Classification of Spatial Data Based on K-means and Voronoï Diagram

Moubaric KABORE, Béné-wendé Odilon Isaïe ZOUNGRANA, Abdoulaye SERE
ED-ST, LAMDI
Université Nazi BONI
BOBO DIOULASSO, BURKINA FASO

***Abstract*—This paper is focusing on the problem of the time taken by different algorithms to search data in a large database. The execution time of these algorithms becomes high, in the case of searching data in a non-redundant data, distributed in different database sites where the research consists of reading on each site for finding data. The main purpose is to establish adapted models to represent data in order to facilitate data research. This paper describes a classification of spatial data using a combination of k-means algorithm and voronoï diagram to determine different clusters, representing different group of database sites. The advantages of classification is made through the k-means algorithm that defines the best number and the centers of required clusters and voronoï diagram which gives definitely the delineation of the area with margins, representing the model of organizing data. A composition of K-mean algorithm followed by voronoï diagram has been implemented on simulation data in order to get the clusters, where future parallel research can be realized on different cluster to improve the execution time. In application to e-health in GIS, a best distribution of medical center and available services, will contribute strongly to facilitate population well-being.**

***Keywords*—Classification; K-means; voronoï diagram; GIS; big data; data research**

I. Introduction

Data research is often made through different requests submitted to the database. The problem of data research become complex in the context of multi database not-connected, distributed in several sites, having big data. For instance, in the structure of a spatial database, where each GIS position corresponds the localization of a database site with big data, the execution time of data research will depend on the number of database site.

Classification is necessary to organize data in order to facilitate data research. The criteria of classification define requirements that must be respected to put together the data in the same cluster. The application of criteria could lead to have partitions, called clusters to store data. Each cluster contains similar elements regarding to criteria definition. Many techniques of Classification have been proposed by different scientists. For instance, Koperski et al. in [8] have proposed an efficient two-step method for classification of spatial data.

There are also supervised classification and unsupervised classification. Supervised classification consists of labeling the dataset and assigning a new data to a pre-existing class. K-nearest neighbors (KNN), decision trees, neural networks, support vector machines(SVN) and Bayes classifiers are supervised classification. For. Instance, Karem et al. in [2] used a supervised classification to study the theory of believing fonction. Kessler et al. in [6] applied neural network to classifier mails.

Unsupervised classification consists of finding a group in unlabeled dataset for a new observation. K–means, K–meloïd, hierarchical classification are unsupervised classification. For instance Kodinarya et al. in [4] have realised a review on determining number of cluster in k–means clustering. Sinaga et al. in [5] proposed unsupervised k–means clustering algorithm.

Meshing space could lead to supervised classification or unsupervised classification. Regular grid or irregular grid are meshing technique. Bottela et al. in [7] generated a mesh for modeling simulation for physique phenomene. Duchaine in [9] also generated a structured mesh kriging with local elliptical refinement. Antoine Vacavant in his thesis, proposed a survey on regular and irregular grids that could be considered as supervised classification.

The initial works of Sere et al. in [1] impose artificial meshing on data that leads to empty clusters used for the map-reduce algorithm. The size of clusters does not take into account the number of data inside.

Our purpose is to reduce the execution time of data research algorithms, in building a model of data structure with relevant clusters. The achievement of this purpose will lead to avoid many artificial clusters and to remove empty clusters.

Our hypothesis is that an adequate model of data structure will contribute to reduce the execution time of data research.

This paper follows and extends the initial works in [1], to explore the combination of k-means clustering and voronoï diagram successively on simulation data, as an alternative to create a model for data structure : the outputs of k–means clustering with different centers are the inputs of voronoï diagram algorithm to get definitely clusters.

The proposed method could be applied to the data set in e-health, representing data related to the sites of pharmacies with drugs characteristics or medical centers. More others fields are concerned with applications, as to find the nearest vehicle seller, the nearest security station in a region.

This paper is organized as follows: Section 2 presents the state of art with classification techniques. Section 3 describes the method while Section 4 shows the results of implementation.

