College of Basic Education Researchers Journal. ISSN: 7452-1992 *Vol. (19), No.(1), (2023)*

Genetic modeling of olive fruit using clustering technique

Hala Nafie Fathee
Collge of Physical Education and sprot sciences

Nima Abdulla AL-Fakhry
Faculty of Administration and Economics

(Received in 5/12/2022 Accepted in 22/1/2023)

Abstract:

Data mining is the process of analyzing a quantity of data and finding relationships between them. It is summarized to obtain recognizable and useful graphic models for its users through the use of a set of automated tools to extract knowledge from its potential without making prior assumptions. The study is a step towards clarifying the principle of the fuzzy assembly algorithm on the practical and theoretical levels. The theoretical section of the research dealt with the concept of data collection and its different types, in addition to an explanation of the FCM method. While the practical part, deal with the selection of the adjacent FCM method. And the olive fruits were chosen in the practical aspect because of the nutritional, economic and commercial characteristics of this fruit, depending on some of the characteristics available in it, as those characteristics belong to more than one variety at the same time. The work of the algorithm is to detect ambiguities between the varieties under study and then distinguish between them. As well as knowing the similarities between them to determine the extent of similarity between the types of olives. Its purpose is to expand the geographical area of olive cultivation and to find new hybrid varieties that have high-quality features. The researchers used the Python language to implement the practical side, and the algorithm proved highly efficient in determining the genetic characteristics of olive fruits in the research sample, The most important finding of the research is that the FCM algorithm was very flexible in dealing with different types of data and incomplete data processing, as well as its smoothness in dealing with different systems and programs.

Key words: Cluster, fuzzy cluster, FCM, olive fruit.

نمذجة الجينات الوراثية لثمار الزيتون باستخدام تقنية العنقدة

هالة نافع فتحي
كلية التربية البدنية وعلوم الرياضة

نعمة عبدالله الفخري
كلية الإدارة والاقتصاد

الملخص:

التنقيب عن البيانات هو عملية تحليل كمية من البيانات وإيجاد العلاقات بينها. تلخص للحصول على نماذج رسومية يمكن التعرف عليها والاستفادة منها عبر استخدام مجموعة من الأدوات لاستخراج المعرفة من مكامنها دون وضع افتراضات مسبقة. يمثل البحث خطوة نحو توضيح مبدأ خوارزمية التجميع الغامض على المستويين العملي والنظري. تناول القسم النظري من البحث مفهوم جمع البيانات وأنواعها المختلفة ، بالإضافة إلى شرح لالية عمل خوارزمية FCM. بينما تناول الجزء العملي ، تطبيق خوارزمية FCM المجاورة. وتم اختيار ثمار الزيتون من الناحية العملية لما تتمتع به هذه الفاكهة من خصائص غذائية واقتصادية وتجارية ، اعتمادًا على بعض الخصائص المتوفرة فيها ، حيث إن تلك الخصائص تنتمي إلى أكثر من صنف في نفس الوقت. يتمثل عمل الخوارزمية في الكشف عن الالتباسات بين الأصناف قيد الدراسة ثم التمييز بينها. وكذلك معرفة أوجه التشابه بين أنواع الزيتون، الغرض منه هو توسيع المساحة الجغرافية لزراعة الزيتون وإيجاد أصناف هجينة جديدة ذات ميزات عالية الجودة. استخدم الباحثون لغة Python لتنفيذ الجانب العملي ، وأثبتت الخوارزمية كفاءة عالية في تحديد الخصائص الجينية لثمار الزيتون في عينة البحث ، وأهم ما توصل إليه البحث هو أن خوارزمية FCM كانت مرنة للغاية في التعامل مع مختلف أنواع البيانات ومعالجة البيانات غير المكتملة ، فضلاً عن سلاسة التعامل مع الأنظمة والبرامج المختلفة.

1-Introduction:

Data mining is one of the branches of artificial intelligence. Its goal is to create graphic models, and such designs are methods, procedures, and so on.

The goal of these models is to connect the input data in order to gain new information based on their shared attributes because of which they can be used to interpret the results and make effective decisions.

Clustering is a data mining application. It is a process of statistical analysis of a number of variables (data) arranged in clusters (vectors). Similarities, differences and relationships between them are studied to obtain different models. In addition, benefit from it in various fields of life, such as the medical, industrial, and agricultural fields and others..

The importance of the research is evident in its theoretical side, as it is a small contribution to clarifying the concepts therefore as minor participation to clarifying the concepts of cluster analysis and fuzzy clustering, as well as an illustration of FCM algorithm model and its phases of process.

While the practical side, the first importance point in this paper is to classify olive fruits based on the common properties between them and to remove the ambiguity surrounding the classification of the olive fruits as they belong to more than one class at the same time. The second importance is the practical side; it comes from applying the Normalization mechanism to the results that obtained from the algorithm in order to obtain a stronger model.

2- Research Methodology:

A- Research aims:

The aims of the research are as follows:

1- Elaboration of a theoretical framework for the concept of data clustering and its types.
2- Highlight the concept of fuzzy clustering.
3- Determine what the FCM algorithm is.

B- Research problem.

With the wide spread of data, and the development of data and information technologies, it has become necessary to find the best and appropriate algorithm for certain groups of data, accordingly, development processes emerged and implementation of algorithms data cluster and on different ranges.

C-Research hypothesis

FCM technique is recognized data as the most accurate in categorizing fuzzy set because that has more than one interpretation.

D-The study sample:

Some properties of olive fruits from different regions were used to implement the FCM algorithm.

First - the theoretical side:

1- The concept of clustering:

It refers to the method of breaking up the initial data into groups, gathering data with related features, and clustering them according to measure how far off they are from one another.
The objective of this assembly process is to develop descriptive models that will make it easier to examine them and generate significant and beneficial outcomes for decision-makers.
The methodology of categorizing data into classes depending on with their involvement in similar features.

This is an undirected division of data, unlike classification, which is widely used in solving many problems. By analyzing a set of data and placing it in the form of categories or sections that can be used later to classify future data using the methods that are used to classify the data. A cluster is a set of data that is similar to each other and is not similar to members belonging to other clusters (Kassambara.2017, PP. 25-28).

In order for the data collection process to be successful, a preprocessing of the data must be carried out, according to the following steps (Zupar and Demsar, 2018, PP.45-47) & (Li et al, 2022, 1573-1580):

A- Choose a feature:

Choosing the attributes (properties), that enables us to analyze the codes of more information related to the task to be solved.

B- Choosing the clustering algorithm:

Choosing the appropriate algorithm, which results from defining a good cluster schema for the data set.

C- The reliability of the research results:

The right approaches that standards used to verify the clustering application's outcome because algorithms, define previously unseen clusters, regardless of manner. As a result, the final segmentation of the data necessitates some pattern of evaluation.

D- Explanation of the study results:

To get to the appropriate conclusions, the expertise in the application field must combine the cluster findings with the remainder of the tested studies.

2- Types of clustering:

The identification of clustering is based on a criterion as well as scale that specifies the similarity among data items entered into the approach. In addition to the theoretical underpinnings and fundamental ideas that support the cluster, analysis method depends. For instance, statistic and fuzzy approach. The development of clustering algorithms is based on related fields like data mining, machine learning, pattern classification, etc. The following categories are used to group these algorithms. (Han & etal, 2016, PP. 15-23) & (G. Mohammed and etal, 2016, P.88):

1. Hierarchy Clustering.
2. Partitioned Clustering.
3. High Dimensional Data Clustering.
4. Using Genetic Properties together Cluster.

