوشه‌های حاصل از الگوریتم Geo-SOM

Table 2: The Average of Socio-Economic Variables in the Clusters Produced by Geo-SOM Algorithm

خوشه‌های Geo-SOM	تراکم جمعیت در فضا	نسبت جمعیت متأهل	میزان اشتغال	نسبت جمعیت جوان	مساحت واحد مسکونی	تعداد واحد مسکونی آپارتمانی	مهاجران واردشده	میزان استحکام بنا	نسبت باسوادی	نسبت محصلان	تعداد بلوک‌های هر خوشه	درصد بلوک‌های هر خوشه
۱	۰/۰٤	۰/۵٦	۰/٤۷	۰/۲۷	۱۱۵/۵	٤۲/٤	۰/۰۵	۵٦	۰/۹۳	۰/۲۱	۱۸۳	۱/۳٦
۲	۰/۰۳	۰/۵٤	۰/٤۷	۰/۲۳	۹۳	٤٦/۳	۰/۰۳	۱۹/۳	۰/۸۲	۰/۲۱	۵۳۳	۳/۹۸
۳	۰/۰۳	۰/۵٦	۰/٤۷	۰/۲٦	۱۱۱/۵	۳۰/٤	۰/۰۳	۲٦/۲	۰/۸۹	۰/۲۱	٤۵۰۸	۳۲/۷۳
٤	۰/۰٤	۰/۵٦	۰/٤۵	۰/۲۹	۹٤/۸	۵۰/۷	۰/۰۵	۳۲	۰/۹۰	۰/۲۳	۱۱۱٦	۸/۳۵
۵	۰/۰۲	۰/۵۵	۰/٤٦	۰/۲۰	۱۲۹/۷	۳٤	۰/۰۳	٤۱/۹	۰/۹٤	۰/۲۲	۳۱۵۳	۲۳/۵۹
٦	۰/۰۲	۰/۵۵	۰/٤٤	۰/۲۲	۱۱٦/۷	٤٤/۳	۰/۰٦	٤۳/۵	۰/۹۵	۰/۲۳	۸۰۹	٦/۰۵

۷	۰/۰۲	۰/۵۵	۰/٤۱	۰/۲٤	۱۲٤/۸	۲۰/۲	۰/۰٦	۲۵/٤	۰/۹۸	۰/۲٦	٦۷۹	۵/۰۵
۸	۰/۰۲	۰/۵۷	۰/٤۷	۰/۲۳	۱۱۲/۷	۱۰	۰/۰۲	۱۰/۱	۰/۸۹	۰/۲۱	۹۸٦	۷/۳۷
۹	۰/۰۲	۰/۵۸	۰/٤۹	۰/۲۳	۱۲۱/۲	۵۱/٦	۰/۰۲	٤۲/۳	۰/۹۳	۰/۲۰	۷٦۸	۵,۷٤
۱۰	۰/۰٤	۰/۵۲	۰/٤۹	۰/۰۲٦	۸۷/۹	٤۷/۷	۰/۰۳	۱۳/۱	۰/۸٤	۰/۲۱	٦۳۰	٤,۷۱
میانگین کل	۰/۰۲	۰/۵۵	۰/٤٦	۰/۲۲	۱۱/۵٦	۳۱/۷	۰/۰۳	۳۰/۳	۰/۹۱	۰/۲۲	۱۳۳٦۲	۱۰۰

منبع: نویسندگان ۱٤۰۲

ارزیابی نتایج خوشه‌بندی SOM و Geo-SOM با مقایسهٔ ضرایب سیلهوته بیانگر این است که در نتایج الگوریتم SOM بیشتر بلوک‌های آماری خوشه‌بندی‌شده، ضرایب سیلهوته منفی و تعداد کمی ضریب سیلهوته مثبت دارند. این درحالی است که ضرایب سیلهوته بلوک‌های آماری خوشه‌بندی‌شدهٔ الگوریتم Geo-SOM مثبت بوده است و تعداد کمی ضرایب سیلهوته منفی داشتند. کیفیت کلی نتایج خوشه‌بندی با استفاده از میانگین ضریب سیلهوته برای کل مجموعه داده بدین صورت است که الگوریتم Geo-SOM با میانگین ضریب سیلهوته ۰/۲۷ نسبت به الگوریتم SOM با میانگین ضریب سیلهوته ۰/۰۲- کیفیت خوشه‌بندی بالاتری دارد. این نتایج بیانگر تأثیر مثبت پارامترهای مکانی در روند خوشه‌بندی است.

