Ontology Space International Journal
ISSN: 2345-6450
OISSN: 2423-7299
Spring 2024, Vol. 13, Issue 2, No. 49, Pages: 15- 24
Journal homepage: soij.qiau.ac.ir

Comparison of machine learning methods in the classification of Bandar Kong windcatchers

Mona Mohtaj [a], Mansoureh Tahbaz [b,*], Atefeh Dehghan Touran Poshti[c]

[a] *Department of Architecture College of Art and architecture, West Tehran Branch, Islamic Azad University, Tehran, Iran*
[b] *Faculty of Architecture and Urbanism, Shahid Beheshti University, Tehran , Iran.*
[c] *Department of Architecture College of Art and architecture, West Tehran Branch, Islamic Azad University, Tehran, Iran*

Received: 01 June 2023 - Accepted: 29 August 2024

Abstract

The hot and humid region of Iran is one of the hardest climates in the world. Due to its proximity to the sea and in order to use of coastal winds, windcatcher is one of the architectural elements of these areas, including Bandar Kong. Classification of architectural types is the first step in understanding the features governing architecture. This research aims to classify the windcatchers of Bandar Kong using machine learning methods. For this purpose, the plans of Bandar Kong have been categorized in two General ways, based on shape and characteristics of plans and the results have been compared. In the first method, the similarity of 35 windcatchers is calculated using the Cosine Distance method in Anaconda3.9. In second step plans are are clustered using using Clust map from Seaborn Library. and classified in Anaconda using complete linkage and average linkage methods from Numpy library. The results show that cosine similaty method is the most appropriate method for clustering plans and divide plans into 6 clusters. the most abandance windcatchers are located at east part of the plan, length to width ratio is 1to 2.2-2.5 and windcatcher rooms have rectangular proportion. The windcatcher is located at south part of the room. The abundance of this cluster is 34 percent.

Keywords: Vernacular architecture; Machine learning; Similarity; Clustering; Anaconda

1. Introduction

Creating the ability to use new scientific knowledge has made it necessary to combine and integrate the sciences of different disciplines and has led to scientific progress and improvement in various fields (Shah Hosseini et al., 1401). Machine learning methods can be used to extract information from data. These systems use the patterns in the data for analysis and inference. Machine learning algorithms can solve complex problems that are impractical to do manually (Varzan, 6, 1400). One of the capabilities of machine learning methods that have been used recently in architectural research is similarity measurement 2 and classification 3 of architectural images. Categorizing and describing the architectural features of each category plays an important role in recognizing architectural types (Araldi, et al, 2021, 113).
Climatic diffrences and a long history of housing in Iran have caused valuable achievements in the field of architecture and urban planning (Tahbaz et al., 2012). Classification and description building typologies plays o fundamental role in understanding architectural types. This paper presented an innovative way for clustering and description of nuilding types. This method allows to identify groups of buildings wich have specific simillar morphological characteristics.

* Corresponding Author Email: m-tahbaz @sbu.ac.ir

This research aims to compare two different methods of machine learning to cluster windcatchers of Bandar Kong. First method is based on windcatcher images and second method is based on characteristics wich are extracted from windcatchers plans.

2. Research Background

Classification and similarity measurement are mainly used in the Data-mining process, the purpose of classification is to sort the data into different groups to find existing patterns. In this part of the article, the research background worked on similarity measurement and classification of architectural plans has been reviewed. The word cluster was used for the first time by Carter and Whitehead (Carter & Whitehead, 1975). Their research aims to classify spaces based on spatial relationships in secondary schools. The aim of this research is to achieve the most optimal possible mode for placing spaces in floors based on function. In this research, hierarchical method 5 has been used to cluster architectural plans. In recent research, machine learning methods have been used more widely in the field of similarity measurement and classification of architectural plans.
In a research, Feist et al. classified the architectural plans of residential apartments located in 6 different blocks. The purpose of this research is the modularization of health

spaces for their implementation. The plans are categorized by the C-O-P-Cummins 6 algorithm and at this stage the plan and its selected cluster are approved, rejected or assigned to another category by the user. This method is used by C+7 plugin which can be implemented in View software. Python software has also been used to perform semi-supervised clustering of 8 plans. Semi-supervised classification includes two parts: first, the classification is done by the software and then it is confirmed by the user (Feist, et al, 2022).

In a research, Shih and Peng have investigated the similarity between architectural plans by using similarity measurement algorithms. Their research shows Euclidean distance method to measure the similarity of the plans did not have favorable results, so the cosine distance method was used to measure the similarity of the plans. (Shih & Peng, 2022)

In 2017, Rodríguez and his colleagues conducted a research on seventy-two residential plans with artificial intelligence algorithms and in Java script by four methods of point distance12, rotation function13, grid matrix14 and tangent distance15, the plan have been converted into vectors.. Therefore, the images have been recovered by these four methods and the hierarchical method, which is one of the unsupervised machine learning methods, has been used for classification. The purpose of this research is to choose the best method for retrieving architectural plans by artificial intelligence. The results of this research show that using the tangent distance method has the best results for classifying plans in this research (Rodrigues et al, 2017).

In Ron Jaio's research, 129 architectural plans related to the contemporary era have been collected, and these plans were designed in 16 architectural offices. Jean Raio used a network method to compare the plans. In the first stage, the design patterns are extracted from the plans. The plans have been divided on a grid regardless of the scale and according to the functions, they have been rotated and reversed in different directions to increase the number of comparisons of the plans. In this method, the similarity of the plans is measured by dividing the number of overlapping cells by the total number of cells. Hierarchical classification method (UPGMA) in Python were used to classify the plans with Skippy library. Finally, Jean Raio has achieved four design patterns used in the examined plans (Xiao, 2021).