Clustering techniques have already been classified by some researchers and use the suggested technique for defining clusters into the following categories (Pande & Thakre,2012):

- Density Based Clustering.
- Grid Based Clustering.

Each of these aforementioned categories seems to have its own distinct and sub-types algorithms for finding the clusters. Accordingly, clustering methods are classed

based on their kind. Of variables present in the data as follows (Jyuti & Kaushok, 2014, PP 183-184):

- Statistic Clustering: This kind employs criteria of similarities to segment patterns that are constrained by numerical data and is based on statistical analysis ideas.
- Clustering concept is a technique for clustering in which modeling is clustered based on the ideas they support but not on their values.
- Fuzzy clustering: fuzzy approached is utilizing to cluster data, which rely on the idea that models may be classified into multiple clusters. Real data that has not been validated is utilized. Crisp Clustering: This type assumes a not overlapping segmentation, this means that the point the data belongs to the cluster or not.

Kohonen Net Clustering: The cluster here depends on the concept of a cluster and a neural network is formed from the external nodes.

3- The concept of fuzzy clustering:

Clustering is a technique for collecting to investigate their patterns using data common characteristics. Fuzzy clustering technique permits components to connect to several classes (types) simultaneously, with different degrees of membership. In addition, in multiple instances groups of fuzzy clusters are more genuine than static groups. Therefor, objects at a boundaries of classes do not have to belong completely to one of the classes (types), rather, a degree of membership valued. Each item is given a number between 0 and 1 to represent their partial membership.(Babuska, 2003:1, 2005: 3) و (Klawonn, and Hoppner). When seeking to discover graphic frameworks or methods for datasets that are overlapping and clustered with each other, the following must be taken into consideration)Francesco,2015: 1):

1- Data representation to transform states of data sets into properties vectors..
2- Using measures of similarities in approaching dataset instances.
3- Hierarchy clustering methods serve as the essential for structuring clustering algorithms, which break down the data structure into smaller parts down to a single part.

4- FCM methodology:

This iterative method creates groups of the dataset using the idea of fuzzy membership. Instead of placing each component in a separate category, within each item there is a separate

membership value for each group. Thus, the items in the center of the set have higher organic values than the items on the edges sets. This technique attempts to identify groups inside the dataset by gaining the least value of the "+".

1- The scheme of the practical side:

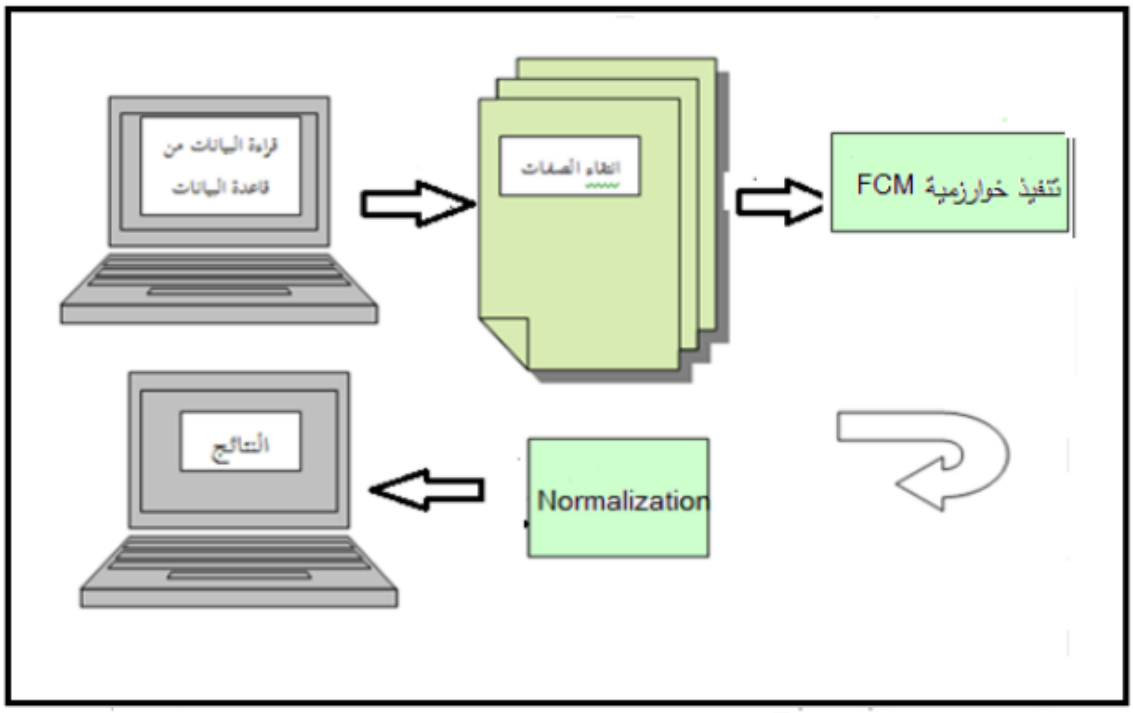

Figure 1 Scheme of the stages of the practical side.

2- Study sample:

Before starting to display, the results obtained from the application of the Fuzzy C-mean algorithm. It is necessary to describe the data that is used in the research. 53 items have been selected, of olive varieties to implement the form for what it carries characteristics and features, for example (the highest percentage of oil in the olive fruit, the lowest percentage of oil in the olive fruit, w eight of the olive fruit, color,... and others), table No. (1) Shows the data of the research sample.

Notes about the data:

When plotting the data as points, it will only show 42 points out of 53 because there is similarity between some of them that makes them stacked, one on top of the other.
Similar data are:
Number (4,6,12) the attributes are {18,14}
Number (10,5) the attributes are {22,16}
Number (45,15) the attributes are {20,18}

.

Table (1)
Varieties of olives and the proportions of oil for each olive fruit

Seq.	Highest oil content in the olive fruit	Lowest oil content in the olive fruit	olive variety
1	**2**	1	Diql
2	5	4	Al-Ajizi Al-Shami
3	6	3	Al-Ajizi Al-Aqsi
4	18	14	**Ashrsi**
5	21	16	Khastawi
6	18	14	Altafahi
7	9	7	Frantoio
8	14	12	Alja
9	13	3	Asklano
10	21	16	Bashiqa (regular)
11	18	15	Dulce

12	18	14	Svalano
13	18	16	Baronial
14	19	16	Alhamid
15	20	18	Alwatiqain
16	20	15	Kalamata
17	20	17	Erbecoin
18	20	15	Shamilali
19	46	23	Fronto
20	20	15	Alshatwi
21	22	16	Manzanillo
22	22	15	Picual
23	22	18	Koratina
24	22	3	Mssion
25	24	16	Kronaki
26	24	19	Wasalati
27	25	20	Almaraqi
28	28	15	Khudairi
29	30	25	Nasuhi
30	30	28	Sourani (Al-Ma'arri)
31	35	25	Zayti
32	40	20	Alnabali
33	37	12	Miski

34	28	22	Zlitny
35	24	16	Koroniki
36	24	22	Kayrusawy
37	15	11	Aldaeaybali
38	22	15	Alnabali Almuhsin
39	37	18	Alsayquz
40	45	25	Alzaraziu
41	25	18	Alzalmatiu
42	12	9	Mahzam Abu Satl
43	22	20	Aljarbueiu
44	20	17	Arbiykuyna
45	20	18	Alquysi
46	23	18	Bishulyn Almaghribiy
47	25	20	Liytshiynu
48	16	12	Sant katiryn
49	26	24	Jbuji
50	24	20	Aldaan
51	16	9	Almahati
52	24	18	Laymili
53	24	18	Aleabaani

3- In this paper, application Fuzzy c-means technology technique and Use Python language for Clustering, We represented the olive fruits in the drawing in the form of points, and the cluster strength scale (PC) was used to know the number of clusters suitable for arranging the types of olives. In this scale, the result is perfect the closer we get to one, because the scale is between (0-1), while the other scale used is a scale, which is the opposite. From the above, that is, the closer we get to zero, the better the result

The results are as follows:.

if the number of clusters are 2

PC =0.361 (0 -1) if the closer to 1, the better the cluster.

