



Evaluating the Efficiency of Contextual Neural Gas Networks in Clustering of Isfahan's Census Blocks Based on Sustainable Urban Development Variables

Hadi Tarigholizadeh ¹ , Babak Mirbagheri ² , Ali Akbar Matkan ³

1. Center for Remote Sensing and GIS Research, Faculty of Earth Sciences, Shahid Beheshti University, Tehran, Iran

Email: h_tarigholizadeh@sbu.ac.ir

2. (Corresponding Author) Center for Remote Sensing and GIS Research, Faculty of Earth Sciences, Shahid Beheshti University, Tehran, Iran

Email: b_mirbagheri@sbu.ac.ir

3. Center for Remote Sensing and GIS Research, Faculty of Earth Sciences, Shahid Beheshti University, Tehran, Iran

Email: a-matkan@sbu.ac.ir

Article Info

Article type:
Research Article

Article History:

Received:

25 August 2023

Received in revised form:

26 November 2023

Accepted:

29 December 2023

Available online:

26 January 2024

Keywords:

Contextual Neural Gas,
Clustering,
Spatial data,
Sustainable development,
Census blocks,
Isfahan city.

ABSTRACT

Clustering is a vital technique for revealing structures and discerning groupings within extensive datasets, particularly in spatial data analysis, where the primary objective is to segregate data into clusters with shared characteristics. Artificial neural networks are established tools for clustering large and multidimensional datasets. This research focuses on clustering census block data, encompassing 21 socio-economic variables and access to services relevant to sustainable urban development. The study employs the Neural Gas (NG) network without spatial parameters. Then, it introduces the geographic coordinates of census blocks as spatial parameters, comparing the outcomes of the two approaches (NG & CNG). The NG algorithm, a prevalent choice for clustering high-dimensional data, and its spatially enhanced version, the Contextual Neural Gas (CNG) algorithm, were employed in clustering Isfahan city's census blocks. Results indicated a notable distinction in the clusters derived from the implementation of the NG and CNG algorithms. Clustering with the NG algorithm yielded heterogeneous clusters, whereas the CNG algorithm produced homogeneous clusters benefiting from spatial parameters. Evaluation of clustering quality, performed by calculating the average Silhouette coefficient for census blocks, showed the superior performance of the CNG algorithm, attaining a silhouette coefficient of 0.29 compared to the NG algorithm's -0.02. This research affirmed the positive impact of spatial parameters on creating homogeneous clusters within the urban environment. Leveraging the CNG algorithm and extracting homogenous areas based on sustainable development variables contributed to streamlined urban planning and management. The clustering of census blocks using variables related to sustainable urban development and a location-based approach using the CNG algorithm is one of the innovations of this research.

Cite this article: Tarigholizadeh, H., Mirbagheri, B., & Matkan, A. A. (2023). Evaluating the Efficiency of Contextual Neural Gas Networks in Clustering of Isfahan's Census Blocks Based on Sustainable Urban Development Variables. *Geographical Urban Planning Research Quarterly*, 11 (4), 91-109.

<http://doi.org/10.22059/JURBANGEO.2024.368861.1885>



© The Author (s).

Publisher: University of Tehran Press

Extended Abstract

Introduction

In recent years, there has been a dramatic increase in the volume of available spatial data. Consequently, it is necessary to comprehensively assess spatial data, considering each location's distinctive characteristics, to extract meaningful insights. With the abundance and diversity of urban spatial data, the primary challenge lies in effectively representing the knowledge derived from these data and illuminating the relationships between the data and their respective locations, incorporating various studied variables. Spatial data mining employs artificial neural networks (ANN) to unveil patterns and unknown relationships within data, transforming this information into new and potentially valuable knowledge. Clustering, a pivotal aspect of unsupervised machine learning, is an effective method for extracting knowledge from spatial data, aiming to segregate data into clusters with similar characteristics. It is crucial to note that the clustering algorithm for spatial data diverges fundamentally from that used for non-spatial data. This study focuses on clustering the census blocks of Isfahan city based on sustainable development data, encompassing socioeconomic information and access to services. The process employs the Contextual Neural Gas (CNG) algorithm, and the results are compared with those obtained from implementing the Neural Gas (NG) algorithm. This comparative analysis sheds light on the efficacy of these algorithms in clustering spatial data and extracting meaningful insights related to sustainable development in the urban texture.

Methodology

In this study, data from the Isfahan census blocks (2015), compiled by the Iran Statistics Center, was utilized, alongside information on medical-emergency, cultural-educational, and transportation service points provided by Isfahan Municipality. The research incorporates 13,361 statistical blocks, with 21 socioeconomic variables and indicators related to various urban services associated with sustainable urban development used for

the clustering process. Both the Neural Gas (NG) and Contextual Neural Gas (CNG) algorithms were deployed to cluster socioeconomic data of census blocks, and the outcomes were subjected to a comparative analysis. The Neural Gas network, a competitive neural network employing an unsupervised learning model, specializes in solving clustering problems and topology learning. In the NG algorithm, neurons, lacking neighboring connections, dynamically distribute in the input space during training, mirroring the behavior of physical gas. During training, input vectors are presented, a specific vector is chosen, and neurons move towards it, with the displacement influenced by neuron ranking, distance to the input vector, learning rate, and neighborhood range. Importantly, NG lacks a predefined topology representing relationships between neurons. Topology learning is facilitated through Hebb's competitive learning in the post-processing step. The Contextual Neural Gas Network (CNG), an extension of the NG algorithm, integrates spatial characteristics of input data vectors into the clustering process. While neuron adaptation remains consistent in both NG and CNG, their distinction lies in the definition of rank order. CNG accommodates spatial autocorrelation between observations and neurons by leveraging spatial ordering. However, due to the absence of a topologically ordered network in CNG, a two-step procedure is employed to determine rank ordering, incorporating spatial autocorrelation. The Silhouette coefficient was employed in this research to evaluate clustering results. This coefficient, calculated for each sample, class, and the entire dataset, measures the similarity within clusters and dissimilarity between clusters. The overall quality of clustering was assessed using the average Silhouette coefficient for the entire dataset, providing a comprehensive evaluation of the effectiveness of both NG and CNG algorithms in clustering the Isfahan census blocks.

Results and discussion

The outcomes underscore a fundamental distinction between the two algorithms, primarily rooted in their approach to

mapping input vectors onto network neurons, resulting in disparate classifications within the respective clusters. The NG algorithm employs a distance criterion to map input vectors, yielding intertwined and heterogeneous clusters. The comparison of the clustered census blocks graph network derived from both algorithms reveals obvious differences in results. Notably, the CNG algorithm, with an average silhouette coefficient of 0.29, demonstrates superior clustering performance compared to the NG algorithm, which yields a notably lower average silhouette coefficient of -0.02. This emphasizes the enhanced ability of the CNG algorithm to form cohesive and meaningful clusters based on socioeconomic and service access data related to sustainable development in Isfahan city.

Conclusion

This research applied neural networks to cluster census blocks in Isfahan, focusing on variables related to sustainable urban development. The study aimed to explore the impact of spatial parameters on neural network clustering results, incorporating geographic coordinates of census block centroids alongside non-spatial inputs. A comparative analysis of algorithm outcomes with and without spatial parameters positively influenced clustering, creating more homogeneous clusters. The Silhouette coefficient and the overall average of the Silhouette, employed for result evaluation, served as affirmative indicators of the beneficial role played by spatial parameters

in the clustering process of the Contextual Neural Gas (CNG) algorithm. Consequently, compared with the Neural Gas (NG) algorithm, the CNG algorithm demonstrated its proficiency in generating appropriate and cohesive clusters of census blocks, emphasizing their similarity and spatial characteristics. This research showed the potential of the CNG algorithm in defining homogeneous regions and identifying similar blocks within a tangible dataset. The utility of this algorithm extends to facilitating urban planning by pinpointing homogeneous areas based on selected variables aligned with a sustainable urban development approach. The findings underscore the practical significance of the CNG algorithm as a valuable tool for informed decision-making in urban development and planning initiatives.

Funding

There is no funding support.

Authors' Contribution

Authors contributed equally to the conceptualization and writing of the article. All of the authors approved the content of the manuscript and agreed on all aspects of the work declaration of competing interest none.

Conflict of Interest

Authors declared no conflict of interest.

Acknowledgments

We are grateful to all the scientific consultants of this paper.

ارزیابی کارایی شبکه عصبی گازی بافتی در خوشه‌بندی داده‌های بلوک‌های آماری شهر اصفهان مبتنی بر متغیرهای توسعه پایدار شهری

هادی تاریقلی‌زاده^۱، بابک میرباقری^۲، علی‌اکبر متکان^۳

- ۱- مرکز مطالعات سنجش‌ازدور و GIS، دانشکده علوم زمین، دانشگاه شهید بهشتی، تهران، ایران. رایانامه: h_tarigholizadeh@sbu.ac.ir
- ۲- نویسنده مسئول، مرکز مطالعات سنجش‌ازدور و GIS، دانشکده علوم زمین، دانشگاه شهید بهشتی، تهران، ایران. رایانامه: b_mirbagheri@sbu.ac.ir
- ۳- مرکز مطالعات سنجش‌ازدور و GIS، دانشکده علوم زمین، دانشگاه شهید بهشتی، تهران، ایران. رایانامه: a-matkan@sbu.ac.ir

چکیده

اطلاعات مقاله

خوشه‌بندی داده‌های بزرگ، ساختارها را آشکار و گروه‌بندی‌ها را شناسایی می‌کند و هدف اصلی آن تفکیک داده‌ها در خوشه‌هایی با ویژگی‌های مشابه است. شبکه‌های عصبی مصنوعی ابزاری استاندارد برای خوشه‌بندی داده‌های بزرگ و چندبعدی هستند. هدف این تحقیق، خوشه‌بندی داده‌های بلوک‌های آماری شامل ۲۱ متغیر اجتماعی-اقتصادی و دسترسی به خدمات مرتبط با رویکرد توسعه پایدار شهری با شبکه عصبی گازی بدون استفاده از پارامترهای مکانی و همچنین با بکارگیری مرکز هندسی بلوک‌های آماری بعنوان پارامتر مکانی در روند خوشه‌بندی و مقایسه نتایج حاصل می‌باشد. الگوریتم شبکه عصبی گازی (NG) متداول‌ترین شبکه برای خوشه‌بندی داده‌های با ابعاد بالا و شبکه عصبی گازی بافتی (CNG) مکانی شده این الگوریتم است. در این مطالعه بلوک‌های آماری شهر اصفهان با آموزش این دو الگوریتم بر اساس متغیرهای منتخب خوشه‌بندی شدند. نتایج بیانگر وجود تفاوت قابل توجه در خوشه‌های حاصل از اجرای این دو الگوریتم است. خوشه‌بندی با استفاده از الگوریتم NG، منتج به خوشه‌های ناهمگن می‌شود و بالعکس اجرای الگوریتم CNG به دلیل استفاده از پارامترهای مکانی منجر به تولید خوشه‌های همگن می‌گردد. در این پژوهش ارزیابی کیفیت خوشه‌بندی با محاسبه متوسط ضریب سیلپه‌توته برای بلوک‌های آماری انجام شد که الگوریتم CNG با متوسط ضریب سیلپه‌توته برابر ۰/۲۹ عملکرد بهتری نسبت به الگوریتم NG با متوسط ضریب سیلپه‌توته ۰/۰۲ دارد. این نتایج بیانگر تاثیر مثبت پارامترهای مکانی در ایجاد خوشه‌های همگن در محیط شهری است. خوشه بندی بلوک‌های آماری شهری با بکارگیری متغیرهای مرتبط با توسعه پایدار و رویکرد مکان مینا با استفاده از الگوریتم CNG از جمله نوآوری‌های این تحقیق به شمار می‌رود.