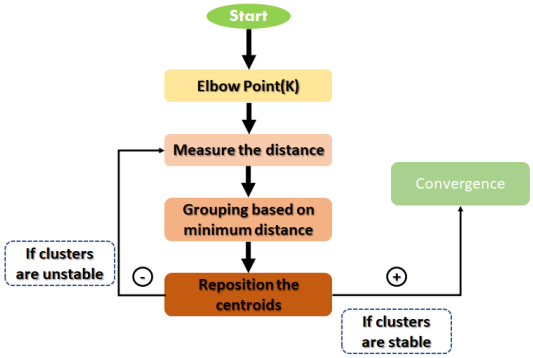

Fig. 1. K-means algorithm.

II. Preliminary

A. Problem Statement

The pharmacies of the city of Ouagadougou have databases within their premises which store the name, the quantity of stock and the price of each product. They are located in different districts. We do not know in its various pharmacies which ones have more influxes. The establishment of a new pharmacy is done without optimization. The acquisition of products is still archaic because the user must always go to a pharmacy before obtaining information on the stock of the product he is looking for. We have developed in our master's thesis and in the article SERE and Al [1], the parallel search for a datum d in several spatial databases by prioritizing the closest database. If we assume B the set of databases of all pharmacies in Ouagadougou and B_i the database associated with each pharmacy with B= $\{B_1, B_2, ..., B_n\}$; we are dealing with a large volume of data commonly called big data.

Our study consists in optimizing the classification of the different cluster pharmacies and the meshing of the clusters obtained for a better spatial distribution of the pharmaceutical sites and better search results.

B. K—means

The k-means algorithm, proposed by MacQueen (1967), is an aggregation technique around mobile centers [10]. The principle in [11] is to place the k chosen points in the space represented by the individuals to be classified. These k points go formerly the first centers of the different clusters. In addition, the barycenter of each cluster is calculated. Finally, the individuals are reassigned according to the proximity to the new center Fig. 1.

C. Mesh Techniques

A mesh is a partition of space or a domain into a set of elementary cells that have coordinates, dimensions and information on the connectivity relationship between its cells (vertices, edges, faces, volume). Bastian and al in [17] used the connectivities for the implementation of a c++ algorithm that generates quadrilateral meshes. Consider a bounded domain noted D of $\mathbb{R}^2$ and $\mathbb{R}^3$, P_h is a mesh of D if:

- $D = U_{K \in P_h} K$
- The interior of any element K of P_h is no empty
- The intersection of the interior of two elements is empty.

There are tree types of mesh as structured, unstructured and hybrid presented in [9].

1) Structured Meshes: Consider any bounded domain noted Q to which a mesh has been applied. The mesh of the domain Q is said to be structured if and only if from a single one of its meshes it is possible to reconstruct the whole identical mesh. It is characterized by a cell that repeats itself identically and ordered in [3], [9].

2) 1D mesh: The mesh is carried out on the right of the abscissa whose domain is a segment [AB] in Fig. 2. This segment is partitioned into m cells with a constant pitch C which is equal to $\frac{B-A}{m}$. Then $X_1 = A, X_{i+1} = X_i + C$ and $X_m = B$ with $i = 1, 2 \ldots m$.

The relationship $X_{i+1} = X_i + C$ expresses the connectivity link between two cells.

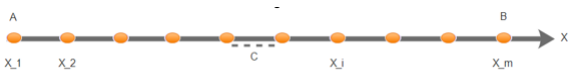

Fig. 2. 1D mesh.

3) 2D mesh: For a rectangular domain (a, b) x (c, d) we have:

- $X_1 = a, X_i + 1 = X_i + C$ with C the pitch of the mesh in the direction X; $C = \frac{b-a}{m}$ and $i = 1, 2...m$
- $Y_1 = a, Y_j + 1 = Y_J + K$ with K the pitch of the mesh in the Y direction; $K = \frac{b-a}{m}$ and $j = 1, 2 \ldots m$
- Cartesian and polar mesh (cylindrical and spherical in 3D)

This mesh is composed of pairs of transverse lines which intersect at the nodes of the mesh and a cell which is repeated with the same number of nodes around a vertex. Each node is obtained thanks to the coordinates (i, j) formed by the indices of the transverse lines. The decrementation and/or the incrementation of the coordinates (i, j) of a node makes it possible to locate its neighbors. Then the incrementation and the decrementation constitute the relations of connectivities between the different cells.