PEC =0.314 (0-1) if the closer to 0, the better the cluster.

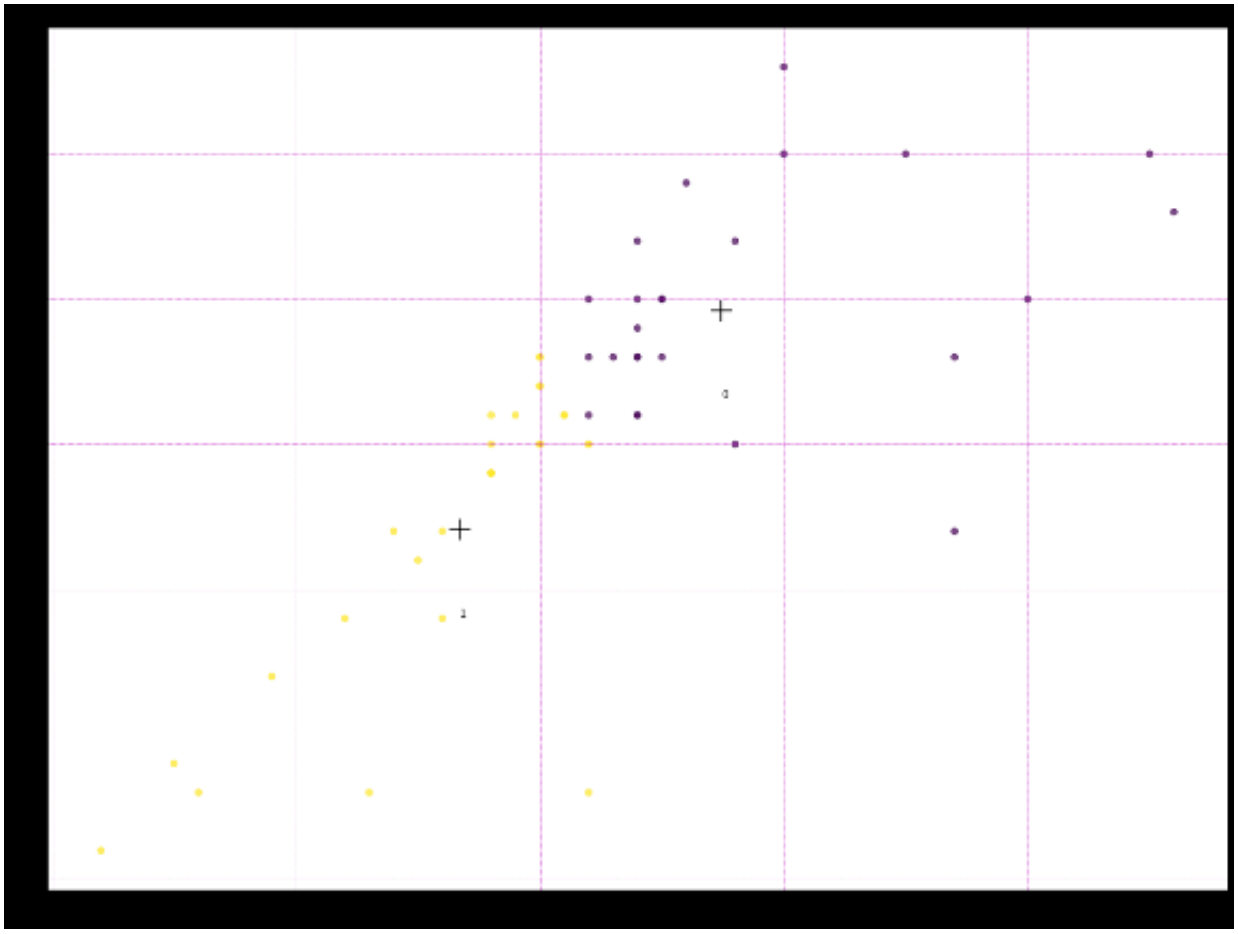

Figure (2) represents the cluster if it were two clusters

The cluster No.1 is represented the following items:

[['1', ' Diql '], ['2', ' Al-Ajizi Al-Shami'], ['3', ' Al-Ajizi Al-Aqsi '], ['4', ' Ashrsi '], ['5',Khastawi '], ['6', ' Altafahi '], ['7', ' Frantoio '], ['8', ' Alja '], ['9', ' Asklano '], ['10', ' Bashiqa '], ['11', ' Dulce '], ['12', 'Svalano '], ['13', ' Baronial '], ['14', ' Alhamid '], ['15', ' Alwatiqain '], ['16', ' Kalamata '], ['17', ' Erbecoin '], ['18', 'Shamilali '], ['20', ' Alshatwi '], ['22', 'Picual '], ['24', ' Mssion '], ['37', ' Aldaeaybali '], ['38', 'Alnabali Almuhsin '], ['42', 'Mahzam Abu Satl '], ['44', ' Arbiykuyna '], ['45', 'Alquysi '], ['48', ' Sant katiryn '], ['51', ' Almahati ']].

The cluster No.2 is represented the following items:

[['19', ' Fronto '], ['21', ' Manzanillo '], ['23', ' Koratina '], ['25', ' Kronaki '], ['26', ' Wasalati '], ['27', ' Almaraqi '], ['28', ' Khudairi '], ['29 ', ' Nasuhi '], ['30 ', ' Sourani '], ['31 ', ' Zayti '], ['32 ', ' Alnabali '], ['33 ', ' Miski '], [' 34', ' Zlitny '], [' 35', ' Koroniki '], ['36 ', ' Kayrusawy '], ['39 ', 'Alsayquz '], [' 40', ' Alzaraziu '], ['41 ', ' Alzalmatiu '], ['43 ', ' Aljarbueiu '], [' 46', ' Bishulyn

Almaghribiy'], ['47 ', ' Liytshiynu ', [' 49', ' Jbuji '], ['50 ', ' Aldaan '], ['52 ', ' Laymili '], ['53 ', ' Aleabaani ']].

2- If the number of clusters is 3:

PC = 0.261 (0 -1) if the closer to 1, the better the cluster.