نتیجه‌گیری

شبکه‌های عصبی خودسازمانده کاربردی‌ترین شبکهٔ استفاده‌شده در مسائل خوشه‌بندی است؛ زیرا پس از آموزش شبکه، می‌توان از ابزارهای بصری‌سازی آن برای تجزیه‌وتحلیل خوشه‌های حاصل استفاده کرد. از نقشه‌های خودسازمانده برای خوشه‌بندی داده‌ها و کاهش ابعاد داده‌ها استفاده می‌شود. در پژوهش حاضر برای خوشه‌بندی داده‌های بلوک آماری از دو الگوریتم شبکه‌های عصبی خودسازماندهٔ SOM و Geo-SOM استفاده شد. هدف اصلی پژوهش حاضر اعمال پارامترهای مکانی در روند خوشه‌بندی شبکه‌های عصبی خودسازمانده بود که مختصات جغرافیایی داده‌های بلوک‌های آماری به‌عنوان پارامتر مکانی در کنار سایر متغیرهای ورودی به شبکه اعمال شد. ضریب سیلهوته و میانگین کلی سیلهوته در ارزیابی نتایج، نشان‌دهندهٔ تأثیر مثبت پارامترهای مکانی در روند خوشه‌بندی الگوریتم Geo-SOM است؛ بنابراین الگوریتم Geo-SOM نسبت به الگوریتم SOM یک خوشه‌بندی مناسب و منسجم از بلوک‌های آماری را باتوجه به شباهت و ویژگی‌های مکانی آنها ایجاد می‌کند. در پژوهش حاضر پتانسیل این الگوریتم در تعریف مناطق همگن و شناسایی بلوک‌های مشابه در یک مجموعه داده واقعی نشان داده شد. الگوریتم Geo-SOM مناطق همگن را براساس متغیرهای منتخب شناسایی کرد که این خود منجر به تسهیل برنامه‌ریزی‌های شهری منطبق با رویکرد توسعهٔ پایدار می‌شود.

منابع

احمدی، علی، حسینی‌فر، سید محسن، و نصیری هندخاله، اسماعیل. (۱۳۹۵). تأثیر مهاجرت بر توسعهٔ شهری با استفاده از مدل (SWOT مورد مطالعه: شهر بابل). *دو فصلنامهٔ پژوهش‌های بوم‌شناسی شهری*، *۷*(۲)، ۵۵-٦٦.

برقی، اسماعیل (۱۳۹۷). سوادآموزی عنصر کلیدی توسعهٔ پایدار. *نشریهٔ راهبرد توسعه*، ۱٤ (٤)،۱۸۷-۲۱۰.

تقوایی، مسعود، و صفرآبادی، اعظم. (۱۳۹۲). توسعهٔ پایدار شهری و برخی عوامل مؤثر بر آن (مورد مطالعه: شهر کرمانشاه). *فصلنامهٔ مطالعات جامعه‌شناختی شهری*، ۳(٦)، ۱-۲٦.

توده فلاح، معصومه، خطیبی، آمنه، صفاکیش، محدثه، و عباسی، محمدباقر. (۱۳۹۷). نگاهی به ساخت سنی و وضع مشارکت در بازار کار مردان و زنان درمعرض ازدواج و طلاق از دریچهٔ سرشماری. *فصلنامهٔ* جمعیت، ۲۰(۵)، ۵۷-۷٦.

سپهوند، رضا، و عارف‌نژاد، محسن. (۱۳۹۲). اولویت‌بندی شاخص‌های توسعهٔ پایدار شهری با رویکرد تجزیه‌وتحلیل سلسله‌مراتبی گروهی (مطالعهٔ موردی: در شهر اصفهان). *مطالعات ساختار و کارکرد شهری*، *۱*(۱)، ٤۳-۵۹.

محمدزاده، رحمت. (۱۳۹٤). بررسی تطبیقی الگوی مجتمع‌های مسکونی ویلائی و آپارتمانی (مطالعهٔ موردی: شهر جدید سهند). *جغرافیا و برنامه‌ریزی*، *۱۹*(۵٤)، ۲۹۷-۳۰۲.

موسوی، میرسعید (۱۳۹۷). بررسی سطح تحقق توسعهٔ پایدار شهر تبریز براساس شاخص ردپای بوم‌شناختی. *جغرافیا و مطالعات محیطی*، *۷*(۲۷)، ٦۱-۷٦.

نصیری دارانی، شهربانو. (۱٤۰۱). *تحلیل حساسیت روش ارزیابی چندمعیارهٔ مکانی به تغییر توابع استانداردسازی و وزن معیارها (مطالعهٔ موردی: سنجش وضعیت پایداری توسعه در شهر اصفهان)*. پایان نامه دکتری، دانشگاه شهید بهشتی تهران.