In 2021, in a research, Sun and Hyun investigated the similarity of plans based on the four criteria of proximity, communication, the shape of spaces and the number of rooms. A 32 x 32 grid was used to retrieve the images and finally this grid was converted into a binary vector. To quantitatively measure the similarity between plans, the cosine distance method is used. This information was entered in the Jason 18 environment and categorized in the Python environment using the JMet for Pi 19 library (Son & Hyun, 2021).

In 2022, Sun and Hyun, in a separate research, investigated the similarity of plans in terms of shape similarity, the number of spaces and the relationship between spaces. In this research, the similarity of the plans has been investigated through the network matrix and the similarity of the spatial communication graphs through the GE-D20 algorithm (Son, & Hyun, 2022).

In a research, Shermin Youssef and Yan have studied the classification of architectural plans with a new method with a plugin in Grasshopper written by Python. The recovery of architectural plans in this research is done through a network matrix. Also, to increase the number of comparisons, the plans have been rotated ,mirrored and compared in these cases. Hungarin's algorithm was used for similarity measurement. This algorithm is considered one of the optimization algorithms and selects the most optimal possible mode among the similarity measurement modes (yousif & Yan, 2019).

In 2021, Aldro and his colleagues have studied the classification and typology of residential houses in France. In this research, the unsupervised method has been used for classification because unsupervised methods lead to classification based on initial characteristics of the plans without the need to predict the number of clusters, the features and characteristics of each clusters (Aldero, et al, 2021, 115). The method used for the unsupervised clustering is the hierarchical classification Average linkage method, which finally reached ten different types, characteristics of each type are described at the end.

Evan and Babakhani have investigated the bridges of the Safavid era in a research. The purpose of this research is to investigate the fact that the form of bridges is related their function. For this purpose, the cross-section similarity of the bridges has been measured using the Cosine distance method (Masawat, Babakhani, 1400). Also, in a research, Shah-Mohammadi investigated the color spectrums of Nasir-ul-Molk Mosque's glasses. For this purpose, the glasses were converted into machine language as color images with RGB codes and their similarity was checked by the cosine similarity relation (Shah Mohammadi, Babakhani, 1400). In another research, the similarity measurement of 30 Qajar house plans was done using the cosine distance method and the houses with a high percentage of similarity were selected (Babakhani, Keifari, 2021).

Regarding the background of research scope, this research aims to cluster windcatcher plans of Bandar kong in two different ways based on shape and characteristics. For this purpose machine learning methods are used with unsupervised method and the plans that are most similar to other plans are identified by Cosine Distance similarity method. the differences between the methods used for clustering and similarity measurement in current research with other researches are shown in table (1). The following results can be extracted from the comparison and review of the research that has been done in the field of similarity measurement and classification of architectural plans.

1. Retrieving the shape of the architectural plan using the grid matrix is the most common form of converting the plan into its binary matrix. This method can also include

the function of space. For example, Jean Raio used a network method to restore the form and function of the plan (Xiao, 2021).
2. In most researches, the similarity measurement has been done by Cosine distance method. In Shih and Peng's research, using the cosine distance method to measure the similarity of plans has better results than the Euclidean distance (Shih & Peng, 2022). The cosine similarity criterion is very suitable for evaluation, especially in sparse vectors, and it is used especially in the positive space which has a range of [0.1] (Farhai and Jamzadeh, 2017, 21).
3. In the scope of research, two clustering methods, supervised and unsupervised, have been used. Unsupervised clustering can be used to discover patterns and relationships in architectural plans. (Xiao, 2021, 188). But if the classification is based on a specific criterion, the supervised classification method can be a more suitable method.

3. Theoretical Framework

The current research aims to categorize and measure the similarity between architectural plans. Therefore, the theoretical foundations can be examined in three sections: retrieving architectural plan, similarity measurement and clustering.

3.1. Retrieving architectural plan

In a general definition, image retrieving is a method of defining its features and characteristics with machine language. The architectural plan can be transformed into a vector or matrix in the following ways.

3.1.1. Point distance method

Chang et al. have presented a method of expressing the shape in machine language based on the distance of the center of the shape (Chang, et al, 1991). In this way, a point is considered in the center of the shape and the distance of all points from this center is measured and recorded in a matrix. In the research conducted by Rodriguez, this method was used to retrieve architectural plans (Rodrigues et al, 2017).

3.1.2. Turning function method

Arkin and colleagues (Arkin, et al, 1991) have presented a method known as the turning function method. In this method, the cosine of the angle of the specified point on the figure, counterclockwise with the x-axis, is measured and written in the matrix of the figure. In this method, enlarging its shape and scale will not affect the result. But the selection of the starting point is very important and it should be noted that in all the studied forms, the starting and ending points are selected according to a specific rule. In their research, Hue and colleagues have used turning function method to recover architectural plans (Hue, et al, 2020).

3.1.3. Tangent distance method

In the tangential distance method, the tangent line passing through each of the points is calculated to the central point and entered in a matrix. In architectural plans, these lines actually form the external walls. The figure is divided into points with a certain distance from the upper left point and counter-clockwise. At each point, a line tangent to the figure is drawn at that point. The distance between this point and the central point is calculated. These numbers are finally normalized between zero and one, and the number one represents the greatest distance (Rodrigues et al, 2017).

3.1.4. Grid based method

For this purpose, a grid is placed on the image. The result of this retrieving is displayed as zero and one codes. In other words, the number one is considered in the parts where the image overlaps with the grid and zero in the empty parts. (Sajjanhar, lu, 2004). Therefore, the shape is displayed as a matrix or vector. According to the review of the conducted researches, the retrieving of architectural plans has been done in the form of a network matrix in most cases.