نوع مقاله:

مقاله پژوهشی

تاریخ دریافت:

۱۴۰۲/۰۶/۰۳

تاریخ بازنگری:

۱۴۰۲/۰۹/۰۵

تاریخ پذیرش:

۱۴۰۲/۱۰/۰۸

تاریخ چاپ:

۱۴۰۲/۱۱/۰۶

واژگان کلیدی:

شبکه عصبی گازی بافتی، خوشه‌بندی، داده‌های مکانی، توسعه پایدار، بلوک‌های آماری، شهر اصفهان.

استناد: تاریقلی‌زاده، هادی؛ میرباقری، بابک و متکان، علی‌اکبر. (۱۴۰۲). ارزیابی کارایی شبکه عصبی گازی بافتی در خوشه‌بندی داده‌های بلوک‌های آماری شهر اصفهان مبتنی بر متغیرهای توسعه پایدار شهری. *پژوهش‌های جغرافیای برنامه‌ریزی شهری*، ۱۱ (۴)، ۹۱-۱۰۹.

<http://doi.org/10.22059/JURBANGEO.2024.368861.1885>

مقدمه

در سال‌های اخیر، حجم داده‌های مکانی در دسترس به شدت افزایش یافته است. از این رو نیاز به ارزیابی جامع داده‌های مکانی و استخراج دانش از آن‌ها با در نظر گرفتن خصوصیات منحصر به فرد هر مکان بیش از پیش احساس می‌شود. با توجه به میزان و تنوع داده‌های مکانی شهری، چالش اصلی، چگونگی به تصویر کشیدن دانش به دست آمده از این داده‌ها و روابط کشف شده میان این داده‌ها و مکان با ترکیبی از چندین متغیر مورد مطالعه است. لزوم تبدیل پایگاه داده‌های حجیم حاصل از سرشماری بلوک‌های آماری شهری به دانشی مفید با در نظر گرفتن ویژگی‌های مکانی، محققان را به استفاده از روش‌هایی با توان محاسباتی بالا که قادر به کاهش پیچیدگی‌ها هستند سوق می‌دهد.

خوشه‌بندی یکی از روش‌های متداول برای کار با مجموعه داده‌های بزرگ است که حجم داده‌ها را با ادغام کردن داده‌های مشابه در خوشه‌ها کاهش می‌دهد به طوری که تفاوت داخلی داده‌ها در خوشه‌ها به طور قابل ملاحظه‌ای کمتر از تفاوت‌های بین خوشه‌ها است (Andrienko et al., 2010). خوشه‌بندی یکی از رویکردهای داده‌کاوی است که در راستای تحلیل و تفسیر داده‌ها به کار گرفته می‌شود و با استفاده از شباهت و تفاوت میان داده‌ها به دنبال کشف ساختار مجموعه داده‌های بزرگ است (Hen et al., 2006). خوشه‌بندی روشی مفید در استخراج دانش از داده‌ها و از اساسی‌ترین مسائل یادگیری ماشین بدون نظارت است که هدف اصلی آن تفکیک داده‌ها در خوشه‌هایی با ویژگی‌های مشابه است (Jain et al., 1999). در واقع از طریق خوشه‌بندی داده‌ها سازمان‌دهی شده و عملیات تجزیه و تحلیل و اکتشاف داده‌ها تسهیل می‌گردد.

داده‌های مکانی اغلب دارای خودهمبستگی مکانی^۱ هستند که یک مفهوم اساسی در علوم مکانی است (Sui, 2004) و بدین معنی است که پدیده‌هایی که از نظر مکانی نزدیک به یکدیگر قرار دارند شباهت‌های بیشتری نسبت به پدیده‌هایی دارند که از هم فاصله دارند (Tobler, 1970). بنابراین الگوریتم‌های خوشه‌بندی مکانی بایستی وابستگی مکانی داده‌ها را نیز در خوشه‌بندی بکار گیرند تا از این داده‌ها برای کشف الگوهای مکانی استفاده کنند (Hagenauer, 2015). خوشه‌بندی مکانی سازمان‌دهی اشیاء مکانی در خوشه‌ها است به گونه‌ای که اشیاء داخل یک خوشه نه تنها از نظر ویژگی‌های توصیفی اشیاء بلکه از نظر مکانی نیز مشابه بوده و با اشیاء داخل خوشه‌های دیگر دارای تفاوت باشند (Miller, 2010). الگوریتم خوشه‌بندی داده‌های مکانی، اساساً با داده‌های غیر مکانی متفاوت است (Grubestic et al., 2014). اگر داده‌های مکانی خوشه‌بندی شوند در حالی که وابستگی مکانی آن‌ها نادیده گرفته شود، منجر به خوشه‌بندی نامناسب داده‌ها می‌شود (Openshaw, 1999).

داده‌کاوی مکانی با به کارگیری شبکه‌های عصبی مصنوعی^۲ (ANN) و پایگاه داده مکانی^۳ به کشف الگوها و روابط ناشناخته در داده‌ها می‌پردازد تا اطلاعات را استخراج کرده و در نهایت آن‌ها را به دانش جدید و مفید تبدیل کند (Yuan et al., 2004). مهم‌ترین مزیت ANN در مقایسه با سایر روش‌های داده‌کاوی، ارائه نتایج به صورت بصری علاوه بر ارائه آن‌ها به صورت برآوردهای عددی است (Stefanovic et al., 2011). بصری سازی به محقق اجازه می‌دهد خوشه‌ها و روابط بین داده‌ها را مشاهده کرده و نتایج خوشه‌بندی را تحلیل و تفسیر کنند (Zhang et al., 2012).

یکی از انواع شبکه‌های عصبی مصنوعی، شبکه عصبی گازی^۴ (NG) است که مارتینتز و شولتن^۵ در سال ۱۹۹۱ با الهام

1. Spatial Autocorrelation
2. Artificial Neural Networks
3. Spatial Database
4. Neural Gas
5. Martinetz and Schulten

گرفتن الگوریتم نقشه‌های خود سازمانده^۱ (SOM) پیشنهاد کردند. شاخصه اصلی الگوریتم NG رقابت انعطاف‌پذیر گره‌ها با الگوی یادگیری غیر نظارت‌شده است که در هر تکرار نه تنها گره برنده به‌روز می‌شود، بلکه همسایگان آن نیز تعدیل می‌شوند. این شبکه می‌تواند به‌گونه‌ای تنظیم شود که داده‌های گروهی و کلاس‌های نسبتاً مشابه با یکدیگر را شناسایی کند (Sheela, 2012). کاربرد اصلی NG در حل مسائل خوشه‌بندی و یادگیری توپولوژی است که آن را به جایگزینی مناسب برای SOM در خوشه‌بندی تبدیل می‌کند (Martinetz & Schulden, 1993); با این حال، الگوریتم NG به‌ندرت در علوم مکانی به کار گرفته شده است.

الگوریتم شبکه عصبی گازی بافتی^۲ (CNG) توسعه‌یافته الگوریتم NG و یک الگوریتم خوشه‌بندی مبتنی بر فرآیند یادگیری رقابتی است که جنبه مکانی دارد و هدف آن اعمال ویژگی‌های مکانی بردارهای ورودی در روند خوشه‌بندی است (Hagenauer & Helbich, 2013). این الگوریتم به‌طور صریح خودهمبستگی مکانی را در روند خوشه‌بندی اعمال می‌کند. ایده اصلی آن بر این اساس است که مشاهدات مکانی نزدیک به نوره‌ها با توجه به رتبه ترتیبی آن‌ها نداشت شده و وابستگی مکانی مستقل از مقادیر ویژگی داده‌ها اعمال شود (Hagenauer & Helbich, 2013).

هدف تحقیق حاضر خوشه‌بندی بلوک‌های آماری شهر اصفهان بر اساس داده‌های مرتبط با توسعه پایدار شامل داده‌های اقتصادی- اجتماعی و دسترسی به خدمات با استفاده از الگوریتم CNG و مقایسه آن با نتایج به‌دست‌آمده از اجرای الگوریتم NG است.

مرور ادبیات تحقیق مرتبط با الگوریتم CNG نشان می‌دهد تاکنون در تحقیقات معدودی به موضوع خوشه‌بندی با استفاده از این الگوریتم پرداخته شده است از جمله Hagenauer و Helbich در سال ۲۰۱۳ ضمن معرفی الگوریتم CNG، برای ارزیابی عملکرد آن، این الگوریتم را جهت خوشه‌بندی دو مجموعه داده مصنوعی و داده دنیای واقعی بکار گرفته و نتایج آن را با نتایج الگوریتم Geo-SOM مقایسه نمودند. یافته‌های تحقیق نشان داد در مجموعه داده‌های مصنوعی الگوریتم CNG خوشه‌های همگن‌تر، با دقت موقعیتی بهتر و خطای کمی سازی^۳ کمتری نسبت به Geo-SOM تولید می‌کند. همچنین نتایج حاصل از خوشه‌بندی مجموعه داده‌های دنیای واقعی نشان داد که خوشه‌های حاصل از CNG از نظر تئوریک صحیح‌تر و منسجم‌تر از نتایج حاصل از Geo-SOM هستند. این مطالعه نشان داد که CNG در خوشه‌بندی بر Geo-SOM برتری دارد.

Hagenauer در تحقیقی در سال ۲۰۱۵، روش جدیدی را معرفی کرد که برای یادگیری توپولوژی و خوشه‌بندی داده‌های مکانی، الگوریتم CNG را با خوشه‌بندی گراف ترکیب می‌کند که می‌توان از آن برای خوشه‌بندی داده‌های مکانی استفاده کرد. نتایج به‌دست‌آمده با این روش پیشنهادی با نتایج مطالعات مشابه که توسط شهرداری شهر فیلاولفیا انجام شده بود مقایسه گردید که افزایش کارایی الگوریتم CNG برای یادگیری توپولوژی و خوشه‌بندی داده‌های مکانی در ترکیب با خوشه‌بندی گراف تأیید شد.

Hagenauer و Helbich در سال ۲۰۱۶ در مطالعه‌ای با به‌کارگیری الگوریتم‌های CNG و Geo-SOM به خوشه‌بندی داده‌های اجتماعی- اقتصادی بلوک‌های آماری ۲۰۱۰ ایالات متحده، با متغیرهای قومیت، سن، مسکن و خانوارهای شهر فیلاولفیا پرداختند. در ابتدا هر دو الگوریتم برای تحلیل داده‌های پرت بکار برده شد. نتایج حاصل تفاوت‌های قابل توجهی دارند که نشانگر مناسب بودن CNG نسبت به Geo-SOM در شناسایی داده‌های پرت است. در مرحله دوم پس از حذف

1. Self-Organizing Map
2. Contextual Neural Gas
3. Vector Quantization

داده‌های پرت، تحلیل همبستگی برای شناسایی و ارزیابی ارتباط بین ویژگی‌های مختلف داده‌ها با استفاده از صفحات عناصر^۱ داده‌ها انجام شد که نتایج نشانگر مناسب بودن Geo-SOM نسبت به CNG در تحلیل همبستگی است. در مرحله تحلیل خوشه‌ای، هر دو الگوریتم خوشه‌های مشابهی را شناسایی کردند با این حال تفاوت‌هایی در ویژگی خوشه‌های شناسایی شده وجود داشت. به‌عنوان مثال Geo-SOM در شناسایی خوشه‌ها در ارتباط با صفحات عناصر به‌منظور بررسی روابط داده‌ها مفید است، ولی خوشه‌بندی CNG از نظر جغرافیایی دقیق‌تر است.