PEC = 0.195 (0-1) if the closer to 0, the better the cluster

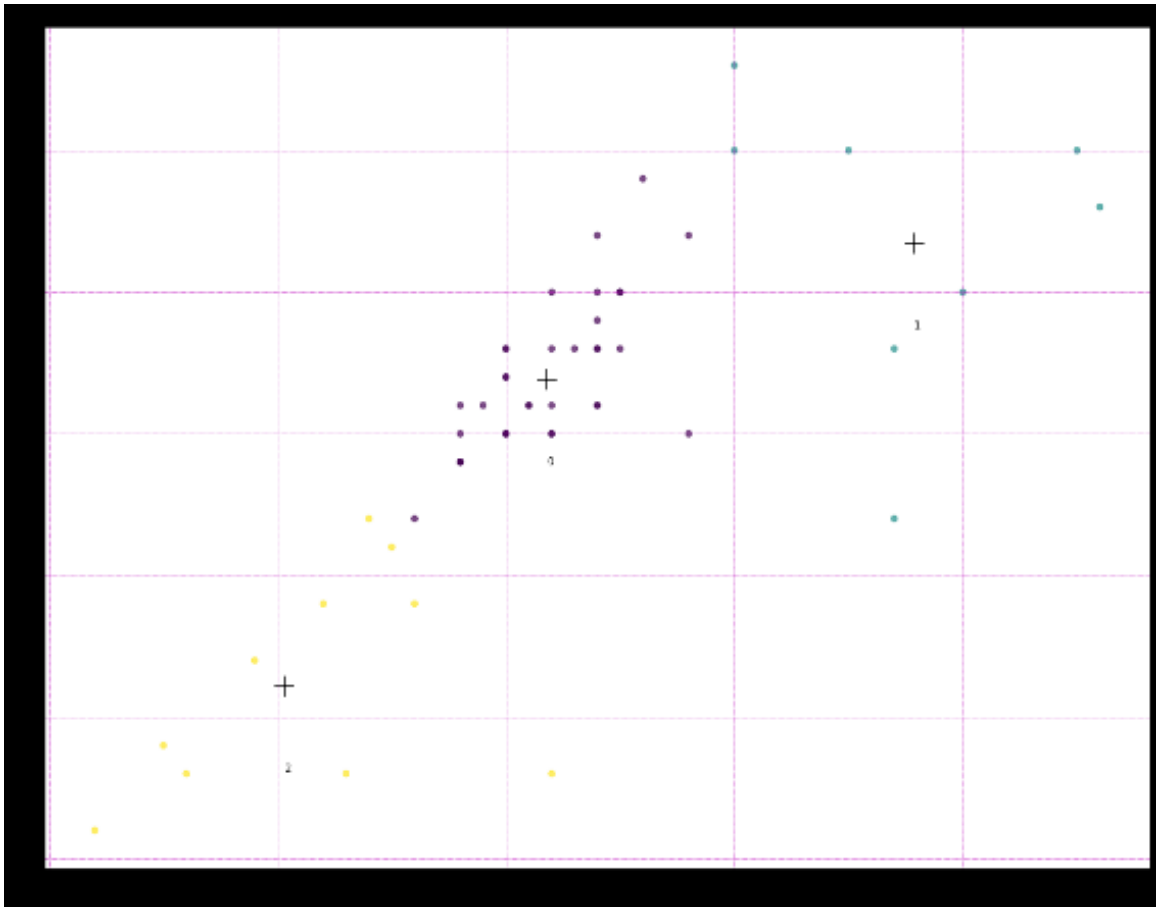

Figure (3) represents a cluster if it were three clusters

Cluster No. 1 included the following items:

[[' 1', ' Diql '], [' 2', ' Al-Ajizi Al-Shami '], ['3 ', ' Al-Ajizi Al-Aqsi '], [' 4', ' Ashrsi '], [' 5', ' Khastawi '], ['6 ', ' Altafahi '], ['7 ', ' Frantoio '], [' 8', ' Alja '], ['9 ', ' Asklano '], ['10 ', ' Bashiqa '], ['11 ', ' Dulce '], ['12 ', 'Svalano '], ['13 ', ' Baronial '], ['14 ', ' Alhamid '], ['15 ', ' Alwatiqain '], ['16 ', ' Kalamata '], ['17 ', ' Erbecoin '], ['18 ', ' Shamilali'], [' 20', ' Alshatwi '], ['22 ', ' Picual'], ['24 ', '

Mssion '], ['37 ', ' Aldaeaybali '], [' 38', 'Alnabali Almuhsin '], ['42 ', 'Mahzam Abu Satl '], ['44 ', ' Arbiykuyna '], ['45 ', 'Alquysi '], ['48 ', ' Sant katiryn '], ['51 ', ' Almahati ']].

While cluster No. 2 represented the following items:

[[' 19', 'Fronto '], [' 21', ' Manzanillo '], [' 23', ' Koratina '], [' 25', ' Kronaki '], [' 26', ' Wasalati '], [' 27', ' Almaraqi '], [' 28', ' Khudairi '], [' 29', ' Nasuhi '], [' 30', ' Sourani '], [' 31', ' Zayti '], [' 32', ' Alnabali'], [' 33', ' Miski '], [' 34', ' Zlitny '], [' 35', ' Koroniki '], [' 36', ' Kayrusawy '], [' 39', ' Alsayquz'], [' 40', 'Alzaraziu '], [' 41', ' Alzalmatiu '], [' 43', ' Aljarbueiu '], [' 46', 'Bishulyn Almaghribiy '], [' 47', ' Liytshiynu '], [' 49', ' Jbuji'], [' 50', ' Aldaan '], [' 52', ' Laymili '], [' 53', 'Aleabaan ']]

2- If the number of clusters are 3:

PC = 0.261 (0-1) if the closer to 1, the better the cluster.

PEC = 0.195 (0-1) if the closer to 0, the better the cluster.