Table 1
background of research scope

Study Number	Author(s)	Year of Publication	**Similarity method**	**Clustering method**	**Functional similarity**
(1)	**Feist et al**	2022	Revit C#	Python COP-KMEANS	-
(2)	**Shih & peng**	2022	layoutGMN Cosine similarity	-	-
(3)	**Son & Hyun**	2022	matrix	-	B Python-GED Gmath4py
(4)	**Aldero et al**	2021	-	Hierarchial average linkage	-
(5)	**Son & Hyun**	2021	Cosine Similarity	-	B Python-GED Gmath4py
(6)	**Keifari and Babakhani**	2021	Cosine Similarity	Pythone hierarchial	-

(7)	**Ran Xiao**	2020	matrix	Scipy sklearn Complete linkage	Polyomino Grid- Based
(8)	**Hu et al**	2020	CNN	-	CNN(Convolutional Neural Network
(9)	**Yusif & Yan**	2019	Hongarian Algorithm & cosine Similarity	K- medoid Grasshopper plugin with python	-
(10)	**Rodrigues et al**	2017	Cosine similarity	hierarchial	-
(11)	**Mosavat 2t al**	۱۳۹۹	Cosine Similarity	-	-
(12)	**Babakhani et al**	1400	Cosine Similarity	-	-
This study			Cosine Similarity	Hierachial average and complete linkage	This study work on windcatchers

3.2. Similarity measurement

in Python software Similarity is used to measure the distance between two points or numbers or matrices. To measure the similarity, the feature vector is extracted from the image and the distance between the images is obtained by one of the similarity methods (Farhai and Jamzadeh, 2017, 16). There are different methods to measure similarity, Euclidean distance, Manhattan distance and cosine distance are some of these methods.

3.2.1. Euclidean distance similarity

When the data are close to each other, this criterion is a suitable tool for measuring similarity. In this method, the distance between two points is obtained through Fithaqorth's theory (Farhai and Jamzadeh, 2017, 20).

3.2.2. Manhattan distance similarity

criterion Manhattan distance is used to calculate the distance between two vectors that has real values. There are two vectors on a uniform grid like a chess board. If the vectors move only in straight paths and cannot move diagonally, then the Manhattan distance determines the distance traveled between two vectors (Farhai and Jamzadeh, 2017).

3.2.3. Cosine distance similarity

Cosine similarity measure by Collins (Collins, Okada, 2012). et al. has been used, it is a measure of similarity between two vectors of an internal multiplicative space, which is based on the cosine of the angle between them. The cosine of zero theta is equal to one and the rest of the angles are smaller than one. Therefore, the measure is the direction of two vectors, not their magnitude (Farhai and Jamzadeh, 2017, 21). The cosine similarity criterion is very suitable for evaluation, especially in sparse vectors, and it is used especially in the positive space that has a range of [0.1] (Farhai and Jamzadeh, 2017, 21). For this reason, this method has been used in evaluating the similarity of architectural plan images due to the variety of data.

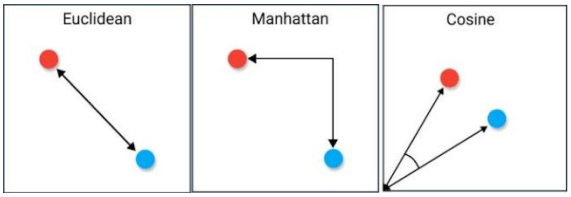

Fig. 1. Similarity Measurment Methods
(Source: Twarddatascience.com)

3.3. Classification

Classification is a process wich divided database into clusters whose members are similar to each other. The purpose of clustering is natural grouping of patterns, objects and points (Serafrazi, 2017, 67). This method is used in scientific researches in most scientific fields(Serafrazi, 2017, 79). There are different methods for classifying data. These methods are generally divided into two main categories: Hierarchical classification and Non-hierarchical classification (Adams, 2019, 1). Hierarchical classification allows the researcher to discover patterns in the database (Xiao, 2021, 188).

3.3.1. Hierarchical clustering

Hierarchical clustering method is used as the most common used cluster analysis method in small database. Hierarchical clustering method first deals with hierarchical analysis of data by considering some criteria. This algorithm itself is divided into two groups .The Agglomerative and Division clustering (Nabi Lu, Daneshpour, 1394, 17). The most conventional clustering methods are linkage method, Single linkage method, Average linkage and Complete linkage. The single linkage method is based on the minimum distance between two points in two different clusters, the average linkage is based on average distances and in the complete linkage method, the farthest distance is calculated for clustering(Jarman, et al, 2020, 7). The single linkage and complete linkage methods are strongly affected by nois data, so among the classification methods, the average

linkage method can provide a more appropriate clusters (Jarman, et al. et al., 2020).

4. Research Methodology

4.1. Scope of research

Bandar Kong is one of the coastal cities of the Persian Gulf. vernacular houses of Bandar Kong (Figure1) have windcatchers often in west and east part of the house. Windcatchers are the main characteristic of Bandar Kong architecture, which are used to benefit the sea breeze. Windcatchers of Bandar Kong are square or rectangular in shape, which are divided into four smaller X-shaped channels. shape of windcatchers are short and wide because they increase the flow of incoming air to reduce the air humidity of the living space.

Fig. 2. The historical context of Bandar-Kong

Fig. 3. Younesi Historical house Plan

Windcatchers of Iran have variety of plans and depending on climate conditions they are divided into several parts [11]. Bandar-Kong is one of the coastal cities of the Persian Gulf, that historical context of it has been preserved. vernacular houses in Bandar Kong generally have courtyard and windcatchers. Most of houses have two windcatcher on east and west part of them. Windcatchers of Bandar-Kong are shorter and wider in comparison with other windcatchers in Iran. They are wider because they increase the flow rate of the incoming air to reduce the air humidity of the living space. Due to the high humidity of the air, the function of these windcatchers is generally cooling by reducing the humidity. The shape of the blades are all X-shaped, and the main reason for those thinness is the poor architecture in the port of Lange.(Zarandi, 2015). Due to the humidity of the air, the function of these windcatchers is generally cooling by reducing the humidity (Zarandi, 2015).