مبانی نظری

شبکه عصبی مصنوعی (ANN)

شبکه‌های عصبی مصنوعی روش‌های محاسباتی نوین برای یادگیری و استخراج دانش از پایگاه داده‌های پیچیده هستند. ایده اصلی این شبکه‌ها، پردازش داده‌ها و اطلاعات به‌منظور یادگیری و ایجاد دانش می‌باشد. ANN ارتباط بین چند مجموعه داده را از طریق آموزش فراگرفته و برای استفاده در موارد مشابه ذخیره می‌کند (Nielsen, 2015). نورون‌ها کوچک‌ترین واحد پردازشگر اطلاعات هستند که اساس عملکرد شبکه‌های عصبی را تشکیل می‌دهند. یک شبکه عصبی مجموعه‌ای از نورون‌هاست که با قرار گرفتن در لایه‌های مختلف، معماری خاصی را بر مبنای ارتباطات بین نورون‌ها در لایه‌های مختلف تشکیل می‌دهند. نورون می‌تواند یک تابع ریاضی غیرخطی باشد، در نتیجه یک شبکه عصبی که از اجتماع این نورون‌ها تشکیل می‌شود، نیز می‌تواند یک سامانه کاملاً پیچیده و غیرخطی باشد. در شبکه عصبی هر نورون به‌طور مستقل عمل می‌کند و رفتار کلی شبکه، برآیند رفتار نورون‌های متعدد است (Kasabov, 1996).

شبکه عصبی گازی (NG)

شبکه عصبی گازی یکی از انواع شبکه‌های عصبی رقابتی با الگوی یادگیری غیر نظارت‌شده است که کاربرد اصلی آن در حل مسائل خوشه‌بندی و یادگیری توپولوژی است. الگوریتم NG از نظر طبقه‌بندی در رده الگوریتم‌های کمی ساز برداری قرار می‌گیرد (Martinetz et al., 1993). NG متشکل از تعداد دلخواهی نورون است که با هیچ نورونی رابطه همسایگی ندارند و به هم متصل نیستند. نورون‌ها به‌طور پویا در فضای ورودی در طول فرآیند آموزش همانند گاز فیزیکی توزیع می‌شوند. در این توزیع هیچ‌گونه محدودیت توپولوژیکی وجود ندارد. فضای خروجی NG در همان ابعاد فضای ورودی است و تحت هیچ‌گونه محدودیت توپولوژیکی قرار نمی‌گیرد (Hagenauer, 2015). NG از مجموعه M نورون تشکیل شده است و هر نورون $(k \in M)$ با بردار نمونه اولیه (m_k) در ارتباط است. هدف از آموزش NG این است که بردارهای نمونه اولیه را در فضای ورودی توزیع کند تا تابع چگالی احتمال آن را برآورد کنند (Hagenauer, 2015; Martinetz & Schulten, 1991). تعداد نورون‌ها در اجرای الگوریتم NG حیاتی است. با این حال، هیچ قانون دقیقی برای تعیین تعداد نورون‌ها وجود ندارد. الگوریتم NG مجموعه داده‌ها را در n خوشه که در آن n تعداد نورون‌ها است، خوشه‌بندی می‌کند. اگر تعداد نورون‌ها کم باشد بردار ورودی با ویژگی‌های غیرمشابه نیز در این نورون‌ها نگاشت می‌شود و اگر تعداد نورون‌ها زیاد باشد در اکثر نورون‌های شبکه هیچ بردار ورودی نگاشت نمی‌شود (Hagenauer & Helbich, 2013). در طی مرحله آموزش، بردارهای ورودی به NG ارائه می‌شوند؛ یک بردار ورودی x از مجموعه بردارهای ورودی انتخاب می‌شود و نورون‌ها به سمت بردار ورودی انتخاب‌شده حرکت می‌کنند. میزان جابجایی به ترتیب رتبه‌بندی نورون‌ها با توجه به فاصله

1. Component Plane

2. Prototype Vector

تا بردار ورودی، نرخ یادگیری و محدوده همسایگی بستگی دارد (Labusch et al. 2009).

عملکرد الگوریتم NG را می‌توان بر اساس روابط زیر توصیف کرد (Martinetz et al., 1993):

$$i_k = \left\{ j \in M : \|x - m_j\| < \|x - m_k\| \right\} \quad \text{رابطه ۱}$$

i_k رتبه نورون‌ها را در M تعریف می‌کند که در آن m_j نورون انتخاب‌شده از مجموعه M و m_k بردار نمونه اولیه می‌باشد که این رابطه عملاً تعداد نورون‌هایی که به x نزدیک‌تر از بردار نمونه اولیه هستند را نشان می‌دهد. سپس هر نورون در M به سمت بردار ورودی ارائه‌شده x با توجه به رتبه‌اش (i_k) بر اساس رابطه (۲) حرکت می‌کند:

$$m_k(t+1) = m_k(t) + \varepsilon(t) e^{-i_k(M)/\lambda(t)} (x - m_k(t)) \quad \text{رابطه ۲}$$

که در آن t مرحله تطبیق، $\varepsilon(t)$ میزان تطبیق‌پذیری $\lambda(t)$ محدوده همسایگی نورون‌هایی است که تطبیق داده می‌شوند. هر دو پارامتر معمولاً به صورت تدریجی در طی آموزش کاهش می‌یابند (Hagenauer, 2015). پس از تعداد کافی تکرارها در مرحله آموزش، بردارهای نمونه اولیه، تابع چگالی احتمال فضای ورودی را با خطای کمی سازی نزدیک به حداقل برآورد می‌کنند (Hagenauer, 2015). الگوریتم NG توپولوژی از پیش تعریف‌شده‌ای ندارد که نشان‌دهنده روابط و شباهت بین نورون‌ها باشد (Martinetz & Schulden, 1991). باین‌حال، توپولوژی بسیار مفید است، زیرا می‌تواند اطلاعات ارزشمندی را در مورد داده‌های مورد مطالعه آشکار کند. برای ایجاد توپولوژی، یادگیری رقابتی هبین را می‌توان در مرحله پس پردازش بدین ترتیب در الگوریتم NG اعمال کرد: برای هر بردار ورودی، دو نورون از نزدیک‌ترین نورون‌ها شناسایی می‌شوند و یک ارتباط بین این دو نورون به کل مجموعه ارتباطات اضافه می‌شود (Martinetz, 1993). نزدیکی معمولاً با فاصله اقلیدسی اندازه‌گیری می‌شود و وقتی همه بردارهای ورودی پردازش شدند، مجموعه ارتباطات حاصل، توپولوژی آموخته‌شده را نشان می‌دهد (Hagenauer, 2015).

قاعده یادگیری در الگوریتم NG بر اساس رابطه (۳) تعریف می‌شود (Martinetz & Schulden, 1991):

$$w_i^{new} = w_i^{old} + a_i (x - w_i^{old}) \quad \text{رابطه ۳}$$

$$a_i = \varepsilon e^{-k_i/\lambda} \quad \text{رابطه ۴}$$

در این روابط w_i نورون‌های ایجادشده در فضای داده می‌باشد که تعداد این نورن‌ها در ابتدا یک مقدار ثابت است و در انتها به منظور کارکرد منطقی و بهینه الگوریتم اصلاح می‌شوند. a_i پارامتری است که نرخ یادگیری را مشخص می‌کند و وابسته به k_i و λ است؛ بدین صورت که اگر λ به سمت بی‌نهایت میل کند یادگیری همه نورون‌ها با هم برابر و اگر به سمت صفر میل کند نزدیک‌ترین نورون شروع به یادگیری می‌کند. حالت‌های حدی λ هیچ کدام به‌تنهایی مناسب نمی‌باشد و معمولاً حالتی بین آن‌ها انتخاب می‌شود. k_i نشان‌دهنده تعداد نورون‌های نزدیک‌تر نسبت به نورون i ام است. ε نیز پارامتری است که نرخ یادگیری را کنترل می‌کند. نکته مهم این است که در طی دوره آموزش بایستی از نرخ یادگیری کاسته شود در غیر این صورت شبکه دچار تکرار شده و چرخه نادرستی ایجاد می‌شود.

روند الگوریتم NG را می‌توان در مراحل زیر خلاصه کرد (Martinetz & Schulden, 1991):

مرحله صفر - موقعیت تصادفی w_i در فضای داده ایجاد می‌شود.

مرحله یک - بردار ورودی x از مجموعه بردارهای ورودی انتخاب می‌شود.

مرحله دو - رتبه‌بندی؛ شامل محاسبه فاصله بین بردار ورودی x و w_i و تخصیص رتبه k_i برای هر نورون بر اساس همین

1. Adaptation

2. Competitive Hebbian learning

فاصله (k_i) برابر تعداد نورون نزدیک‌تر از نورون w_i نسبت بردار ورودی x است).

مرحله سه- مرحله تطبیق یا یادگیری: در این مرحله بر اساس رابطه (۵) برای هر نورون میزان انطباق نسبت به بردار ورودی x محاسبه می‌شود.

$$w_i^{new} = w_i^{old} + \varepsilon e^{-k_i/\lambda} (x - w_i^{old}) \quad \text{رابطه (۵)}$$

مرحله چهار- ایجاد همسایگی: برای ایجاد همسایگی، یالی بین اولین نورون (i) و دومین نورون (j) از نظر نزدیکی ایجاد می‌شود و سن این یال برابر صفر در نظر گرفته می‌شود. برای هر نورون $c_{i,j} \in \{0,1\}$ نشانگر وجود یا عدم وجود یال یا همسایگی $t_{i,j} \in \{0,1,2,3, \dots\}$ نشانگر فاصله زمانی (سن) از آخرین دیدار یا تجدید یال بین نورون i و j است که اگر از یک اندازه بیشتر شود همسایگی شکسته می‌شود. این رویکرد کمک می‌کند تا توپولوژی به شبکه عصبی آموزش داده شود.

مرحله پنج- سن همه یال‌ها افزایش می‌یابد ($t_{i,j} \rightarrow t_{i,j} + 1$).

مرحله شش- در این مرحله $k_i = 0$ فرض می‌شود و به ازای هر j که در آن $t_{i,j} > T$ باشد؛ $c_{i,j} = 0$ در نظر گرفته می‌شود. در این مرحله نیز بایستی T که مشخص‌کننده واحد زمانی تجدید یال است در طی دوره آموزش افزایش یابد تا از میزان سخت‌گیری کاسته شده و اجازه داده شود یال‌ها بیشتر دوام داشته باشند.

مرحله هفت- در صورت برآورد نشدن شرایط خاتمه، از مرحله یک تکرار می‌شود و در غیر این صورت الگوریتم پایان می‌یابد.

الگوریتم شبکه عصبی گازی بافتی (CNG)

همان‌طور که پیش‌تر اشاره شد تفاوت اصلی الگوریتم CNG در مقایسه با الگوریتم NG اعمال ویژگی‌های مکانی بردارهای ورودی در روند خوشه‌بندی است. با اینکه که مرحله انطباق نورون‌ها در الگوریتم‌های NG و CNG یکسان است، ولی این دو در تعریف رتبه k_i برای نورون‌ها متفاوت هستند. الگوریتم CNG برای تعیین رتبه با استفاده از ترتیب مکانی نورون‌ها، خودهمبستگی را در یک روش دو مرحله‌ای بین بردارهای ورودی و نورون‌ها لحاظ می‌کند (Hagenauer & Helbich, 2016). در مرحله اول، نورون‌ها بر اساس نزدیکی مکانی مرتب می‌شوند و نزدیک‌ترین نورون شناسایی می‌شود. در مرحله دوم، نورون نهایی (بهترین نورون) از بین نزدیک‌ترین نورون‌ها در شبکه با توجه به شباهت با بردار ورودی بر اساس سایر ویژگی‌های توصیفی تعیین می‌شود. در الگوریتم CNG رتبه مکانی نورون‌ها در مجموعه M طبق رابطه (۶) تعریف می‌شود:

$$i_k^{geo} = \left| \left\{ j \in M : \|x_{geo} - m_{j,geo}\| < \|x_{geo} - m_{k,geo}\| \right\} \right|. \quad \text{رابطه (۶)}$$

در این رابطه i_k^{geo} رتبه مکانی نورون k ، x_{geo} ویژگی مکانی (مختصات) بردار ورودی x ، $m_{j,geo}$ بردار ویژگی مکانی نورون j از مجموعه M و $m_{k,geo}$ بردار ویژگی مکانی نورون اولیه k می‌باشد. با استفاده از رتبه مکانی به دست آمده و بر اساس رابطه (۷) یک زیرمجموعه S_l تعیین می‌شود که شامل l همسایه نزدیک‌تر نورون k از نظر مکانی در مجموعه M است که این رابطه عملاً تعداد نورون‌هایی که به x نزدیک‌تر از بردار نمونه اولیه هستند را نشان می‌دهد.