The structured mesh reveals its weaknesses when it comes to the meshing of a domain and its borders. Moreover one cannot control neither the form nor the distribution of the cells in the field, one cannot either mesh according to the simulated phenomenon. To make up for this shortcoming, we are going to switch to the unstructured mesh.

4) Unstructured meshes: The mesh of an unstructured bounded domain Q is characterized by unordered cells and all its cells are not identical [9]. Owen described the majority of unstructured meshes which we summarize in [3].

After the generation of the mesh algorithms necessarily comes the step of improving the overall quality of the elements.

The two types of mesh enhancement smoothing and cleaning developed by Owen in [3].

- Smoothing is the adjustment of nodes without any change in connectivity between elements.
- Cleaning is the adjustment that modifies the connectivity of the elements.

Bastian and al in [17] presents an unstructured quadrilateral meshing algorithm that generates the mesh by recursively subdividing domains.

5) Hybrid meshes: it is a mesh that combines both structured and unstructured elements. source [9]

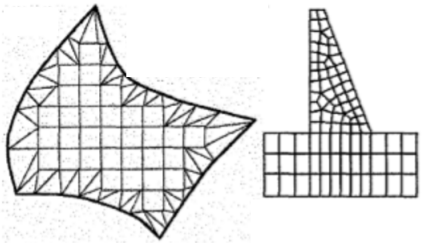

Fig. 3. Hybrid mesh.

6) Description of the Voronoï-Delaunay mesh: Johann Peter Gustav Lejeune Dirichlet in [13] introduced the Voronoï diagram in 1850, which is how a given domain could be systematically decomposed into a set of compact convex polygons. Georgy Voronoï in [16] formalized this notion in the general case in 1908.

Definition 1: Voronoï diagram [12], [14], [15]: Let P be a set of points p_i such that $1 \leq i \leq n$. Let E be a Euclidean vector space such that $P \in E$. Consider a point among the n points of P and denote the p_i. Let us delimit the space which gathers all the points of E which are closer to p_i than the rest of the points of P. This delimitation forms a cell called the Voronoï cell of the point p_i and is denoted C_i. In the cell C_i all the points of E are closer to p_i than the rest of the points of P. Let p_i be the set of points of P except p_i and q the set of points of C_i we have: $d(q, p_i) \leq d(q, p_j)$; the point p_i is called the germ of the cell C_i.
Next, take another point of P and delimit its corresponding Voronoï cell. Let us do the same for all the points of P. We note that the space E is subdivided into cells C_i associated with the points p_i with $1 \leq i \leq n$. This subdivision of space E into cells C_i associated with points p_i with $1 \leq i \leq n$ is called Voronoï diagram and is denoted Vor(P). Note that the union of cells C_i is equivalent to the Voronoï diagram of P: $Vor(P) = U_{pi \in P} C_i$ The boundary between two cells is called the Voronoï edge and is halfway between the two seeds corresponding to the two cells in 2D, it is a segment of the perpendicular bisector of the line joining these two seeds.
Drawing:
Case 1: two cinemas A and B are at different positions. From my position, if I want to go to the nearest cinema, it is in my interest to head towards a Fig. 4.

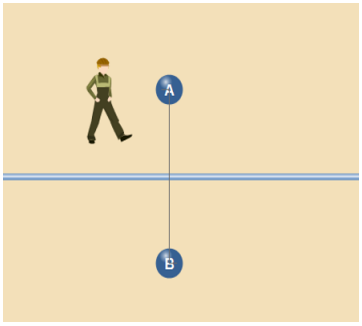

Fig. 4. Closest distance between two sites.

If all the pairs of points that have an edge in common are connected by straight lines, the result is a triangulation within the convex hull of the set of points $\{Pi\}$. This tessellation is known as the Delaunay triangulation in [12], [14], [15].

There is a particular similarity between the k-means classification and the space mesh with the voronoï diagram. The synthesis of the two methods could be a basis for the grouping and distribution of sites in clusters in order to organize the distribution of data in space and to facilitate their search.