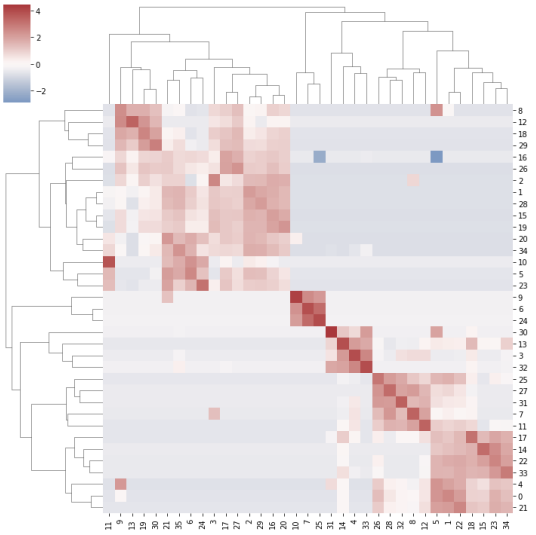

Fig. 4. Clustmap created by anaconda

In this research, 35 windcatchers of vernacular houses in Bandar Kong were selected for clustering. The samples have been classified in two general ways. The first classification method is based on shape similarity, and the second classification method is based on extracted features. In the second method,three method are used for clustering average linkage, single linkage and complete linkage.

1.first method: classification based on shape similarity:

In this method, first, the plan of 35 windcatchers are saved as 400*400 pixel images , these images are converted into a 400*400 matrix in the Anaconda 3.9 based on RGB codes in images. each matrixes have been compared to other matrixes (wich are extracted from images) 34 times. The percentage of similarity is caculated with cosine distance method and entered in a 35*35 matrix. this matrix is clustered using clustmap from skippy library. Here is the prompt wich used for similarity and clustering plans.

```
similarity = -1 * (spatial.distance.cosine(plan_array1, plan_array2) - 1)
sns.clustermap(building, z_score=0, cmap="vlag", center=0)
```

2. second method: clustering based on characteristics:

In this method windcatcher features such as length, width, location and direction are extracted and entered in an excel file. These characteristics are clustered using unsupervised machine learning methods in Anaconda3.9 by Sklearn.cluster library. The plans are clustered in 3 different modes: single linkage, average linkage and complete linkage. Fig shows three clusters dedrogram. Here in the prompt to cluster the plans wich is used in thie research:

```
from scipy.cluster.hierarchy import linkage, dendrogram, fcluster
hierachical = linkage(kong, metric='euclidean',method='average')
```

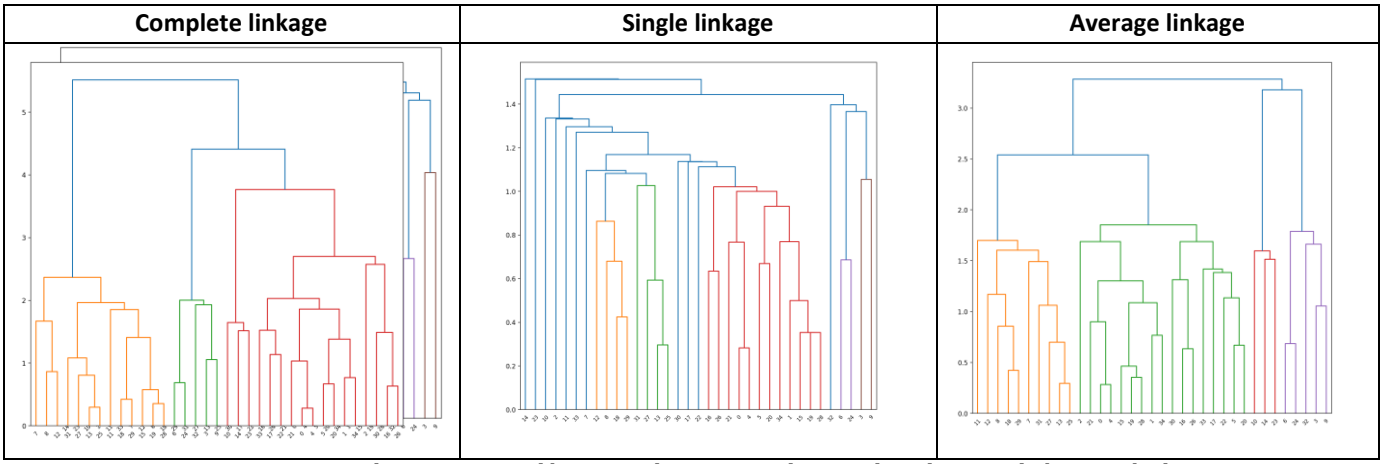

Fig. 5. Dendograms extracted by anaconda using complete, single and average linkage methods

5. Results and Discussion

The diagrams extracted from Anaconda called dendrogram. The more diffrences between a cluster objects, the less clusters.

By cutting dendrogram of complete linkage, the plans are divided in 4 clusters. In average linkage method the number of clusters are 5 clusters and in single linkage method there is no specific cluster and most of plans belong in a group containing one plan. In first method that clusters are based on similarity, there is 6 clusters.