References

Ahmadi, A., Hosseini Far, S. M., & Nasiri Handkhale, I. (2015). The impact of immigration on urban development using the model (SWOT case study: Babol city). Journal *of Urban Ecology Research*, *7*(2), 55-66. **[**In Persian**]**.

Aldegheishem, A. (2014). Evaluating the urban sustainable development on the basis of AHP: A case study for Riyadh city. *Journal of Sustainable Development*, *7*(2), 113. Doi: 10.5539/jsd.v7n2p113

Andrienko, G., Andrienko, N., Bak, P., Bremm, S., Keim, D., von Landesberger, T., & Schreck, T. (2010). A framework for using self-organising maps to analyze spatio-temporal patterns, exemplified by analysis of mobile phone usage. *Journal of Location Based Services*, *4*(3-4), 200-221. Doi: 10.1080/17489725.2010.532816

Bação, F., Lobo, V., & Painho, M. (2004). Geo-self-organizing map (Geo-SOM) for building and exploring homogeneous regions. In *International Conference on Geographic Information Science* (pp. 22-37). Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-30231-5_2

Barghi, I**.** (2017). Literacy is the key element of sustainable development**.** *Strategic Development Journal*, *14*(4), 187-210. **[**In Persian**]**.

Fayyad, U., Piatetsky-Shapiro, G., & Smyth, P. (1996). From data mining to knowledge discovery in databases. *AI Magazine*, *17*(3), 37-37. https://doi.org/10.1609/aimag.v17i3.1230

Galutira, E. F., Fajardo, A. C., & Medina, R. P. (2019). A novel Kohonen self-organizing maps using exponential decay average rate of change for color clustering. In *Intelligent and Interactive*

Computing (pp. 23-33). Springer, Singapore. https://doi.org/10.1007/978-981-13-6031-2_28
Grubesic, T. H., Wei, R., & Murray, A. T. (2014). Spatial clustering overview and comparison: Accuracy, sensitivity, and computational expense. *Annals of the Association of American Geographers*, *104*(6), 1134-1156. https://doi.org/10.1080/00045608.2014.958389
Hagenauer, J., & Helbich, M. (2016). SPAWNN: A Toolkit for spatial analysis with self-organizing neural networks. *Transactions in GIS*, *20*(5), 755-774. https://doi.org/10.1111/tgis.12180
Hsu, K. C., & Li, S. T. (2010). Clustering spatial–temporal precipitation data using wavelet transform and self-organizing map neural network. *Advances in Water Resources*, *33*(2), 190-200. https://doi.org/10.1016/j.advwatres.2009.11.005
Jain, A. K., Murty, M. N., & Flynn, P. J. (1999). Data clustering: A review. *ACM Computing Surveys (CSUR)*, *31*(3), 264-323. https://doi.org/10.1145/331499.331504
Kassambara, A. (2017). *Practical guide to cluster analysis in R: Unsupervised machine learning* (Vol. 1). Sthda.
Kaufman, L., & Rousseeuw, P. J. (2009). *Finding groups in data: an introduction to cluster analysis*. John Wiley & Sons.
Klösgen, W., & Żytkow, J. M. (1996). Knowledge discovery in databases terminology. In Advances in knowledge discovery and data mining (pp. 573-592).
Kohonen, T. (1995). *Self-Organizing Maps, Springer Series in Information Science* (Vol. 30). Springer.
Lez'er, V., Semerianova, N., Kopytova, A., & Truntsevsky, Y. (2019). Youth entrepreneurship as a basis for sustainable urban development: Social and legal aspect. In *E3S Web of Conferences* (Vol. 110, p. 02093). EDP. https://doi.org/10.1051/e3sconf/201911002093
Li, T., Sun, G., Yang, C., Liang, K., Ma, S., & Huang, L. (2018). Using self-organizing map for coastal water quality classification: Towards a better understanding of patterns and processes. *Science of the Total Environment*, *628*, 1446-1459. https://doi.org/10.1016/j.scitotenv.2018.02.163
Liao, X., Tao, H., Gong, X., & Li, Y. (2019). Exploring the database of a soil environmental survey using a geo-self-organizing map: A pilot study. *Journal of Geographical Sciences*, *29*, 1610-1624. https://doi.org/10.1007/s11442-019-1644-8
Löhr, S. C., Grigorescu, M., Hodgkinson, J. H., Cox, M. E., & Fraser, S. J. (2010). Iron occurrence in soils and sediments of a coastal catchment: A multivariate approach using self organising maps. *Geoderma*, *156*(3-4), 253-266. https://doi.org/10.1016/j.geoderma.2010.02.025
Miller, H. J. (2010). The data avalanche is here. Shouldn't we be digging?. *Journal of Regional Science*, *50*(1), 181-201. https://doi.org/10.1111/j.1467-9787.2009.00641.x
Mohammadzadeh, R. (2014). A comparative study of the pattern of villa and apartment residential complexes (A case study of the new city of Sahand). *Scientific Journal of Geography and Planning*, *19*(54), 279-302. **[**In Persian**]**.
Mousavi, M. (2017). Examining the level of realization of sustainable development of Tabriz city based on ecological footprint index. *Geography and Environmental Studies Quarterly*, *7*(27), 61-76. **[**In Persian**]**.
Nasiri Darani**,** Sh. (2022). *Sensitivity analysis of spatial multi-criteria evaluation method to change the standardization functions and weight of criteria (Case study: assessing the sustainability of development in Isfahan*). PhD Thesis. Shahid Beheshti University. [In Persian**]**.
Openshaw, S. (1999). Geographical data mining: Key design issues. In *Proceedings of GeoComputation* (Vol. 99).
Park, Y. S., Chon, T. S., Bae, M. J., Kim, D. H., & Lek, S. (2018). Multivariate data analysis by means of self-organizing maps. In *Ecological Informatics* (pp. 251-272). Springer, Cham. https://doi.org/10.1007/978-3-319-59928-1_12
Patel, P., & Patel, A. (2021). Use of sustainable green materials in construction of green buildings for sustainable development. *IOP Conference Series: Earth and Environmental Science* (Vol. 785, No. 1, p. 012009). IOP Publishing. Doi: 10.1088/1755-1315/785/1/012009
Rodrigues, M., & Franco, M. (2020). Measuring the urban sustainable development in cities through a composite index: The case of Portugal. *Sustainable Development*, *28*(4), 507-520. https://doi.org/10.1002/sd.2005