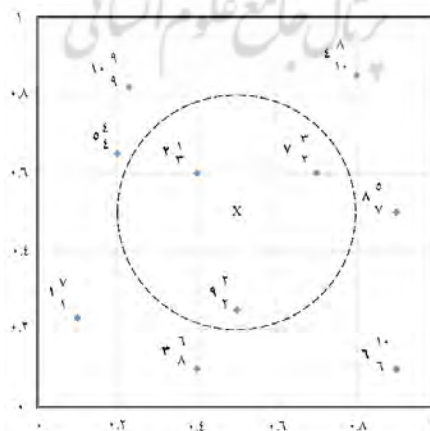
$$S_l := \{k \in M : i_k^{geo} < l\}. \quad \text{رابطه (۷)}$$

در این رابطه پارامتر l اندازه همسایگی فضایی^۱ (SNS) است. SNS خودهمبستگی مکانی لحاظ شده در فرآیند نگاهت CNG را تعیین می‌کند. اگر $l = 1$ باشد، مجموعه S_l فقط شامل نزدیک‌ترین نورون به لحاظ مکانی می‌باشد. در نتیجه، رتبه نهایی برابر رتبه مکانی ($i_k = i_k^{geo}$) است و انطباق نورون‌ها منحصراً توسط فواصل مکانی نورون‌ها تا بردار ورودی تعیین می‌شود و مقادیر ویژگی داده‌های ورودی نادیده گرفته می‌شوند. اگر l افزایش یابد اندازه S_l نیز افزایش می‌یابد و نورون‌های S_l بر اساس شباهت ویژگی‌هایشان به x_{att} (ویژگی‌های توصیفی (غیر مکانی) بردار ورودی x) دوباره رتبه‌بندی می‌شوند. رتبه نهایی نورون k برای انطباق نورون‌ها طبق رابطه (۸) تعریف می‌شود:

$$i_k = \begin{cases} \left\{ j \in S_l : \|x_{att} - m_{j,att}\| < \|x_{att} - m_{k,att}\| \right\} & \text{if } k \in S_l \\ i_k^{geo} & \text{otherwise} \end{cases} \quad \text{رابطه (۸)}$$

با افزایش l تأثیر فواصل مکانی برای رتبه‌بندی نهایی نورون‌ها کمتر شده و انطباق نورون‌ها نیز با تأثیر کمتر مجاورت مکانی تعیین می‌شود و در نتیجه خودهمبستگی مکانی تأثیر کمتری در انطباق نورون‌ها دارد. اگر $l = M$ باشد مجموعه S_l شامل همه نورون‌های M شده و رتبه مکانی i_k^{geo} منسوخ می‌شود؛ زیرا رتبه نهایی نورون‌ها صرفاً بر اساس شباهت ویژگی‌های غیر مکانی آن‌ها تعیین می‌شود و بنابراین هیچ‌گونه خودهمبستگی مکانی در CNG لحاظ نمی‌شود (Hagenauer & Helbich, 2016).

شکل (۱) به صورت نمونه نحوه تعیین رتبه نورون‌ها در الگوریتم CNG را نشان می‌دهد که در آن مجموعه M شامل ۱۰ نورون و $l = 3$ در نظر گرفته شده است. نقاط، نشانگر مرکز نورون‌ها، اعداد نشانگر شماره نورون‌ها، اعداد بالانویس بیانگر رتبه نورون‌ها با توجه به فاصله آن‌ها تا مختصات جغرافیایی بردار ورودی (x_{geo}) هستند. اعداد زیرنویس نیز نشانگر رتبه‌های نورون‌ها با توجه به شباهت به سایر ویژگی‌های بردار ورودی (x_{att}) هستند. دایره نشان‌دهنده نورون‌های زیرمجموعه S_3 است. با توجه به روند الگوریتم CNG در تعریف رتبه مکانی نورون‌ها، ابتدا رتبه نورون‌ها در مجموعه M طبق رابطه (۶) بر اساس ویژگی‌های مکانی تعریف می‌شود و سپس بر اساس پارامتر اندازه همسایگی فضایی، ویژگی‌های نورون‌ها در رتبه‌بندی اعمال می‌شود و رتبه مکانی نهایی طبق رابطه (۸) تعریف می‌شود. بدین ترتیب نورون ۷ با داشتن x_{att} و x_{geo} مشابه بردار ورودی x رتبه مکانی یکم را می‌گیرد و سایر نورون‌ها نیز همانند نورون ۷ با اعمال x_{geo} و x_{att} آن‌ها رتبه‌بندی مکانی شدند که نورون ۲ رتبه دوم، نورون ۹ رتبه سوم، نورون ۵ رتبه چهارم، نورون ۸ رتبه پنجم، نورون ۳ رتبه ششم، نورون ۱ رتبه هفتم، نورون ۴ رتبه هشتم، نورون ۱۰ رتبه نهم و نورون ۶ رتبه دهم را گرفتند.



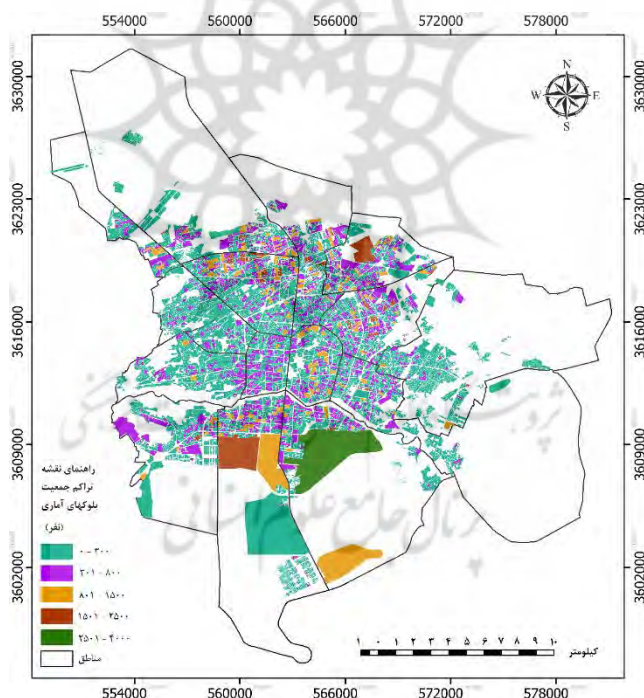
شکل ۱. نمونه رتبه‌بندی نورون‌ها در الگوریتم CNG

گراف شبکه عصبی گازی

گراف یک نمایش انتزاعی از یک شبکه است که در علم داده، برای مدل‌سازی روابط بین اشیاء استفاده می‌شوند. یک گراف را می‌توان به صورت $G=(V,E)$ تعریف کرد که در آن V مجموعه‌ای از گره‌ها (نورون‌ها) و E یال‌های بین آن‌ها است. مهم‌ترین کاربرد گراف به‌عنوان ساختار ریاضی مدل‌سازی پدیده‌های گوناگون و بررسی روابط آن‌هاست (علیپور، ۱۳۸۳). یکی از تحلیل‌های مهمی که روی گراف‌ها انجام می‌شود خوشه‌بندی گره‌های گراف است. خوشه‌بندی گره‌ها در گراف در حقیقت همان مسئله تشخیص انجمن‌های گراف^۱ است مشروط بر این که گره‌های گراف متناظر با نقاط داده و کوتاه‌ترین فاصله میان نودها متناظر با فاصله میان نقاط در نظر گرفته شود. انجمن‌ها گروه‌هایی از نورون‌های شبکه هستند که ارتباط تنگاتنگی با هم دارند و با نورون‌های دیگر گروه‌ها ارتباط نسبتاً کمی دارند (Marino, 2018).

محدوده مورد مطالعه

کلان‌شهر اصفهان سومین شهر پر جمعیت ایران دارای طول جغرافیایی ۵۴ درجه و ۳۹ دقیقه و ۴۰ ثانیه شرقی و عرض جغرافیایی ۳۵ درجه و ۳۸ دقیقه و ۳۰ ثانیه شمالی در بخش مرکزی ایران واقع شده است. این کلان‌شهر با ۵۵۱ کیلومتر مربع مساحت دارای ۱۵ منطقه شهری می‌باشد که پس از تهران دومین کلان‌شهر ایران از نظر مساحت می‌باشد. شکل (۲) مناطق ۱۵ گانه شهر اصفهان را به همراه بلوک‌های آماری آن‌ها نمایش می‌دهد.



شکل ۲. نقشه تراکم جمعیت بلوک‌های آماری شهر اصفهان به تفکیک مناطق

روش پژوهش

داده‌های مورد استفاده در این پژوهش، شامل داده‌های بلوک‌های آماری سرشماری جمعیت سال ۱۳۹۵ کلان‌شهر اصفهان تهیه شده توسط مرکز آمار ایران و همچنین نقاط خدمات درمانی - اضطراری، فرهنگی - آموزشی و حمل‌ونقل تهیه شده

توسط شهرداری اصفهان می‌باشد. تعداد کل بلوک‌های آماری مورد استفاده در این تحقیق برابر ۱۳۳۶۱ عدد می‌باشد. همان‌طور که اشاره شد هدف تحقیق حاضر خوشه‌بندی بلوک‌های جمعیتی بر اساس متغیرهای مرتبط با توسعه پایدار شهری می‌باشد. مفهوم توسعه پایدار توسط کمیسیون جهانی محیط‌زیست و توسعه^۱ (WCED) به‌عنوان توسعه‌ای که نیازهای نسل حاضر را بدون به خطر انداختن توانایی نسل‌های آینده برای برآوردن نیازهای خود مرتفع می‌سازد، تعریف شده است (Aldegheshi, 2014). پایداری شهری شامل ابعاد فیزیکی، اقتصادی و فرهنگی است (Rodrigues & Franco, 2020). هدف اصلی توسعه پایدار، تأمین نیازهای اساسی، بهبود و ارتقای سطح زندگی برای همه، حفظ و اداره بهتر اکوسیستم‌ها و آینده‌ای امن‌تر و سعادتمند است. توسعه پایدار شهری تنها مبتنی بر یک بعد نیست و بدون حل مسائل در تمامی ابعاد رسیدن به توسعه پایدار محقق نخواهد شد (برزگر و همکاران، ۱۳۹۷). لذا بر اساس مرور تحقیقات صورت گرفته در خصوص متغیرهای مرتبط با توسعه پایدار شهری، متغیرهای ذکر شده در جدول (۱) شامل ۲۱ متغیر اقتصادی- اجتماعی و دسترسی به انواع خدمات شهری جهت خوشه‌بندی بلوک‌های آماری مورد استفاده قرار گرفتند. در این جدول نام متغیرها و پژوهش‌هایی که این متغیرها در آن‌ها معرفی شده‌اند ذکر شده است:

جدول ۲. متغیرهای مورد مطالعه

ردیف	متغیر مورد مطالعه	منبع
۱	تراکم جمعیت در فضا	(سپهوند و عارف نژاد، ۱۳۹۲)
۲	نسبت جمعیت متأهل	(موسوی، ۱۳۹۷)
۳	میزان مهاجرین وارد شده	(نصیری هندخاله و همکاران، ۱۳۹۵)
۴	نرخ اشتغال	(عباسی و همکاران، ۱۳۹۷)
۵	نسبت جمعیت جوان	(لزتر ^۲ و همکاران، ۲۰۱۹)
۶	نسبت باسوادی	(برقی، ۱۳۹۷)
۷	نسبت محصلان	(نصیری دارانی، ۱۴۰۱، ۳۸)
۸	متوسط مساحت واحد مسکونی	(نصیری دارانی، ۱۴۰۱، ۳۹)
۹	تعداد واحد مسکونی آپارتمانی	(محمدزاده، ۱۳۹۴)
۱۰	میزان استحکام بنا	(پاتل و پاتل ^۳ ، ۲۰۲۱)
۱۱	دسترسی به داروخانه	(نصیری دارانی، ۱۴۰۱، ۳۸)
۱۲	دسترسی به درمانگاه	(موسوی، ۱۳۹۷)
۱۳	دسترسی به بیمارستان	(نصیری دارانی، ۱۴۰۱، ۳۹)
۱۴	دسترسی به مراکز آتش‌نشانی	(وادهوا ^۴ ، ۲۰۰۰)
۱۵	دسترسی به مراکز پلیس	(عباسی و همکاران، ۱۳۹۷)
۱۶	دسترسی به ایستگاه اتوبوس	(نصیری دارانی، ۱۴۰۱، ۳۹)
۱۷	دسترسی به ایستگاه مترو	(وادهوا، ۲۰۰۰)
۱۸	دسترسی به معابر عمومی	(وادهوا، ۲۰۰۰)
۱۹	دسترسی به مراکز ورزشی	(امیری و همکاران، ۱۳۹۲)
۲۰	دسترسی به مدارس	(برقی، ۱۳۹۷)
۲۱	دسترسی به کتابخانه	(موسوی، ۱۳۹۷)

در ادامه پس از محاسبه مقادیر هر یک از متغیرهای فوق برای بلوک‌های آماری (برای لایه‌های دسترسی متوسط مقادیر

1. World Commission on Environment and Development
2. Lez'er
3. Patel & Patel
4. Wadhwa

دسترسی به هر بلوک تخصیص داده می‌شود، خوشه‌بندی بلوک‌ها بر اساس روش‌های الگوریتم‌های NG و CNG انجام و نتایج آن‌ها با یکدیگر مقایسه می‌گردد.

روش ارزیابی خوشه‌بندی

ارزیابی و مقایسه نتایج خوشه‌بندی در این تحقیق با استفاده از ضریب سیلهوته^۱ انجام شد. این ضریب برای هر داده بر اساس رابطه (۹) محاسبه شده و درجه شباهت میان داده‌های داخل یک خوشه و عدم شباهت با خوشه‌های دیگر را اندازه‌گیری می‌کند (Kaufman & Rousseeuw, 2009).

$$S(i) = \frac{(b(i) - a(i))}{\max[a(i), b(i)]} \quad \text{رابطه ۹}$$

در رابطه فوق مقدار $S(i)$ ضریب سیلهوته داده i ام است. $a(i)$ میانگین عدم تشابه^۲ میان داده i ام با سایر داده‌ها در یک خوشه مشابه و $b(i)$ کمینه میانگین عدم تشابه داده i ام نسبت به سایر داده‌ها در خوشه‌های دیگر می‌باشد. مقدار $S(i)$ بین -1 و $+1$ قرار دارد. اگر $S(i)$ به $+1$ نزدیک‌تر باشد، به معنی مناسب بودن خوشه پیشنهادشده برای داده موردنظر و مقادیر نزدیک‌تر به -1 نشان‌دهنده نامناسب بودن خوشه پیشنهادی است.

در نهایت کیفیت کلی خوشه‌بندی با استفاده از محاسبه میانگین ضریب سیلهوته برای کل مجموعه داده بر اساس رابطه (۱۰) ارزیابی می‌شود:

$$SC = \frac{1}{N} \sum_{i=1}^N s(i) \quad \text{رابطه ۱۰}$$

که در آن N تعداد کل داده‌ها است. مقادیر بالاتر SC نشان‌دهنده مناسب‌تر بودن کیفیت خوشه‌بندی است (Hsu & Li, 2010).

یافته‌ها

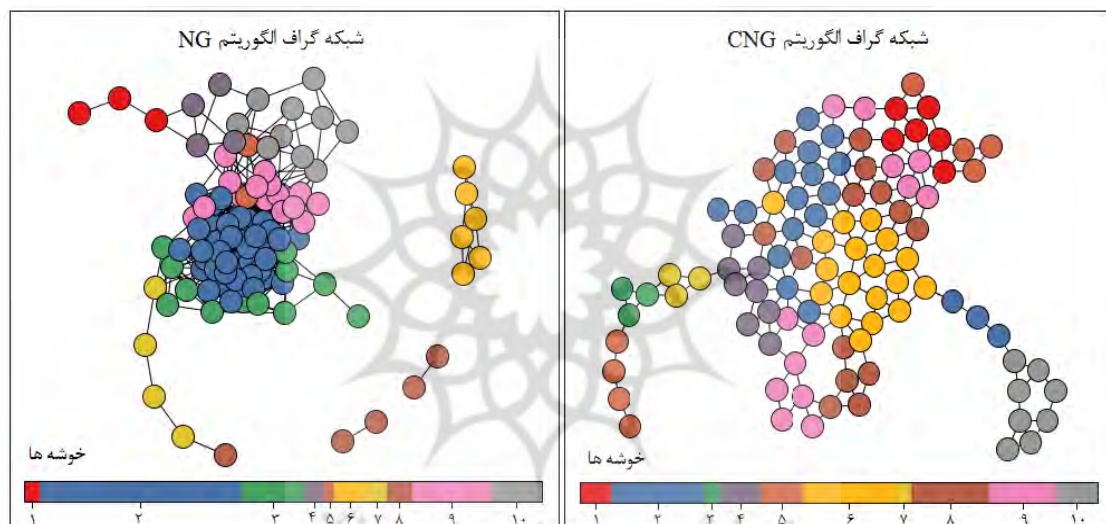
گراف حاصل از شبکه عصبی گازی

گراف حاصل از خوشه‌بندی بلوک‌های آماری کلان‌شهر اصفهان با استفاده از دو الگوریتم NG و CNG در شکل ۳ نشان داده شده است. در این مطالعه پس از بررسی تعداد نورون‌های مختلف در خوشه‌بندی بلوک‌های آماری، تعداد ۱۰۰ نورون و مقدار $l = 3$ برای الگوریتم NG و CNG انتخاب شد. در گراف‌های حاصل از این خوشه‌بندی، گره‌ها نشانگر مراکز نورون‌های شبکه و خطوط نشانگر ارتباط بین نورون‌ها هستند. همچنین رنگ‌های مختلف نشان می‌دهند که نورون‌های شبکه متعلق به کدام خوشه می‌باشند. بررسی گراف حاصل از دو الگوریتم NG و CNG نشان می‌دهد که نتایج آن‌ها کاملاً متفاوت هستند. در الگوریتم NG بر اساس نتیجه به‌دست‌آمده از ۱۰۰ نورون، ۴۰ نورون در خوشه ۲ و ۱۵ نورون در خوشه ۹ دسته‌بندی شده‌اند که به ترتیب دارای بیشترین نورون هستند و همچنین ۳ نورون در خوشه ۱ و ۴ نورون در خوشه ۵ دسته‌بندی شده هستند که به ترتیب دارای کمترین نورون هستند. در الگوریتم CNG نه تنها از متغیرهای مرتبط با توسعه پایدار به‌عنوان معیار اندازه‌گیری شباهت و عدم شباهت استفاده شده، بلکه موقعیت مکانی مراکز هندسی بلوک‌های آماری نیز در روند خوشه‌بندی در نظر گرفته شده و سپس نورون‌های شبکه دسته‌بندی شده‌اند. بنابراین در این خوشه‌ها

1. Silhouette
2. Dissimilarity

نورون‌ها نه تنها از لحاظ ویژگی‌های توصیفی بلکه از لحاظ موقعیت مکانی نیز نزدیک به یکدیگر هستند. بررسی نتایج نشان داد که استفاده از موقعیت مکانی متغیرها باعث شده است شبکه گراف الگوریتم CNG شکل منظم‌تری نسبت به شبکه گراف الگوریتم NG داشته باشد.

بررسی نتایج حاصل از خوشه‌بندی داده‌های مورد مطالعه نشان داد که نورون‌های منفردی که به نورون‌های شبکه گراف متصل نیستند نورون‌هایی هستند که دارای هیچ بردار ورودی نیستند و یا بردار ورودی کمتری در آن‌ها نگاشت شده و اطلاعات توپولوژیکی ندارند. برای نمونه در خوشه 8 در الگوریتم NG که از چهار گره جدا از شبکه اصلی تشکیل شده است تنها چهار بردار ورودی نگاشت شده و گره‌ها رابطه توپولوژیکی با هم و با شبکه اصلی ندارند و خوشه 6 از شش گره تشکیل شده است که گره‌ها رابطه توپولوژیکی با هم دارند ولی با شبکه گراف ندارند و همچنین 59 بردار ورودی در این گره‌ها نگاشت شده است و در مقایسه با خوشه‌بندی NG، در نتایج حاصل از خوشه‌بندی CNG نورون‌های منفردی که به نورون‌های شبکه گراف متصل نیستند وجود نداشت که این می‌تواند به دلیل استفاده ویژگی‌های مکانی بردارهای ورودی در روند خوشه‌بندی باشد.



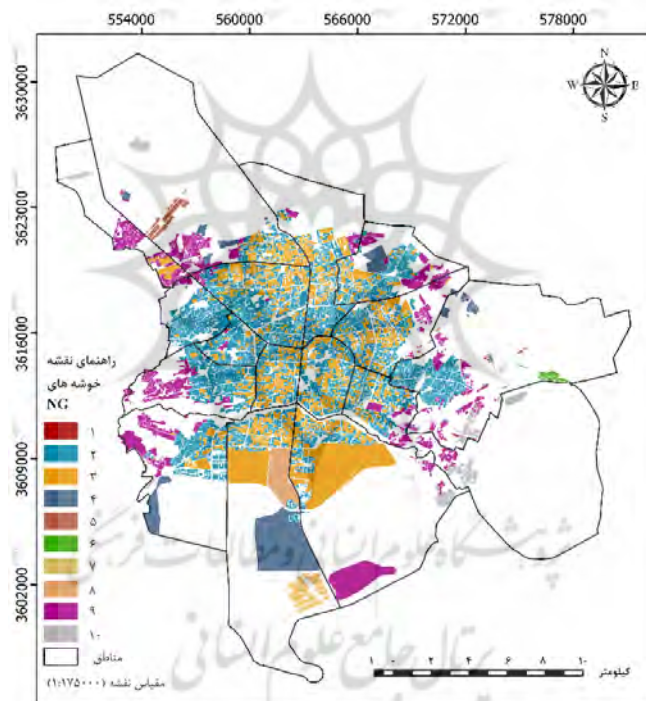
شکل ۳. شبکه گراف بلوک‌های آماری خوشه‌بندی شده شهر اصفهان حاصل از الگوریتم NG و CNG