5.1. Clustering based on image similarity

In this method there is 6 clusters . first cluster cotain four east windcatchers with smaller width and length. In this cluster windcatcher is located in north part of the room. Second cluster contain 12 east windcatchers wich are bigger than first cluster and the windcatcher are mostly located in north part of the room. third cluster contains 3 north windcatchers that are located in east part of the room. Forth cluster are south wind catchers. Fifth cluster are 5 west windcatchers with smaller lengh and width and sixth cluster belongs to west windcatchers with bigger size than fifth cluster and are located mostly in south part of the room.

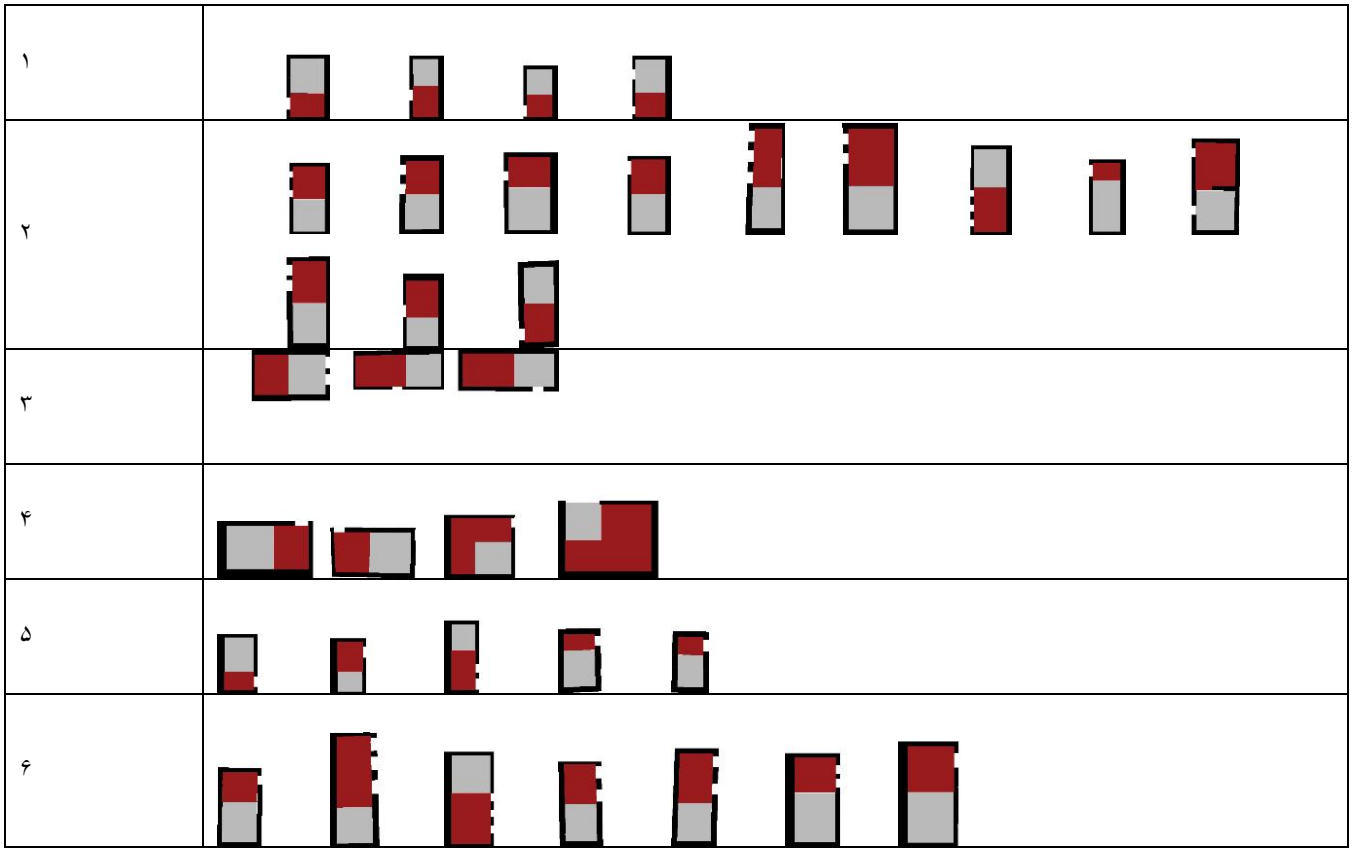

Fig. 6. Clustering based on cosine similarity method

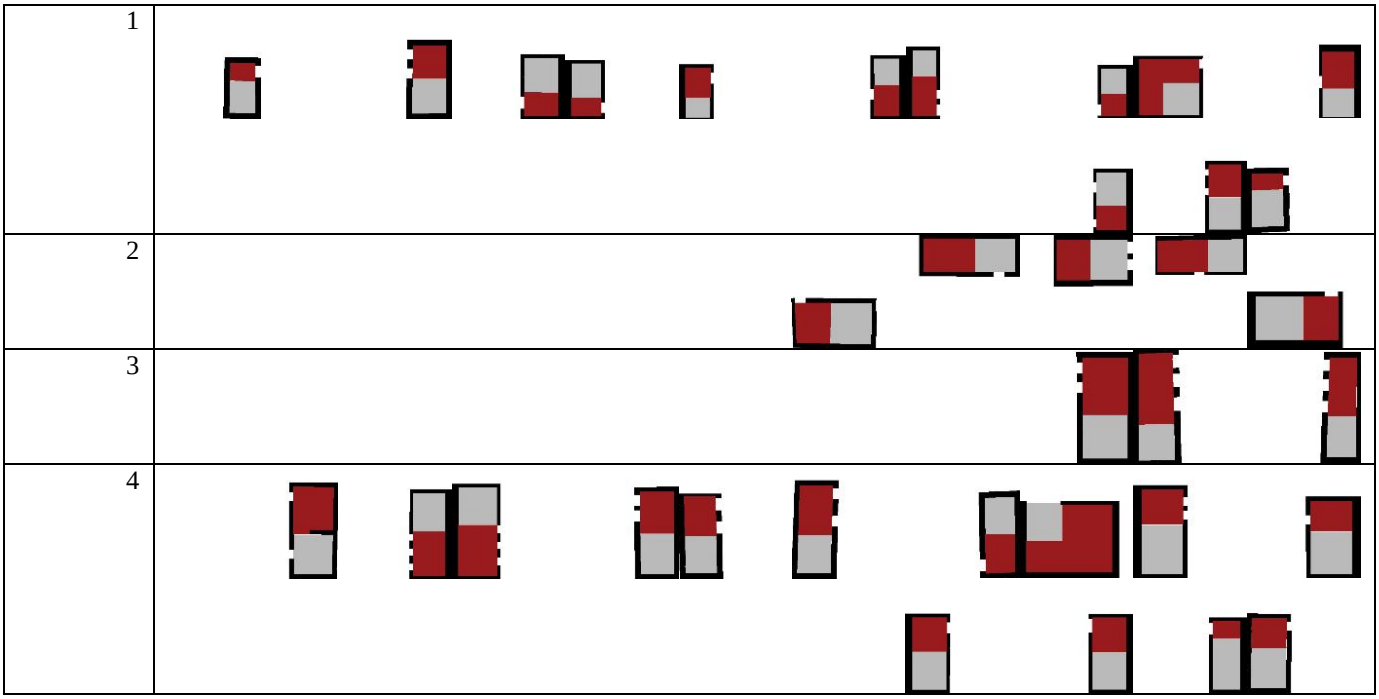

Fig. 7. Clustering based on average linkage method

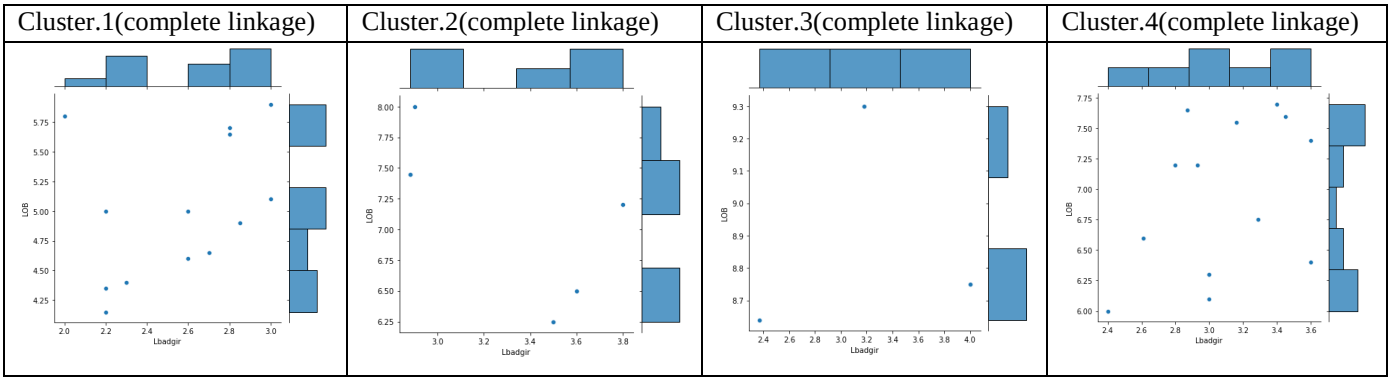

Fig. 8. Scatter plot forwidth and lenght of windcachers of clusters based on complete linkage method

5.2. Clustering based on average linkage method.

In this method 5 clusters are defined based on characteristics. These clusters are according to characteristics, specially width and length of windcatcher room, regardless to direction and location of wincatchers. By using Seaborn library, scatter plot of clusters are drown and The range of each characteristics is specified. According to the scatter plots In first cluster windcatchers are in east and west part of the house length of windcatcher is between 2-3 and width is between 4-5. In second cluster the length is between 2.6- 3.6 and the width is between 6.2-7.7 and windcatchers are in east and west part of the house. In third cluster, length is between 2.4-4 and the width is between 8.6-9.3 and wincatchers are in east and west part of the house. This cluster is the same as cluster3 in complete linkage method. Windcatchers in this cluster are the biggest one. In forth cluster the length is between 2.8- 3.8 and width is between 6.2-7.5 and widcatchers are located in east and west part of the house. In fifth cluster the length is between 2.4 – 3.2 and the width is between 5.6- 6.7 and windcatchers are located in east and west.