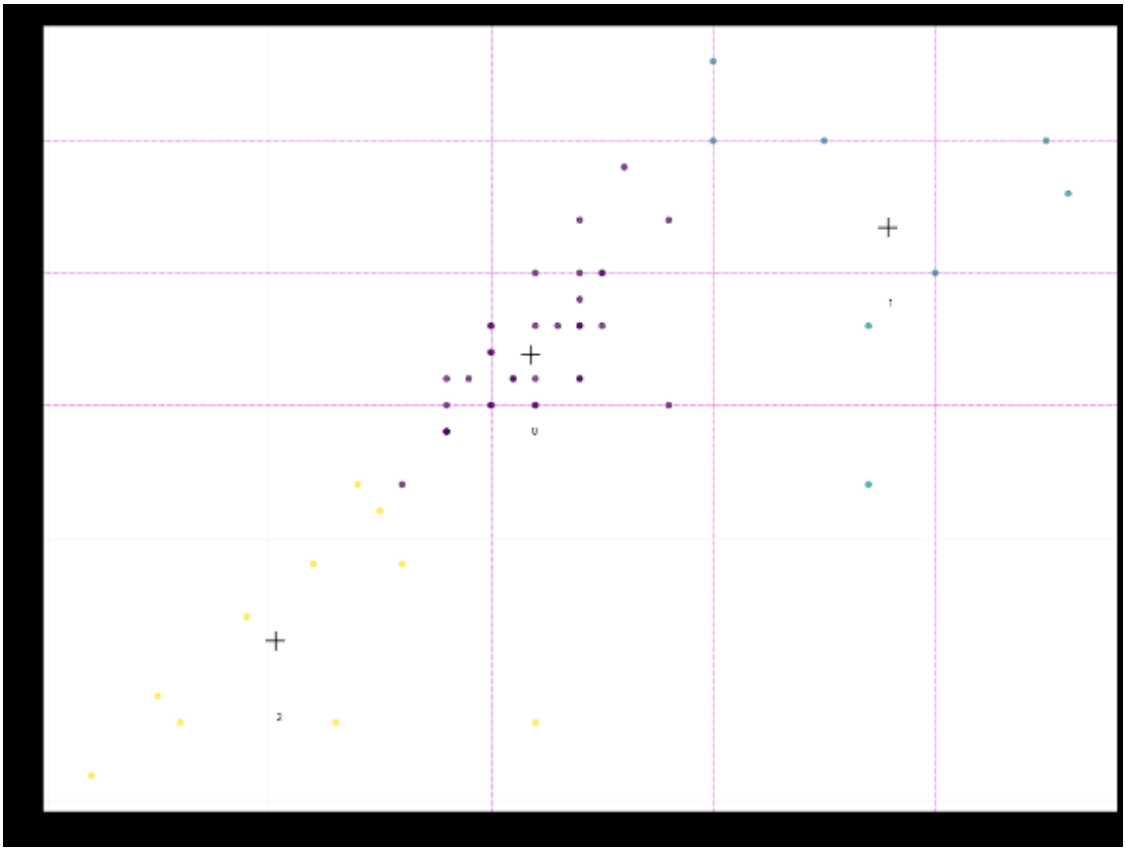

Figure (3) represents a cluster if it were three clusters

Cluster No. 1 included the following items:

[['1 ', ' **Diql** '], [' 2', ' **Al-Ajizi Al-Shami** '], ['3 ', ' **Al-Ajizi Al-Aqsi** '], [' 7', ' Frantoio '], ['8 ', ' Alja '], ['9 ', ' Asklano '], [' 24', ' Mssion '], ['37 ', ' Aldaeaybali '], [' 42', ' Mahzam Abu Satl'], [' 51', ' Almahati ']]

Cluster No. 2 included the following items:

[['4 ', ' **Ashrsi** '], ['5 ', ' Khastawi '], ['6 ', ' Altafahi '], ['10 ', ' Bashiqa '], [' 11', ' Dulce '], [' 12', 'Svalano '], ['13 ', ' Baronial '], ['14 ', ' Alhamid '], [' 15', ' Alwatiqain '], [' 16', ' Kalamata '],

[' 17', ' Erbecoin '], ['18 ', 'Shamilali '], [' 20', ' Alshatwi '], ['21 ', ' Manzanillo '], ['22 ', ' Picual'], ['23 ', ' Koratina '], ['25 ', ' Kronaki '], ['26 ', ' Wasalati '], ['27 ', ' Almaraqi '], [' 28', ' Khudairi '], ['34 ', ' Zlitny '], ['35 ', ' Koroniki '], [' 36', ' Kayrusawy '], [' 38', 'Alnabali Almuhsin '], [' 41', ' Alzalmatiu '], [' 43', ' Aljarbueiu '], ['44 ', ' Arbiykuyna '], [' 45', 'Alquysi '], ['46 ', 'Bishulyn Almaghribiy '], ['47 ', ' Liytshiynu '], ['48 ', ' Sant katiryn '], ['49 ', 'Jbuji '], ['50 ', ' Aldaan '], ['52 ', ' Laymili '], ['53 ', 'Aleabaan ']]

Cluster No. 3 included the following items:

[['19 ', ' Fronto'], [' 29', ' Nasuhi '], ['30 ', ' Sourani '], ['31 ', ' Zayti '], ['32 ', 'Alnabali '], [' 33', ' Miski '], ['39 ', 'Alsayquz '], [' 40', 'Alzaraziu '], the figure number (4) Represents the number of clusters 2,3,4,5,6,7.

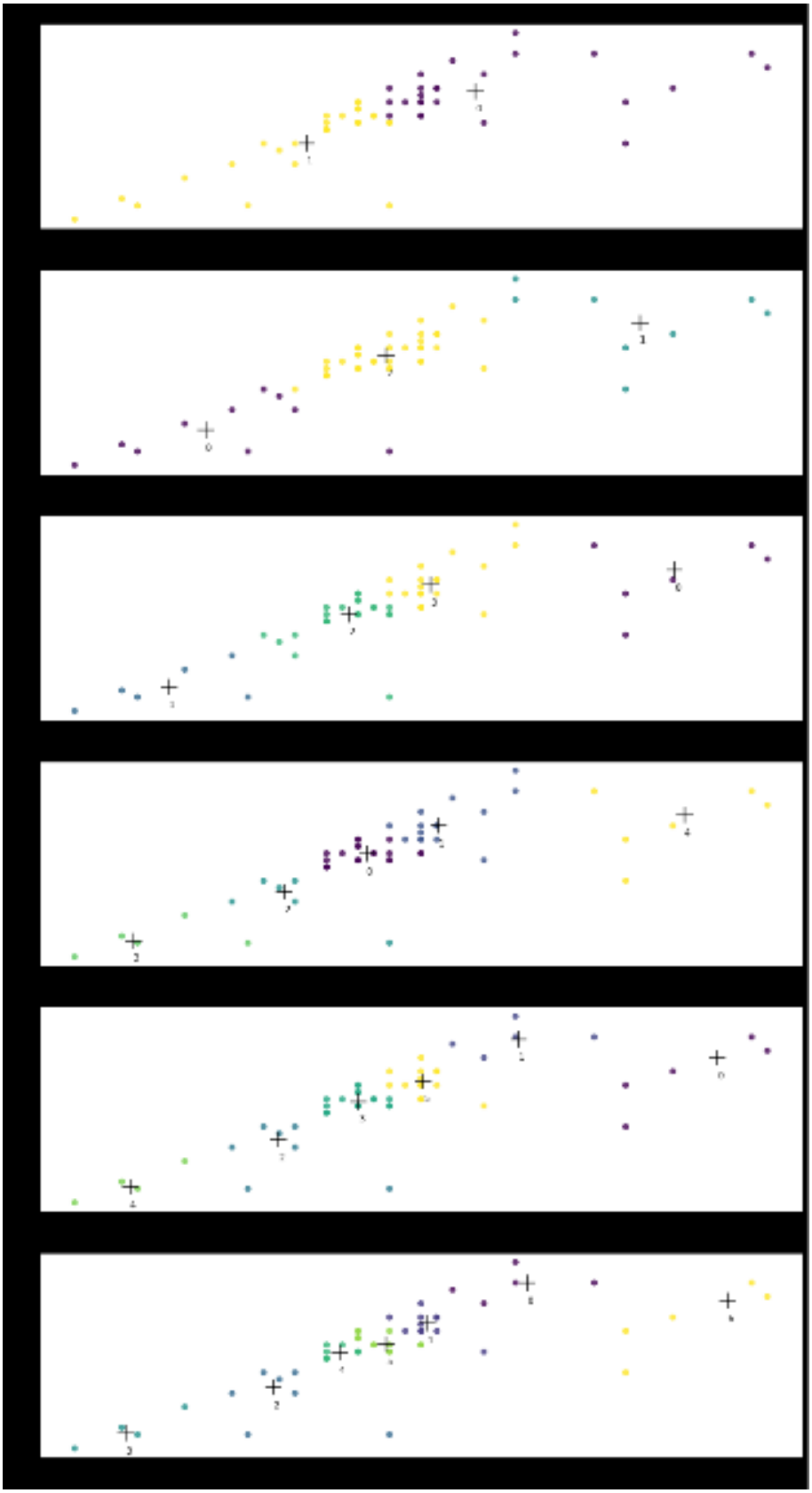

Figure (4) shows the number of clusters 2, 3, 4, 5, 6, 7.