نقشه بلوک‌های آماری خوشه‌بندی شده

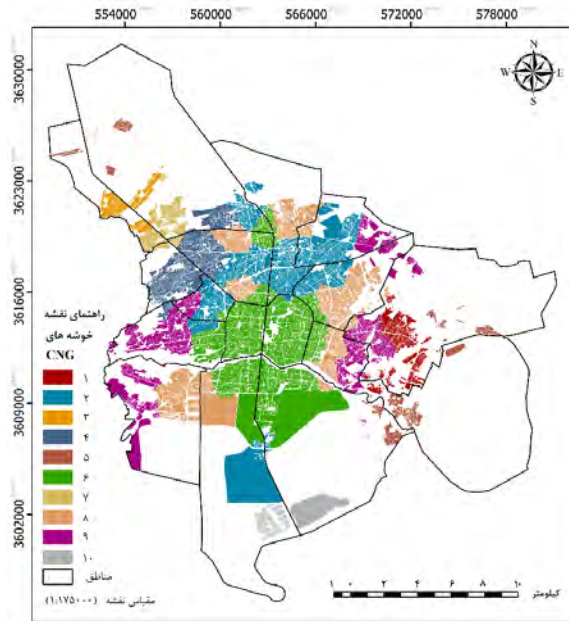
نقشه بلوک‌های آماری خوشه‌بندی شده کلان شهر اصفهان حاصل از الگوریتم NG در شکل (۴) و CNG در شکل (۵) نشان داده شده است. همان طور که از بررسی بصری نقشه‌ها مشاهده می‌شود نتایج حاصل از خوشه‌بندی هر دو الگوریتم کاملاً متفاوت است. الگوریتم NG، داده‌های بلوک‌های آماری را بر اساس شباهت میان متغیرهای مرتبط با توسعه پایدار شهری در خوشه‌هایی ناهمگن و پراکنده خوشه‌بندی کرده است. این بدان دلیل است که الگوریتم NG بلوک‌های آماری مشابه را از تمام نقاط شهر بدون در نظر گرفتن ویژگی‌های مکانی آن‌ها خوشه‌بندی کرده است. یکی از ویژگی‌های بارز داده‌های مکانی خودهمبستگی مکانی است یعنی داده‌هایی که از لحاظ فاصله مکانی نزدیک به یکدیگر هستند، تمایل به داشتن ویژگی‌های مشابه نیز دارند. پس اگر داده‌هایی نظیر بلوک‌های آماری که وابسته به مکان هستند خوشه‌بندی شوند در حالی که ویژگی‌های مکانی آن‌ها لحاظ نشده باشد، ممکن است به ایجاد خوشه‌هایی منجر شود که تنها از نظر چند متغیر مورد بررسی مشابه باشد و خوشه‌بندی مطلوب به دست نیامده باشد که کاملاً در نتایج خوشه‌بندی الگوریتم NG مشهود

است. الگوریتم CNG با به‌کارگیری ویژگی‌های مکانی و استفاده از معیار شباهت متغیرهای مرتبط با توسعه پایدار شهری در روند آموزش الگوریتم، داده‌های بلوک آماری را در خوشه‌هایی متجانس و همگن دسته‌بندی کرده است. از آنجائی که روند خوشه‌بندی و متغیرهای بردارهای ورودی هر دو الگوریتم یکسان و تنها تفاوت این دو الگوریتم در به‌کارگیری مختصات جغرافیایی مراکز هندسی بلوک‌های آماری به‌عنوان ویژگی‌های مکانی در روند خوشه‌بندی الگوریتم CNG است؛ بنابراین نتایج نشان‌دهنده تأثیر مثبت پارامترهای مکانی بر نتایج خوشه‌بندی بلوک‌های آماری است.

الگوریتم‌های NG و CNG، ۱۳۳۶۱ بلوک آماری را در ۱۰ خوشه طبق جداول ۲ و ۳ خوشه‌بندی کرده‌اند. در جدول (۲) که مربوط به نتایج حاصل از الگوریتم NG است، خوشه ۸ با تقریباً ۰,۰۲ درصد دارای کمترین تعداد (۴ بلوک) و خوشه ۲ با ۷۲,۱ درصد از کل بلوک‌ها دارای بیشترین تعداد (۹۶۴۲ بلوک) است. در نتایج حاصل از خوشه‌بندی مربوط به الگوریتم CNG، جدول (۳)، خوشه ۷ با ۳۳ بلوک آماری معادل ۰,۲۴ درصد و خوشه ۳ با ۱۴۳ بلوک آماری معادل ۱,۰۷ درصد کل بلوک‌ها دارای کمترین تعداد و خوشه ۶ با ۲۸۹۵ بلوک آماری معادل ۲۱,۶ درصد و خوشه ۲ با ۲۸۵۲ بلوک آماری معادل ۲۱,۳ درصد دارای بیشترین تعداد بلوک‌های آماری هستند.



شکل ۴. نقشه بلوک‌های آماری خوشه‌بندی‌شده شهر اصفهان حاصل از الگوریتم NG



شکل ۵. نقشه بلوک‌های آماری خوشه‌بندی شده شهر اصفهان حاصل از الگوریتم CNG

نتایج به دست آمده از اجرای الگوریتم CNG نشان می‌دهد که بر خلاف نتایج الگوریتم NG، بلوک‌های خوشه‌بندی شده اکثراً در یک محدوده جغرافیایی خاص محدود شده‌اند. بر این اساس می‌توان ویژگی‌های هر خوشه را به صورت ناحیه‌ای تحلیل نمود؛ بدین معنی که هر یک از خوشه‌ها عمدتاً در نواحی جغرافیایی مشخصی قرار دارند و در نتیجه این نواحی ویژگی‌های خوشه مورد نظر را به خود اختصاص خواهند داد. بنابراین با توجه به قرار گرفتن خوشه‌ها در یک منطقه خاص و در نظر گرفتن تمام متغیرهای مورد استفاده می‌توان ارزیابی کلی از وضعیت نواحی مختلف شهر اصفهان را از منظر توسعه پایدار ارائه نمود. به عنوان مثال بلوک‌های خوشه شماره ۶ واقع در مرکز و جنوب شهر اصفهان نه تنها از نظر معیارهای دسترسی به درمانگاه، بیمارستان، مراکز آتش‌نشانی، مراکز پلیس، مراکز ورزشی و کتابخانه در وضعیت مناسب‌تری نسبت به میانگین دسترسی کل شهر قرار دارند بلکه از نظر دیگر معیارها نظیر میانگین مساحت واحد مسکونی نیز میزان بالاتری را نسبت به میانگین کل شهر نشان می‌دهند. این خوشه با تراکم جمعیت و تعداد مهاجرین کمتر، نسبت جمعیت جوان متوسط، همچنین نسبت جمعیت متأهل، نرخ اشتغال، مساحت مسکونی، نسبت باسواد و نسبت محصلان بالاتر از میانگین، از وضعیت بهتری نسبت به کل شهر برخوردار است که بیانگر وضعیت مطلوب این ناحیه از منظر توسعه پایدار شهری است. در مقابل خوشه ۷ واقع در شمال غربی و خوشه ۲ در شمال شهر با داشتن تراکم جمعیت و تعداد مهاجرین بیشتر، نسبت جمعیت جوان بالاتر و همچنین نسبت جمعیت متأهل، مساحت مسکونی و نسبت باسواد کمتر از میانگین، از وضعیت نامناسب‌تری نسبت به کل شهر برخوردار هستند. همچنین در این خوشه‌ها دسترسی به درمانگاه، بیمارستان، مراکز آتش‌نشانی، مراکز پلیس، مراکز ورزشی و کتابخانه نیز در وضعیت نامطلوب‌تری نسبت به کل شهر قرار دارد. که در مجموع نشان می‌دهد این مناطق نسبت به مناطق جنوبی از لحاظ شاخص‌های توسعه پایدار در سطح پایین‌تری قرار دارند. در مقایسه با این دو خوشه، بلوک‌های خوشه ۱ واقع در مناطق شرقی شهر از لحاظ وضعیت شاخص‌های توسعه پایدار نزدیک به میانگین کل شهر قرار دارند.

جدول ۲. میانگین متغیرهای اجتماعی-اقتصادی در خوشه‌های حاصل از الگوریتم NG

میانگین	۱۰	۹	۸	۷	۶	۵	۴	۳	۲	۱	خوشه‌های NG
0/02	0/02	0/02	0/02	0/03	0/02	0/02	0/06	0/03	0/03	0/55	تراکم جمعیت در فضا
0/55	0/57	0/56	0/56	0/56	0/58	0/6	0/29	0/54	0/56	0/19	نسبت جمعیت متأهل
0/03	0/01	0/04	0/03	0/07	0/01	0/03	0/47	0/02	0/02	0/72	میزان مهاجرین واردشده
0/46	0/45	0/45	0/45	0/52	0/49	0/47	0/23	0/47	0/47	0/22	نرخ اشتغال
0/22	0/21	0/22	0/21	0/25	0/2	0/26	0/6	0/23	0/74	0/74	نسبت جمعیت جوان
۱۶۲۴	۱۱۸۳	۱۰۶۳	۱۰۲۱	۳۲۶۹	۵۵۴۵	۵۶۷۱	۸۳۴	۱۵۲۸	۳۰۵۱	۱۶۸۱	دسترسی به داروخانه
۱۰۷۸	۸۹۵	۹۲۳	۷۹۷	۳۱۳۲	۵۶۵۹	۱۹۷۷	۵۶۷	۹۰۲	۲۳۸۸	۱۳۱۵	دسترسی به درمانگاه
۱۹۹۶	۲۰۶۴	۱۷۵۵	۱۴۷۹	۲۸۸۴	۵۴۶۴	۲۴۴۷	۱۱۸۱	۲۲۸۱	۳۲۰۹	۱۴۰	دسترسی به بیمارستان
۱۴۳۰	۱۰۶۱	۱۲۸۶	۱۲۶۳	۳۴۷۶	۶۲۴۶	۳۲۲۸	۸۹۴۸	۱۱۹۷	۲۰۵۰	۱۰۲۱	دسترسی به آتش‌نشانی
۱۷۸۲	۱۸۵۵	۱۳۰۳	۱۳۴۱	۲۷۵۷	۴۹۰۷	۴۲۲۴	۱۳۴۷	۱۸۲۵	۲۶۱۳	۲۱۲۵	دسترسی به مراکز پلیس
۱۱۵	۱۱۳	۱۲۲	۱۲۶	۱۰۴	۱۲۲	۱۱۸	۱۰۹	۱۰۰	۱۱۰	۵۵	متوسط مساحت واحد مسکونی
۳۲	۶	۲۶۸	۲۳	۲۰	۴۱	۱۳	۱۱۰	۲۸	۲۲	۱۵۴	تعداد واحد مسکونی آپارتمانی
۳۰	۴	۳۱۰	۳۹	۱۰	۷۲	۱۳	۲۷	۱۷	۱۷	۶۱	میزان استحکام بنا
0/91	0/72	0/95	0/95	0/8	0/8	0/92	0/98	0/88	0/88	0/8	نسبت باسoadی
0/22	0/14	0/23	0/23	0/2	0/2	0/2	0/64	0/21	0/21	0/71	نسبت محصلان
۸۸۹	۱۱۲۱	۶۴۰	۶۹۹	۳۸۷۳	۷۵۴۶	۲۷۴۶	۶۱۲	۸۰۸	۱۰۶۸	۵۱۸	دسترسی به مراکز ورزشی
۲۸۸	۲۳۷	۲۹۹	۲۶۳	۶۵۴	۴۲۱	۵۴۱	۲۳۸	۲۵۲	۴۰۶	۳۸۸	دسترسی به مدرسه
۷۲۶	۶۷۲	۶۹۳	۶۴۳	۲۳۶۷	۵۵۰۷	۱۷۱۰	۵۷۹	۶۱۵	۸۱۴	۱۰۵۳	دسترسی به کتابخانه
۴۴۳	۴۵۵	۲۳۲	۲۶۶	۲۹۷۱	۶۲۸۰	۱۲۴۱	۵۵۱	۳۰۰	۶۰۳	۱۳۵۳	دسترسی به ایستگاه اتوبوس
۳۷۰۳	۵۱۰۴	۱۹۹۶	۲۵۹۸	۹۵۴۱	۳۸۴۳	۸۲۳۸	۲۶۳۱	۳۹۶۷	۵۳۱۵	۳۴۷۵	دسترسی به ایستگاه مترو
۱۵	۱۲	۱۷	۱۳	۳۰۲	۲۶	۲۸	۱۶	۱۲	۲۳	۳۲	دسترسی به معابر عمومی
۱۳۳۶۱	۴۵۳	۱۴۴۸	۴	۲۰	۵۹	۴۹	۴۶	۱۶۲۵	۹۶۴۲	۱۵	تعداد بلوک‌های هر خوشه
۱۰۰	3/39	10/8	0/02	0/14	0/44	0/36	0/34	12/1	72/1	0/11	درصد بلوک‌های هر خوشه