Sepahvand, R., & Arifnejad, M. (2012). Prioritization of sustainable urban development indicators with a group hierarchical analysis approach (Case study: in Isfahan city). *Studies of Urban Structure and Function*, *1*(1), 43-59. **[**In Persian**]**.

Sheela, K. G., & Deepa, S. N. (2012). An efficient hybrid neural network model in renewable energy systems. *IEEE International Conference on Advanced Communication Control and Computing Technologies (ICACCCT)* (pp. 359-361). IEEE. Doi: 10.1109/ICACCCT.2012.6320802

Stefanovic, P., & Kurasova, O. (2011). Visual analysis of self-organizing maps. *Nonlinear Analysis: Modeling and Control*, *16*(4), 488-504. DOI: 10.15388/NA.16.4.14091.

Stefanovic, P., & Kurasova, O. (2018). Outlier detection in self-organizing maps and their quality estimation. *Neural Network World*, *28*(2), 106-117. DOI:10.14311/NNW.2018.28.006

Sui, D. Z. (2004). Tobler's first law of geography: A big idea for a small world?. *Annals of the Association of American Geographers*, *94*(2), 269-277. https://doi.org/10.1111/j.1467-8306.2004.09402003.x

Taghavai, M., & Safarabadi, A. (2012). Sustainable urban development and some factors affecting it (Study: Kermanshah city)**.** Quarterly Journal of Urban Sociological Studies, *3*(26), 1-26. [In Persian**]**.

Tobler, W. R. (1970). A computer movie simulating urban growth in the Detroit region. *Economic Geography*, *46*(sup1), 234-240. https://doi.org/10.2307/143141

Tude Fallah, M., Khatibi, A., Safakish, M., & Abbasi, M. B. (2017). Looking at the structure of age and participation in the labor market of men and women subject to marriage and divorce through the lens of the census. *Population Quarterly*, *20*(5), 57-76. **[**In Persian**]**.

Vesanto, J., & Alhoniemi, E. (2000). Clustering of the self-organizing map. *IEEE Transactions on Neural Networks*, *11*(3), 586-600. DOI: 10.1109/72.846731

Wankhede, S. B. (2014). Analytical study of neural network techniques: SOM, MLP, and classifier-a survey. *IOSR J. Comput. Eng. Ver. VII*, *16*(3), 2278-661.

Yuan, M., Buttenfield, B., Gahegan, M., & Miller, H. (2004). Geospatial data mining and knowledge discovery. In A research agenda for geographic information science (pp. 365-388). CRC Press.

Zhang, J., & Fang, H. (2012). Using self-organizing maps to visualize, filter, and cluster multidimensional bio-omics data. *Applications of Self-Organizing Maps*, 181-204.

پژوهشگاه علوم انسانی و مطالعات فرهنگی
پرتال جامع علوم انسانی