4- To enhance the results, Executing the Normalization operation over data with dividing by the Standard Deviation, the standard deviation was calculated by using this equation:

$$s = \sqrt{\frac{(x_i - \bar{x})^2}{n-1}}$$

1- Number of clusters are 2:

PC =0.388 (0-1) if the closer to 1, the better the cluster.

PEC =(0.260) if the closer to 0, the better the cluster.

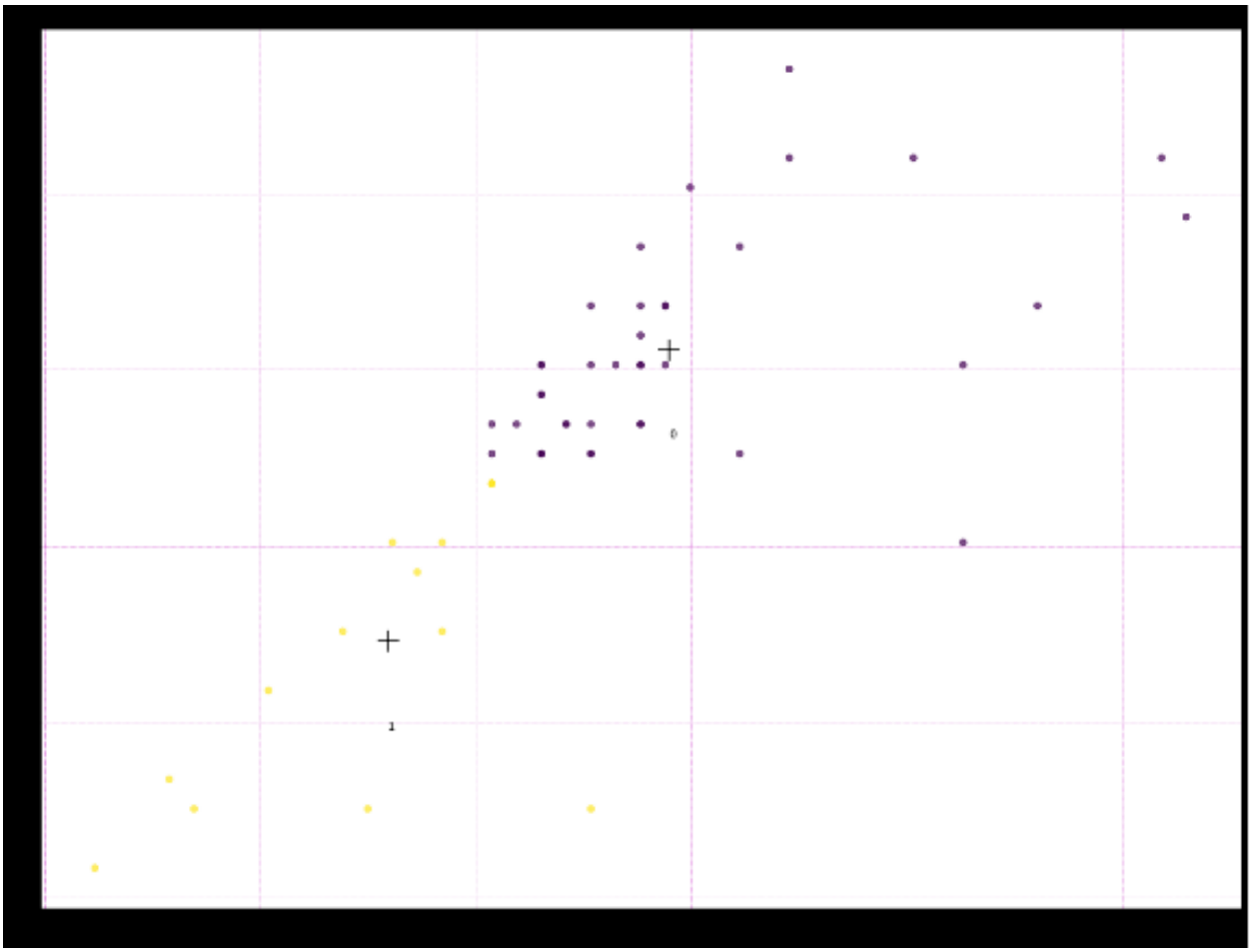

Figure (5) Number of clusters are two after Normalization

Cluster No. 1 contained the following items:

[[' 1', ' Diql '], [' 2', ' Al-Ajizi Al-Shami '], [' 3', ' Al-Ajizi Al-Aqsi '], [' 4', ' Ashrsi '], [' 6', ' Altafahi '], [' 7', 'Frantoio '], [' 8', ' Alja '], [' 9', ' Asklano '], [' 12', 'Svalano '], [' 24', ' Mssion '], [' 37', ' Aldaeaybali '], [' 42', ' Mahzam Abu Satl'], [' 48', ' Sant katiryn '], [' 51', ' Almahati']]

Cluster No. 2 contained the following items:

[[' 5', ' Khastawi '], ['10 ', ' Bashiqa '], ['11 ', ' Dulce '], ['13 ', ' Baronial '], ['14 ', ' Alhamid '], [' 15', ' Alwatiqain '], [' 16', ' Kalamata '], [' 17', ' Erbecoin '], [' 18', 'Shamilali '], [' 19', ' Fronto'], ['20 ', ' Alshatwi '], ['21 ', ' Manzanillo '], [' 22', ' Picual'], ['23 ', ' Koratina '], [' 25', ' Kronaki '], ['26 ', ' Wasalati '], [' 27', ' Almaraqi ']]

Cluster No. 3 contained the following items:

[[' 28', ' Khudairi '], [' 29', ' Nasuhi '], [' 30', ' Sourani '], [' 31', ' Zayti '], [' 32', ' Alnabali '], [' 33', ' Miski '], [' 34', ' Zlitny '], [' 35', ' Koroniki '], [' 36', ' Kayrusawy '], [' 38', ' Alnabali Almuhsin'], [' 39', 'Alsayquz '], [' 40', 'Alzaraziu '], [' 41', ' Alzalmatiu '], [' 43', ' Aljarbueiu '], [' 44', ' Arbiykuyna '], [' 45', ' Alquysi'], [' 46', 'Bishulyn Almaghribiy '], [' 47', ' Liytshiynu '], [' 49', ' Jbuji'], [' 50', ' Aldaan '],[' 52', ' Laymili '], [' 53', ' Aleabaan']]

The number of clusters are 3:

PC = 0.257 (0-1) the closer to 1, the better the cluster

PEC = 0.205 (0-1) the closer to 0, the better the cluster

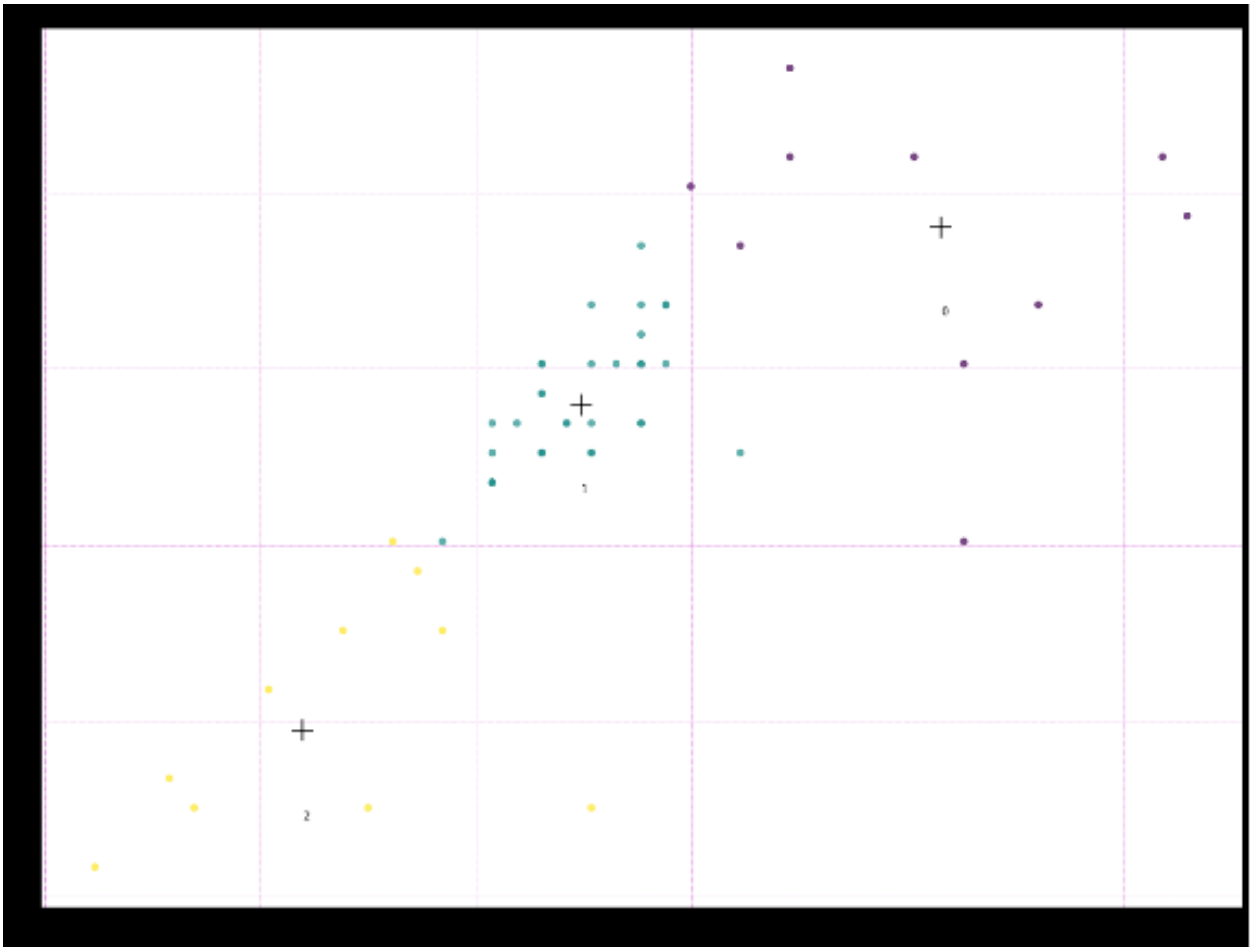

Figure (6) Number of clusters are 2 after Normalization

Cluster No.1

[[' 1', ' **Diql** '], [' 2', ' **Al-Ajizi Al-Shami** '], [' 3', ' **Al-Ajizi Al-Aqsi** '], [' 7', ' Frantoio'], [' 8', ' Alja '], [' 9', ' Asklano '], [' 24', ' Mssion '], [' 37', ' Aldaeaybali '], [' 42', ' Mahzam Abu Satl'], [' 51', ' Almahati ']].

Cluster No.2:

[[' 4', ' **Ashrsi** '], [' 5', ' Khastawi '], [' 6', ' Altafahi '], [' 10', ' Bashiqa '], [' 11', ' Dulce '], [' 12', 'Svalano '], [' 13', ' Baronial '], [' 14', ' Alhamid '], [' 15', ' Alwatiqain '], [' 16', ' Kalamata '], [' 17', ' Erbecoin '], [' 18', ' Shamilali'], [' 20', ' Alshatwi '], [' 21', ' Manzanillo '], [' 22', ' Picual'], [' 23', ' Koratina '], [' 25', ' Kronaki '], [' 26', ' Wasalati '], [' 27', ' Almaraqi '], [' 28', ' Khudairi '], [' 35', ' Koroniki '], [' 36', ' Kayrusawy '], [' 38', ' Alnabali Almuhsin '], [' 41', ' Alzalmatiu '], [' 43', ' Aljarbueiu '], [' 44', ' Arbiykuyna '], [' 45', 'Alquysi '], [' 46', 'Bishulyn Almaghribiy '], [' 47', ' Liytshiynu '], [' 48', ' Sant katiryn '], [' 50', ' Aldaan '], [' 52', ' Laymili '], [' 53', 'Aleabaan ']].

Cluster No.3:

[[' 19', 'Fronto '], ['29 ', ' Nasuhi '], ['30 ', ' Sourani '], ['31 ', ' Zayti '], ['32 ', 'Alnabali '], ['33 ', ' Miski '], [' 34', ' Zlitny '], ['39 ', 'Alsayquz '], [' 40', 'Alzaraziu '], ['49 ', ' Jbuji']]

After performing the Normalization operation, the shape of the clusters became as follows in figure no.7 and noting that each cluster is indicated by (+) symbol.