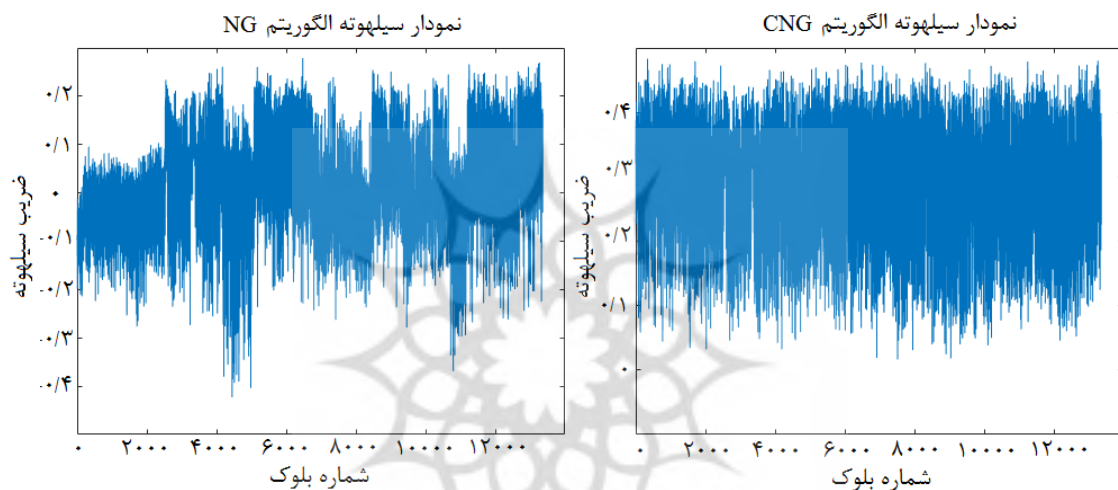
جدول ۳. میانگین متغیرهای اجتماعی-اقتصادی در خوشه‌های حاصل از الگوریتم CNG

میانگین	۱۰	۹	۸	۷	۶	۵	۴	۳	۲	۱	خوشه‌های CNG
0/02	0/03	0/01	0/02	0/03	0/03	0/03	0/03	0/02	0/02	0/02	تراکم جمعیت در فضا
0/55	0/55	0/56	0/54	0/54	0/55	0/58	0/55	0/59	0/43	0/56	نسبت جمعیت متأهل
0/03	0/07	0/01	0/04	0/04	0/03	0/03	0/03	0/02	0/04	0/03	میزان مهاجرین واردشده
0/46	0/41	0/48	0/45	0/42	0/44	0/47	0/46	0/54	0/46	0/48	نرخ اشتغال
0/22	0/24	0/22	0/22	0/24	0/2	0/23	0/23	0/22	0/28	0/23	نسبت جمعیت جوان
۱۶۲۴	۸۹۰	۲۶۹۷	۱۳۲۵	۲۳۴۶	۷۲۱	۶۳۰۸	۹۲۳	۵۲۳۹	۱۴۶۰	۱۵۰۶	دسترسی به داروخانه
۱۰۷۸	۸۸۷	۱۳۵۱	۱۰۲۳	۳۳۷۹	۵۰۷	۲۳۶۸	۱۰۸۱	۴۷۰۹	۱۰۸۷	۱۱۵۵	دسترسی به درمانگاه
۱۹۹۶	۱۰۵۶	۲۸۲۸	۱۷۵۰	۴۲۵۳	۸۴۰	۲۹۱۳	۳۵۸۶	۱۵۳۱	۲۶۵۰	۱۱۲۲	دسترسی به بیمارستان
۱۴۳۰	۱۰۱۵	۱۹۲۹	۱۶۷۵	۱۱۷۸	۱۰۳۵	۳۸۲۵	۱۱۰۱	۲۱۰۷	۱۶۳۷	۱۶۵۸	دسترسی به آتش‌نشانی
۱۷۸۲	۱۳۵۸	۲۸۸۵	۱۳۶۳	۲۴۰۲	۱۰۷۱	۴۶۰۰	۱۰۳۹	۱۴۲۷	۱۹۰۲	۲۱۵۲	دسترسی به مراکز پلیس
۱۱۵	۱۱۸	۱۳۸	۱۲۲	۸۴	۱۴۹	۱۱۹	۱۰۸	۱۲۸	۹۲	۱۲۰	متوسط مساحت واحد مسکونی
۳۲	۲۱	۲۰	۳۹	۲۶	۳۸	۱۳	۳۳	۳۲	۴۰	۲	تعداد واحد مسکونی آپارتمانی
۳۰	۲۶	۱۵	۴۱	۲۲	۴۴	۱۳	۲۹	۳۵	۲۹	۲	میزان استحکام بنا
0/91	0/88	0/89	0/84	0/91	0/93	0/89	0/88	0/9	0/89	0/9	نسبت باسoadی
0/22	0/27	0/27	0/21	0/22	0/22	0/2	0/21	0/2	0/21	0/22	نسبت محصلان
۸۸۹	۴۴۸	۸۶۱	۷۷۷	۵۱۳	۵۳۱	۳۴۶۰	۸۶۱	۸۵۲	۶۶۵	۲۳۶۹	دسترسی به مراکز ورزشی
۲۸۸	۲۹۵	۳۰۴	۲۷۸	۲۷۸	۲۲۲	۵۴۳	۳۲۱	۳۹۰	۲۷۹	۳۱۰	دسترسی به مدرسه
۷۲۶	۸۳۰	۷۳۷	۶۵۵	۵۶۸	۵۵۵	۱۸۵۸	۵۸۲	۸۹۷	۹۷۴	۱۱۰۳	دسترسی به کتابخانه
۴۴۳	۲۶۶۵	۲۸۹	۲۱۷	۲۱۹	۱۵۹	۱۸۴۰	۲۸۰	۳۵۵	۲۰۷	۱۲۸۳	دسترسی به ایستگاه اتوبوس
۳۷۰۳	۴۶۵۷	۶۳۸۳	۳۸۴۲	۱۰۰۵	۱۸۵۷	۸۹۴۹	۲۱۷۹	۳۷۹۰	۲۴۶۲	۸۵۴۹	دسترسی به ایستگاه مترو
۱۵	۱۶	۱۷	۱۴	۱۹	۱۲	۳۲	۱۳	۱۸	۱۳	۲۶	دسترسی به معابر عمومی

۱۳۳۶۱	۵۰۰	۱۷۹۸	۲۳۳۴	۳۳	۲۸۹۵	۵۲۲	۱۲۸۷	۱۴۳	۲۸۵۲	۶۹۷	تعداد بلوک‌های هر خوشه
۱۰۰	3/74	13/4	17/4	0/24	21/6	3/9	9/63	1/07	21/3	5/21	درصد بلوک‌های هر خوشه

ارزیابی کیفیت خوشه‌بندی

مقادیر ضریب سیلهوته به دست آمده برای هر یک از بلوک‌های آماری خوشه‌بندی شده توسط الگوریتم‌های NG و CNG (شکل ۶) نشان می‌دهد که مقادیر این ضریب برای اکثر بلوک‌های آماری خوشه‌بندی شده توسط الگوریتم NG منفی است. این در حالی است که ضرایب سیلهوته محاسبه شده برای بلوک‌های آماری خوشه‌بندی شده توسط الگوریتم CNG دارای مقادیر مثبت هستند. همچنین کیفیت کلی نتایج خوشه‌بندی با استفاده از میانگین ضرایب سیلهوته برای کل مجموعه بلوک‌های آماری نشان می‌دهد که الگوریتم CNG با مقدار میانگین ۰/۲۹ نسبت به الگوریتم NG با میانگین ۰/۰۲- کیفیت خوشه‌بندی مناسب‌تری دارد. این نتایج بیانگر تأثیر مثبت پارامترهای مکانی در فرایند خوشه‌بندی است.



شکل ۶. نمودار مقادیر سیلهوته بلوک‌های آماری خوشه‌بندی شده با الگوریتم‌های NG و CNG

بحث

مقایسه نتایج حاصل از خوشه‌بندی بلوک‌های آماری با به‌کارگیری الگوریتم CNG در این تحقیق، در مقایسه با سایر تحقیقات مشابه نشان از همسویی نتایج به دست آمده دارد. به‌عنوان مثال Hagenauer و Helbich در سال ۲۰۱۳، برای ارزیابی عملکرد الگوریتم CNG در خوشه‌بندی داده‌ها با اعمال ویژگی‌های مکانی، دو مجموعه داده مصنوعی و دنیای واقعی را با این الگوریتم خوشه‌بندی کردند. یافته‌های این تحقیق نشان داد الگوریتم CNG در مقایسه با NG در هر دو مجموعه داده مصنوعی و واقعی خوشه‌هایی همگن‌تر و از نظر تئوریک صحیح‌تر را دسته‌بندی می‌کند که نشانگر تأثیر مثبت ویژگی‌های مکانی داده در روند خوشه‌بندی است. همچنین در تحقیق Hagenauer و Helbich در سال ۲۰۱۶، نتایج خوشه‌بندی داده‌های اقتصادی-اجتماعی بلوک‌های آماری شهر فیلادلفیا نشانگر تفاوت‌های قابل توجه خوشه‌های به دست آمده با اعمال ویژگی‌های مکانی در فرایند خوشه‌بندی است که این موضوع با نتایج به دست آمده در تحقیق حاضر انطباق دارد.

نتیجه‌گیری

در این تحقیق شبکه‌های عصبی گازی برای خوشه‌بندی بلوک‌های آماری شهر اصفهان مبتنی بر متغیرهای مرتبط با توسعه پایدار شهری بکار گرفته شد. در پژوهش حاضر با هدف بررسی تأثیر پارامترهای مکانی در نتایج خوشه‌بندی شبکه‌های عصبی گازی، مختصات جغرافیایی مراکز هندسی بلوک‌های آماری به‌عنوان پارامتر مکانی در کنار سایر متغیرهای ورودی به الگوریتم شبکه‌های عصبی گازی ارائه شدند. بررسی نتایج حاصل از الگوریتم با اعمال و بدون اعمال پارامترهای مکانی نشانگر تأثیر مثبت پارامترهای مکانی در روند خوشه‌بندی و ایجاد خوشه‌های همگن است. میانگین کلی ضریب سیلهوته در ارزیابی بلوک‌های خوشه‌بندی شده بیانگر این است که الگوریتم CNG، نسبت به الگوریتم NG یک خوشه‌بندی مناسب و منسجمی از بلوک‌های آماری را با توجه به ویژگی‌های مکانی و توصیفی آن‌ها ایجاد می‌کند. در این تحقیق پتانسیل الگوریتم CNG در تعریف مناطق همگن و شناسایی بلوک‌های مشابه در یک مجموعه داده واقعی نشان داده شد. شناسایی مناطق همگن در قالب خوشه‌های مجزا با خصوصیات معین، خود عاملی مؤثر جهت برنامه‌ریزی‌های محیطی متمرکز کوتاه‌مدت و بلندمدت است. در نتیجه به‌کارگیری الگوریتم‌های خوشه‌بندی مبتنی بر مکان نظیر CNG، امکان استخراج مناطق همگن جهت تسهیل در مدیریت و برنامه‌ریزی شهری را فراهم می‌سازد.