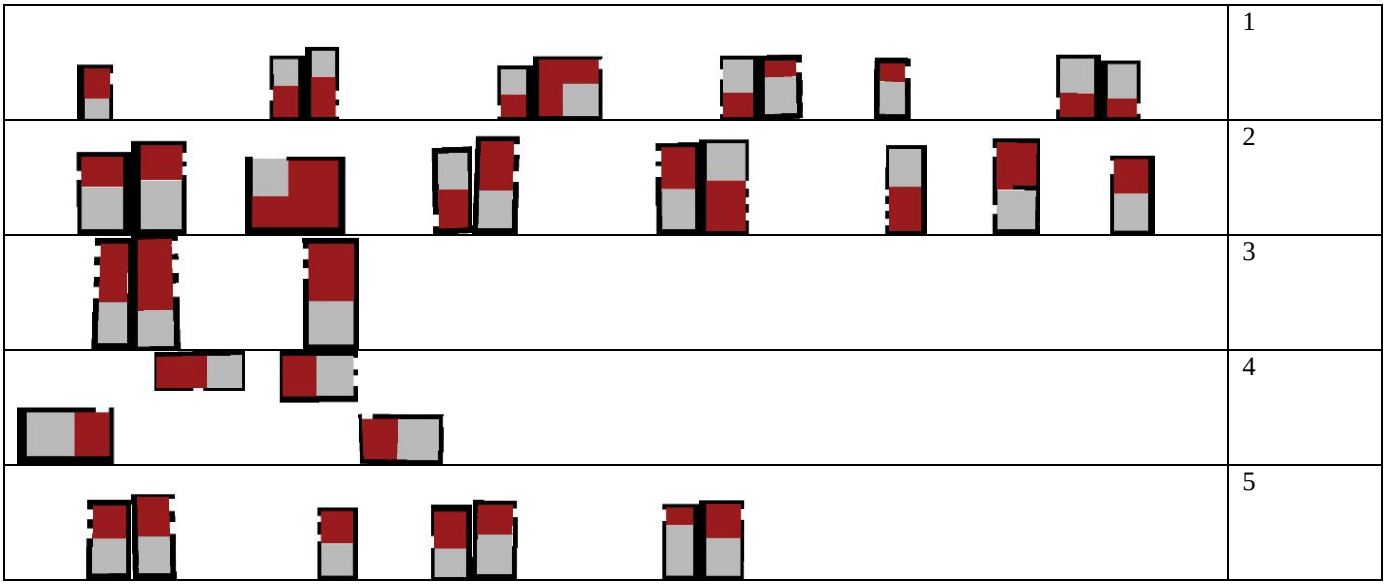

Fig. 9. clustering based on average linkage method

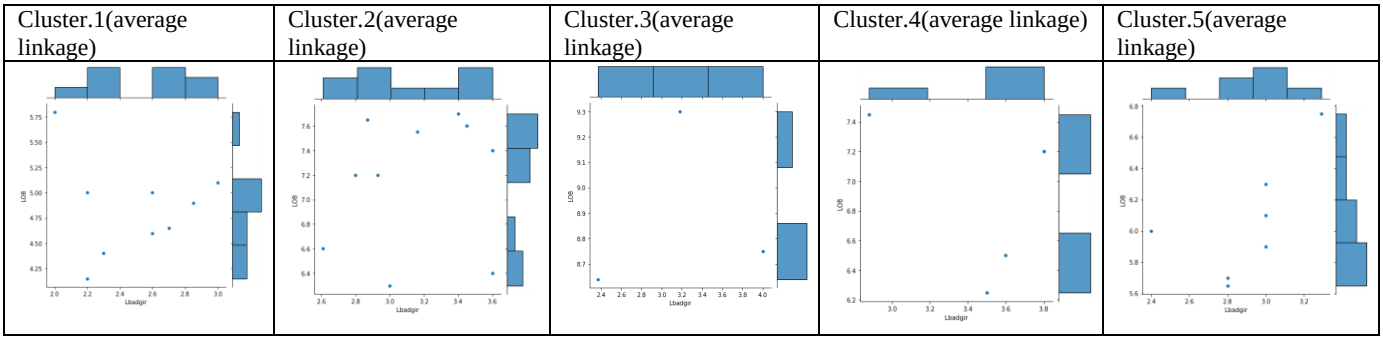

Fig. 10. Scatter Plot for Width and Lenght of windcatchers based on average linkage Method

6. Conclusion

Based on the findings of the research, the classification using the plan similarity percentage method based on the cosine distance criterion has put the plans that are more similar in terms of shape into one category. Therefore, it can be concluded that if the goal is architectural plan typology, the use of Ksionsi similarity measurement method is a more suitable method.

Unsupervised classification is a suitable method for extracting patterns and information in plans. In the single link classification method, the number of categories is very large and no conclusions can be drawn from this category. In the classification by complete link method, the classification is more general. The average link method is a more appropriate method in hierarchical classification because the average distance between data is the basis of calculations.

Based on results the plans of Bandar-e-kong can be classified in 6 clusters by cosine smimilarity method and their characteristecs are as follow :

1. In First cluster windcatchers are located at East part of the plan, length to width ratio is 1to 2-2.25 and windcatchers have square proportion. The windcatcher is located at north part of the room. The abundance of this cluster is 11.5 percent.
2. In second cluster windcatchers are located at east part of the plan, length to width ratio is 1to 2.2-2.5 and windcatcher rooms have rectangular proportion. The windcatcher is located at south part of the room. The abundance of this cluster is 34 percent.
3. I this cluster windcatchers are located in north part of the housee length to width ratio is 1 to 1to 2-2.5 and the abundance of this cluster is 11 percent.
4. In this cluster windcatchers are located at the south part of the house. Windcatcher have square proportion and the abundance of this cluster is 11 percent.
5. In fiftht cluster windcatchers are located at west part of the house, length to width ratio is 1to 2-2.5 and windcatcher rooms have square proportion. The windcatcher is located at south part of the room. The abundance of this cluster is 14 percent.
6. In this cluster windcatchers are located at west part of the plan, length to width ratio is 1to 2.2-2.5 and windcatcher rooms have rectangular proportion. The windcatcher is located at south part of the room. The abundance of this cluster is 20 percent.

References

Araldi, A., Emsellem, D., Fusco, G., Tettamanzi, A., & Overal, D. (2021). Exploring building typologies through fast iterative Bayesian clustering. In *SAGEO'2021* (pp. 113-124).

Arkin, E. M., Chew, L. P., Huttenlocher, D. P., Kedem, K., & Mitchell, J. S. (1989). *An efficiently computable metric for comparing polygonal shapes*. Cornell University Operations Research and Industrial Engineering.

Babakhani, R., & Keifari, A. (2021). Explaining the evolution of Iranian traditional house spaces based on distance measurement method of plan similarity vector. *Space Ontology International Journal*, *10*(4), 97-103.