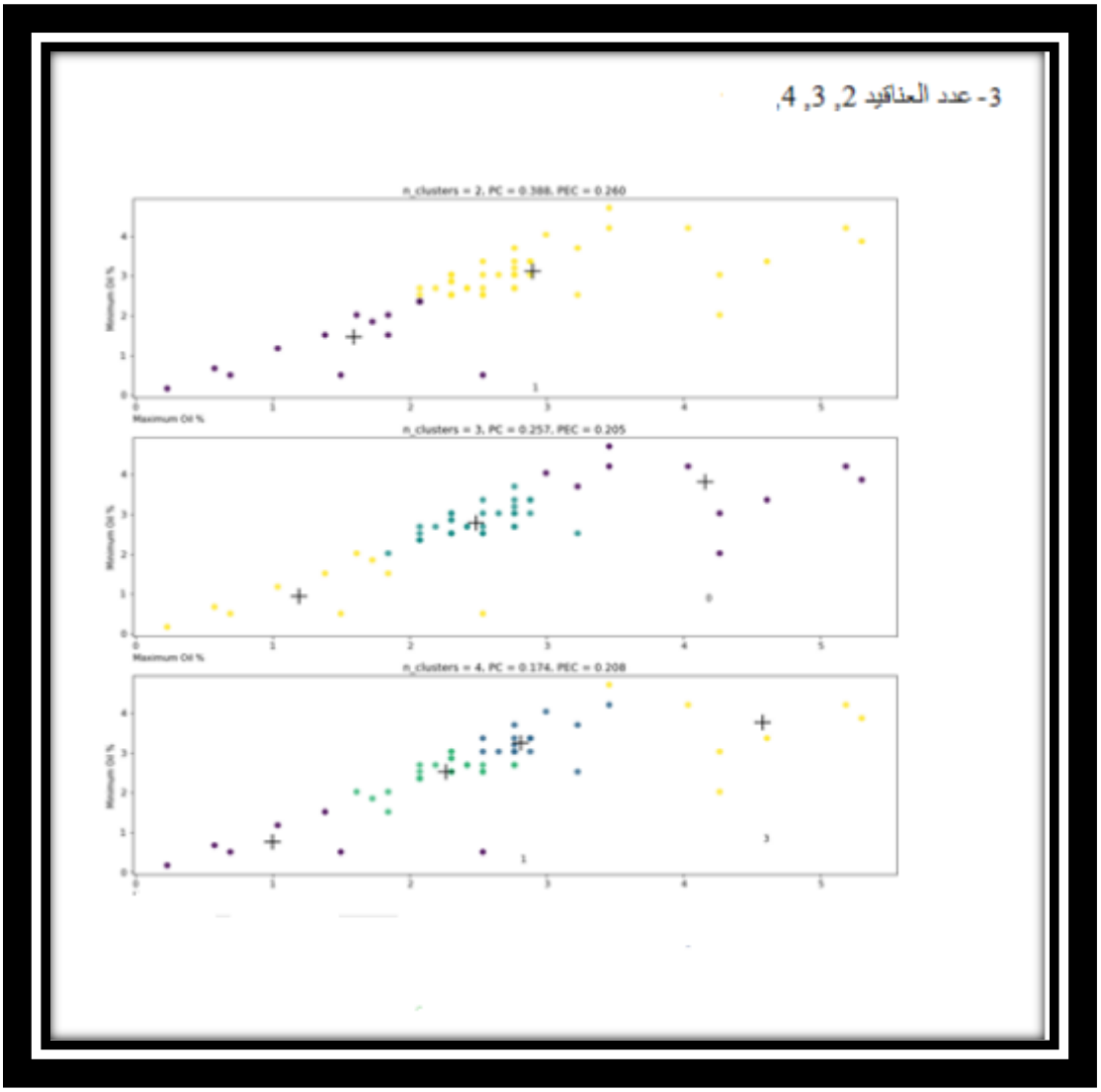

Figure (7) clarify the last shape of the clusters after the Normalization

5- Interpretation of results:

1- From Figure (7), we notice the convergence of the distances between the varieties, (Diql, Al-Ajizi Al-Shami, Al-Ajizi Al-Aqsi, Frantoio, Asklano, Aldaeaybali, Mahzam Abu Satl, Almahat)i, It is the smallest of the clusters, all of them are located in the first cluster, most of them which are Mediterranean species. While the varieties that are not located on the Mediterranean, they can be grown in the countries of the Mediterranean because the distances and their genetic characteristics are closed.
2- The following items are clustered: Khastawi, Bashiqa (regular), Dulce, Baronial, Alhamid, Alwatiqain, Kalamata, Erbecoin, Shamilali, Fronto, Alshatwi, Manzanillo, Picual, Koratina, Kronaki, Wasalati, Almaraqi, Khudairi, Nasuhi, Sourani (Al-Ma'arri), Zayti, Alnabali, Miski, Zlitny, Koroniki, Kayrusawy, Alnabali Almuhsin, Alsayquz, Alzaraziu, Alzalmatiu, Aljarbueiu, Arbiykuyna, Alquysi, Bishulyn Almaghribiy, Liytshiynu, Jbuji, Aldaan, Laymili, Aleabaani. This cluster is the largest, that is to say the items within this cluster are closed, even if they are from different origins.
3- Items are clustered: Fronto, Nasuhi, Sourani (Al-Ma'arri), Zayti, Alnabali, Miski, Zlitny, Alsayquz, Alzaraziu, Jbuji. The varieties here have similar characteristics, but the convergence is weak.
4- The best result of clustering was in two clusters only.
5- The result was not good and not clear when the number of clusters was 7.

Conclusions:

1- The FCM algorithm gives great and clear results; its results are still impressive and rival its predecessor’s algorithms.
2- FCM showed an algorithm Great flexibility in dealing with different types of data and data processing incomplete.
3- The FCM algorithm is smooth in dealing with various systems and software.
4- Absence of studies related to linking fuzzy clustering techniques to the axis of agriculture in all its forms.
5- The fewer clusters the better the results

Suggestions:

1- Keeping developments in the field of data mining and related algorithms.

2- Adopting graphic models by adopting fuzzy algorithms to create an appropriate environment for making agricultural decisions that it represents one of the modern treatment methods for decision-making mechanisms during an environment full of variables.
3- The work can be developed in different directions and in many real-life applications.
4- Working on making the farming systems all controlled by the computer because of the consequent making of it controlling the results of the diagnosis.

- **References:**

- Almazrooie, Mishal,Vadiveloo, Mogana, Abdullah, Rosni, (2016) ,"**GPU-Based Fuzzy C-Means Clustering Algorithm for Image Segmentation**", Universiti Sains Malaysia, 11800 USM, P.,Pulau Pinang, Malaysia.
- Francesco, Masulli (2015)" **Fuzzy Clustering**", Dip. Informatica, Bioingegneria, Robotica e Ingegneria dei Sistemi University of Genova - Via Dodecaneso 35, 16146 Genoa – ITALY.
- G., Mohammed & F., Ahmed, Mohammed, & Uddin, Mohammed Y. (2016), "**Clustering as Datamining technique in risk factors analysis of Diabetes, Hypertension and obesity**", Vol.1, No.6, EJERS, European Journal of Engineering Research and Science.
- Han, Jiawei, Kamber, Micheline, Pei, Jian (2016), "**Data mining concepts and techniques**", 3rd edition, ELSEVER.
- Jyuti, Neha, & Kaushok, Rekha (2014), "**Review Paper on Clustering and Validation Techniques**", International Journal for Research in Technology (IJRASET), Vol.2, Issue V.
- Kassambara, Albokadel (2017), **"Practical guide to cluster analysis in R unsupervised machine learning**", Ed.1, STHDA
- Klawonn, Frank and Hoppner , Frank, (2003) "**What is Fuzzy About Fuzzy Clustering? Understanding and Improving the Concept of the Fuzzifier**", Department of Computer Science, University of Applied Sciences Braunschweig / Wolfenbuettel , Salzdahlumer Str. 46/48 D-38302 Wolfenbuettel, Germany.

- Li. Li & sang, Qifa, & Yang, Xi (2022), "**K-means Clustering of Overweight and Obese Population using Quantile-Transformed Metabolic Data**", http://doi.org/10.2147/DMSO.S.206640
- Pande ,S. R., Sambare ,Ms. S. S., Thakre,V. M.,(2012) **" Data Clustering Using Data Mining Techniques**", International Journal of Advanced Research in Computer and Communication Engineering Vol. 1, Issue 8, India. web site: www.ijarcce.com.
- Zupan, Blaz and Demsar, Janez(2018), "**Introduction to data mining**", University of Ljubljand, Slovenia, CC by: NC-ND.

- Internet sites:

- Deokkar, Siddharth, "WEIGHTED K-NEAREST NEIGHBOR", (2009) CS 8751, http://www.csee.umbc.edu/
- http:// homes.di.unimi.it/ valenti/SlideCorsi/Bioinformatica05/ Fuzzy- Clustering-lecture-Babuska.

- Data sources:

- Report of the Food and Agriculture Organization of the United Nations (2020).
- Abdul Qadir, Muayyad Salih and Muhammad, Asma Jassem (2011), "Olives (Green Gold) in the Qur'an and Sunnah and the Prospects for Developing Its Production in Iraq - in the Light of the Experiences of Some Countries", Iraqi Journal of Market Research and Consumer Protection, Volume (3), Number (5) , Iraq.
- http://ejabat.google.com.
- http://ar.wikipedia.org
- http://arshad2010.yoo7.com/t179-topic.
- www.aoad.org/olive/jor