از دیگر الگوریتم‌های خوشه‌بندی می‌توان به الگوریتم نقشه‌های خود سازمانده (SOM) اشاره کرد. SOM مبتنی بر یک شبکه عصبی بدون نظارت است که معمولاً برای نگاشت داده‌های با ابعاد بالا به نقشه‌های ویژگی عمده‌تاً دو بعدی که شبکه واحدها یا نورون‌ها هستند استفاده می‌شود (Wankhede, 2014). در طول فرایند آموزش در این الگوریتم هر داده با تمام واحدها مقایسه می‌شود و شبیه‌ترین واحد انتخاب می‌شود. الگوریتم خوشه‌بندی Geo-SOM نیز الگوریتمی بر پایه SOM است با این تفاوت که جنبه مکانی دارد؛ به طوری که جستجوی شبیه‌ترین واحد را علاوه بر داده‌های توصیفی، با توجه به ویژگی‌های مکانی بردارهای ورودی نیز انجام می‌دهد (Hagenauer and Helbich, 2016). لذا پیشنهاد می‌شود در تحقیقات آتی خوشه‌بندی بلوک‌های آماری مبتنی بر متغیرهای توسعه پایدار شهری با الگوریتم Geo-SOM نیز انجام و نتایج آن با نتایج حاصل از الگوریتم CNG مقایسه گردد.

حامی مالی

این اثر حامی مالی نداشته است.

سهام نویسندگان در پژوهش

نویسندگان در تمام مراحل و بخش‌های انجام پژوهش سهم برابر داشتند.

تضاد منافع

نویسندگان اعلام می‌دارند که هیچ تضاد منافی در رابطه با نویسندگی و یا انتشار این مقاله ندارند.

تقدیر و تشکر

نویسندگان از همه کسانی که در انجام این پژوهش به ما یاری رساندند، به‌ویژه کسانی که کار ارزیابی کیفیت مقالات را انجام دادند، تشکر و قدردانی می‌نمایند.

منابع

- امیری، مجتبی؛ رحمانیان، مجید و غفاری، علی. (۱۳۹۲). بررسی وضعیت عوامل فرهنگی مؤثر بر مدیریت توسعه پایدار شهر تهران. *مدیریت دولتی*، ۵ (۴)، ۱-۱۹. doi:10.22059/jipa.2013.50386
- برزگر، صادق؛ دیوسالار، اسدالله؛ صفرعلی‌زاده، اسماعیل و فنی، زهره. (۱۳۹۷). تحلیل شاخص‌های پایداری کالبدی در شهرهای کوچک (مطالعه موردی: شهرهای کوچک استان مازندران). *فصلنامه فضای جغرافیایی*، ۱۸ (۶۱)، ۱۶۱-۱۸۰.
- برقی، اسماعیل. (۱۳۹۷). سوادآموزی عنصر کلیدی توسعه پایدار. *نشریه راهبرد توسعه*، ۱۴ (۴)، ۱۸۷-۲۱۰.
- سپهوند، رضا و عارف‌نژاد، محسن. (۱۳۹۲). اولویت‌بندی شاخص‌های توسعه پایدار شهری با رویکرد تجزیه و تحلیل سلسله مراتبی گروهی مطالعه موردی در شهر اصفهان. *مطالعات ساختار و کارکرد شهری*، ۱ (۱)، ۴۳-۵۹. doi: 20.1001.1.20085362.1391.23.4.12.6
- عباسی، محمدباقر؛ توده فلاح، معصومه؛ خطیبی، آمنه و صفاکیش، محدثه (۱۳۹۷). نگاهی به ساخت سنی و وضع مشارکت در بازار کار مردان و زنان در معرض ازدواج و طلاق از دریچه سرشماری، *فصلنامه جمعیت*، ۲۵ (۱۰۵ و ۱۰۶)، ۵۷-۷۶.
- علیپور، علی‌رضا. (۱۳۸۳). *آشنایی با نظریه گراف*. تهران، چاپ اول، انتشارات فاطمی.
- محمدزاده، رحمت. (۱۳۹۴). بررسی تطبیقی الگوی مجتمع‌های مسکونی ویلائی و آپارتمانی (مطالعه موردی شهر جدید سهند). *نشریه جغرافیا و برنامه‌ریزی*، ۱۹ (۵۴)، ۲۷۹-۳۰۲. doi: 20160515142421-9918-210
- محمودی، محمدرضا؛ اسلامیان، سعید؛ گوهری، علیرضا و طحانیان، معین (۱۴۰۱). بررسی عملکرد شبکه‌های عصبی گازی در خوشه‌بندی هیدرولوژیک. *مجله مدیریت آب و آبیاری*، ۱۲ (۲)، ۳۵۹-۳۷۳. doi:10.22059/JWIM.2022.339537.972
- موسوی، میرسعید. (۱۳۹۷). بررسی سطح تحقق توسعه پایدار شهر تبریز بر اساس شاخص ردپای بوم‌شناختی. *فصلنامه جغرافیا و مطالعات محیطی*، ۷ (۲۷)، ۶۱-۷۶.
- نصیری دارانی، شهربانو. (۱۴۰۱). تحلیل حساسیت روش ارزیابی چند معیاره مکانی به تغییر توابع استانداردسازی و وزن معیارها (مطالعه موردی: سنجش وضعیت پایداری توسعه در شهر اصفهان). *پایان‌نامه کارشناسی ارشد*، دانشکده علوم زمین، دانشگاه شهید بهشتی تهران.
- نصیری هندخاله، اسماعیل؛ حسینی فر، سید محسن و احمدی، علی. (۱۳۹۵). تأثیر مهاجرت بر توسعه شهری با استفاده از مدل SWOT (مورد مطالعه: شهر بابل)، *دو فصلنامه پژوهش‌های بوم‌شناسی شهری*، ۷ (۲)، ۵۵-۶۶. doi: 20.1001.1.25383930.1395.7.14.4.6

References

- Abbasi, M., Tude Fallah, M., Khatibi, A., & Safakish, M. (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. *Journal Population*, 20(5), 57-76 [In Persian].
- Aldegheshem, 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/jisd.v7n2p113.
- Alipour, A (1383). *Familiarity with graph theory*. Tehran, first edition, Fatemi Publications. [In Persian].
- Amiri, M., Rahmanian, M., & Ghaffari, A. (2012). Investigating the status of cultural factors affecting the management of sustainable development in Tehran. *Public Administration*, 5(4), pp. 1-19. doi: 10.22059/jipa.2013.50386. [In Persian].
- Andrienko, G., Andrienko, N., Bak, P., Bremm, S., Keim, D., von Landesberger, T., & Schreck, T. (2010). A framework for using self-organising maps to analyse 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.
- Barghi, E (2017). Literacy the key element of sustainable development. *Journal of Development Strategy*, 14(4), 187-210. [In Persian].
- Barzegar, S., Divsalar, A., Safaralizadeh, I., & Fanni, Z. (2017). The Analysis of indicators of

- physical sustainability in small cities, (case study: small cities of Mazandaran province). *Journal of Geographical Space*, 18(61), pp. 161-180. [In Persian].
- Grubestic, 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. doi: 10.1080/00045608.2014.958389.
- Hagenauer, J. (2015). Clustering contextual neural gas: a new approach for spatial planning and analysis tasks. *Computational approaches for urban environments*, 77-94. doi.org/10.1007/978-3-319-11469-9_4.
- Hagenauer, J., & Helbich, M. (2013). Contextual neural gas for spatial clustering and analysis. *International Journal of Geographical Information Science*, 27(2), 251-266. doi:10.1080/13658816.2012.667106.
- 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.
- Hagenauer, J., & Helbich, M. (2018). *The Application of the SPAWNN Toolkit to the Socioeconomic Analysis of Chicago, Illinois*. Trends in Spatial Analysis and Modelling: Decision-Support and Planning Strategies, 75-90. doi: 10.1007/978-3-319-52522-8_5
- Han, J., Kamber, M., & Mining, D. (2006). Concepts and techniques. *Morgan kaufmann*, 340, 94104-3205.
- 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. doi: 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. doi: 10.1145/331499.331504.
- Kasabov, N. K. (1996). *Foundations of neural networks, fuzzy systems, and knowledge engineering*. Marcel Alencar.
- Kaufman, L., & Rousseeuw, P. J. (2009). *Finding groups in data: an introduction to cluster analysis*. John Wiley & Sons.
- Labusch, K., Barth, E., & Martinetz, T. (2009). Sparse coding neural gas: Learning of overcomplete data representations. *Neurocomputing*, 72, 1547-1555. https://doi.org/10.1016/j.neucom.2008.11.027.
- 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. doi: 10.1051/e3sconf/201911002093.
- Mahmoudi, M., Islamian, S., Gohari, A., & Tahanian, M (2021). Investigation of the performance of neural gas networks in hydrological clustering. *Journal of Water and Irrigation Management*, 12(2), 359-373. doi:10.22059/JWIM.2022.339537.972. [In Persian].
- Marino, A. (2018). Graph Clustering Algorithms. Ph.D. Course on Graph Mining Algorithms, Università di Pisa.
- Martinetz, T. (1993). Competitive Hebbian learning rule forms perfectly topology preserving maps. In *ICANN'93: Proceedings of the International Conference on Artificial Neural Networks Amsterdam, The Netherlands 13-16 September 1993* 3 (pp. 427-434). Springer London. doi: 10.1007/978-1-4020-0000-0_27.
- Martinetz, T. M., Berkovich, S. G., & Schulten, K. J. (1993). Neural-gas' network for vector quantization and its application to time-series prediction. *IEEE transactions on neural networks*, 4(4), 558-569. doi: 10.1109/72.238311.
- Martinetz, T., & Schulten, K. (1991). *A "neural-gas" network learns topologies*. Artificial Neural Networks.
- Miller, H. J. (2010). The data avalanche is here. Shouldn't we be digging?. *Journal of Regional Science*, 50(1), 181-201. doi:abs/10.1111/j.1468-2421.2009.00641-1.x.
- Mohammadzadeh, R. (2014). Compatibility Survey of Detached and Apartment Residential Complexes Pattern in Sahand New Town. *Journal of Geography and Planning*, 19(54), 279-302. doi: 20160515142421-9918-210 [In Persian].
- Mousavi, M. (2017). Evaluation of the sustainable development level in Tabriz city based on

- ecological footprint index. *Journal of Geography and Environmental Studies*, 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)*. MSc Thesis. Shahid Beheshti University [In Persian].
- Nasiri Hende Khaleh, E., Hoseinifar, S. M., & Ahmadi, A. (2017). The Impact of Migration on Urban Development Using SWOT, Case study: Babol city. *Journal of Urban Ecology Researches*, 7(14), 55-66. doi: 20.1001.1.25383930.1395.7.14.4.6 [In Persian].
- Nielsen, M. A. (2015). *Neural networks and deep learning* (Vol. 25, pp. 15-24). San Francisco, CA, USA: Determination press.
- Openshaw, S. (1999, July). Geographical data mining: key design issues. *In Proceedings of Geo Computation (Vol. 99)*. doi: 10.1007/978-3-642-17316-5_55.
- Patel, P., & Patel, A. (2021, June). Use of sustainable green materials in construction of green buildings for sustainable development. *In 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. doi: 10.1002/sd.2005.
- Sepahvand, R., & Arifnejad, M. (2012). Prioritization of indicators of urban permanent development with a group analytic hierarchy proces (Case study: in Isfahan city). *Journal of Urban Structure and Function studies*, 1(1), 43-59. doi: 20.1001.1.20085362.1391.23.4.12.6 [In Persian].
- Sheela, K. G., & Deepa, S. N. (2012, August). An efficient hybrid neural network model in renewable energy systems. *In 2012 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: Modelling and Control*, 16(4), 488-504. doi:10.15388/NA.16.4.14091.
- 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. doi:10.1111/j.1467-8306.2004.09402003.
- Tobler, Waldo R. A. (1970). computer movie simulating urban growth in the Detroit region. *Economic geography*, 46, no. sup1, 234-240. doi.org/10.2307/143141.
- Van Dongen, S. M. (2000). *Graph clustering by flow simulation* (Doctoral dissertation).
- Wadhwa, L. C., (2000). Sustainable transportation: the key to sustainable cities, doi: 10.2495/URS000301.
- 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. doi:10.9790/0661-16378692
- Yuan, M., Battenfield, B., Gahegan, M., & Miller, H. (2004). Geospatial data mining and knowledge discovery. *A research agenda for geographic information science*, 3, 365.
- 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. doi:10.5772/51702.