Carter, D. J., & Whitehead, B. (1975). The use of cluster analysis in multi-storey layout planning. *Building Science*, *10*(4), 287-296.

Chang, C. C., Hwang, S. M., & Buehrer, D. J. (1991). A shape recognition scheme based on relative distances of feature points from the centroid. *Pattern recognition*, *24*(11), 1053-1063.

Collins, J., & Okada, K. (2012, September). A Comparative Study of Similarity Measures for Content-Based Medical Image Retrieval. In *CLEF (Online Working Notes/Labs/Workshop)* (pp. 2-7).

Ezavan, Amin, Babakhani, Reza (1400). Analysis of the structure of the opening form of the bridges of the Safavid period. Environmental science and technology (105)

Farhadi, Marzieh, Jamzadeh, Mansour (2017) Examining similarity criteria in content-based image retrieval. Computer Science, (21), 13-27.

Feist, S., Sanhudo, L., Esteves, V., Pires, M., & Costa, A. A. (2022). Semi-supervised clustering for architectural modularisation. *Buildings*, *12*(3), 303.

Grosswendt, A., & Roeglin, H. (2017). Improved analysis of complete-linkage clustering. *Algorithmica*, *78*, 1131-1150.

Hu, R., Huang, Z., Tang, Y., Van Kaick, O., Zhang, H., & Huang, H. (2020). Graph2plan: Learning floorplan generation from layout graphs. *ACM Transactions on Graphics (TOG)*, *39*(4), 118-1.

Jarman, A. M. (2020). Hierarchical cluster analysis: Comparison of single linkage, complete linkage, average linkage and centroid linkage method. *Georgia Southern University*, *29*.

Mousavi, M., & Afzalian, K. (2019). Morphological Analysis of Modern Residential Architecture in Turkey and Iran (Case Study: Chankaya Palace and Sa'ad Abad Palace). *Journal of Iranian Architecture & Urbanism (JIAU)*, *10*(1), 113-126.

Nabi Lo, Maryam, Daneshpour, Negin, 2014, a new hybrid clustering algorithm in category data approach

RahmatNia, A., & Hayati, H. The role of traditional medicine and human physiology in Iranian bath architecture: A case study of Kahyar Dehdasht bath, Ali Gholi Agha public bath in Isfahan, Vakil bath in Shiraz, and Bokan bath in Behbahan.

Rodrigues, E., Sousa-Rodrigues, D., de Sampayo, M. T., Gaspar, A. R., Gomes, Á., & Antunes, C. H. (2017). Clustering of architectural floor plans: A comparison of shape representations. *Automation in Construction*, *80*, 48-65.

Ryan P.Adams, 2019, Hierarchical clusterin, Cos 324-Element of Machine Learning, Prinston University.

Sajjanhar, A., & Lu, G. (1997). A grid-based shape indexing and retrieval method. *Australian Computer Journal*, *29*(4), 131-140.

Saxena, A., Prasad, M., Gupta, A., Bharill, N., Patel, O. P., Tiwari, A., ... & Lin, C. T. (2017). A review of clustering techniques and developments. *Neurocomputing*, *267*, 664-681.

Seawright, J., & Gerring, J. (2008). Case selection techniques in case study research: A menu of qualitative and quantitative options. *Political research quarterly*, *61*(2), 294-308.

Serafrazi, Abbas, 2017, half a century after clustering: review and evaluation of clustering approaches and methods with multi-criteria decision analysis, Research Quarterly in Engineering Sciences and Technology, Volume 4, Number 2, Taistan 2017, 65-84

Shah Hosseini, Habib, Khandani, Nadia, Korepaz, Rana (1401) Systematic evaluation of architectural articles in Iranian scientific-research journals and Elsevier publications (q1). Bagh Nazar (115), 5-20

Shah Mohammadi, Nima, Babakhani, Reza (1400). Analysis of the writing pattern of Nasir al-Molk mosque sashes based on the effects of environmental psychology on the psychological perceptions of users. Environmental Science and Technology, Journal 23(8): 37-46

Shih, C. Y., & Peng, C. H. (2022). Floor Plan Exploration Framework Based on Similarity Distances. *arXiv preprint arXiv:2211.07331*.

Son, K., & Hyun, K. H. (2021). A framework for multivariate data based floor plan retrieval and generation. In *26th International Conference of the Association for Computer-Aided Architectural Design Research in Asia: Projections. The Association for Computer-Aided Architectural Design Research in Asia, Hong Kong* (Vol. 29, No. 03, pp. 2021-01).

Son, K., & Hyun, K. H. (2022). Designer-centric spatial design support. *Automation in Construction*, *137*, 104195.

Strauss, Anselm, Corbin, Juliet, (2017) Fundamentals of qualitative research - techniques and stages of producing grounded theory. Translated by Ebrahim Afshar, Tehran: Nei Publishing House, 7th edition.

Vakilinejad, Roza, Mofidi, Majid, Mahdizadeh Siraj, Fatemeh (2012), The combined effect of building envelope and ventilation patterns on energy consumption in residential buildings.

Varzan, Milad (1400), Machine Learning and General Data: Basics, Concepts, Algorithms and Tools, Tehran: Miyad Andisheh.

Xiao, R. (2021). Comparing and clustering residential layouts using a novel measure of grating difference. *Nexus Network Journal*, *23*(1), 187-208.

Yousif, S., & Yan, W. (2019). Shape clustering using k-medoids in architectural form finding. In *Computer-Aided Architectural Design." Hello, Culture" 18th International Conference, CAAD Futures 2019, Daejeon, Republic of Korea, June 26–28, 2019, Selected Papers 18* (pp. 459-473). Springer Singapore.