III. Method Description

Pharmacies in Ouagadougou are located in the different districts with the aim of bringing their services closer to beneficiaries. This requires a good distribution or network of sites in the locality while projecting the coverage of users by each pharmacy. Pharmaceuticals are widely diversified. Thus it is difficult or even impossible for a pharmacy to have all types of products in stock. Consequently, users must go to other pharmacies for cases of missing product. To do this, we are going to classify the pharmacies into clusters to guide users to search for products first in the cluster in which they are housed, then in the nearest neighboring cluster in the event that they have not obtained all of their products within their cluster.

Our objective is to optimize the distribution or meshing of sites in space, to classify them into groups or clusters in order to facilitate the search for a service according to the position of the user in relation to the different groups of sites. To do this we will make a synergy between the k-means algorithm and the voronoï diagram.

We will first form the clusters. Each cluster consists of a set of pharmacy databases. It occupies a well-defined space that we call a cell. When a new pharmacy is created, it is assigned to the nearest cell. We will then look for the cluster closest to the position of the user who is looking for information on the disposal of a pharmaceutical product.

A. Cluster Creation

Let k $\in \mathbb{N}$ and k less than the pharmacy numbers. We will use the k–means algorithm to group database sites into k-clusters. The centroids of the different clusters are called seeds. We will then use the seeds to build the Voronoï diagram. We get k-cells. We've just applied the function kmeans$\circ$Voronoï or f$\circ$g if we consider that kmeans is the function f and

Voronoï is the function g. Each cell houses a cluster and each cluster groups pharmaceutical database sites. The GPS(x,y) coordinates of each seed and all of its database sites are collected. The following algorithm makes it possible to create the k-clusters and to accommodate them in cells thanks to the Voronoï diagram.

1) Cluster creation algorithm: Cluster creation is define by the procedure:

- Group objects into k-clusters using the k-means algorithm.
- Using the centroids of the different clusters as seeds, separate the locality into k-cells using the Voronoï diagram. Each cell gathers the sites of the cluster and the barycenter becomes the seed of the cell.
- Locate the GPS coordinates (x, y) of the seed of each cell and all of its sites.

Note: This algorithm is called kmeans∘Voronoï.

2) Nearest cell search algorithm: The following algorithm determines the cell closest to any point. The coordinates of this point are determined using GPS. The distance used is the Euclidean distance.

- Locate the GPS(x,y) coordinates of the person concerned.
- Calculate the Euclidean distance between the position of interest and the different seeds of the cells $d(i,j) = \sqrt{(x_{i1} - x_{j1})^2 + (x_{i2} - x_{j2})^2}$.
- Compare distances.
- Choose the minimum distance as the nearest cell.

Given a set S composed of n points $(y_1.....y_n)$ in a space E, k an integer less than n and x another point not belonging to the set S. We are looking for the k nearest neighbor of x. The neighborhood problem first involves knowing the distance between x and the different points of S. The algorithm below calculates and displays the Euclidean distance between two points A and B.

For k=1 the nearest neighbor search is limited to the single nearest neighbor of x. We are looking for the smallest distance between x and y_i for i ranging from 1 to n. Let d_i=d(x,y_i) then the nearest neighbor algorithm amounts to finding the smallest d_i for i ranging from 1 to n. Let d_i=d(x,y_i) then the nearest neighbor algorithm amounts to finding the smallest d_i for i ranging from 1 to n.

Algorithm 1 NearestNeighbor

```
Variable i, n, index: integer
procedure index ⟵ 1
    for i ranging from 2 to n do
        if D[index] > D[i] then index =i
        end if
    end for
    Display D[index] is the smallest distance
    Show y[index] is nearest cell
end procedure
```

We can use the two previous algorithms to determine the nearest neighbors to the farthest neighbors. For k=1, it is the nearest neighbors algorithm. We will rank the nearest neighbors to the farthest neighbors. So we're going to sort the distances from the smallest to the largest.

Algorithm 2 RankNearestNeighbor

```
variable i, j, integer
procedure
    for i going from n to 1 do
        for j ranging from 2 to i do
            if D[j-1] < D[j] then
                Exchange (D[j-1],D[j])
            end if
        end for
    end for
end procedure
```

IV. Experimental Result

The choices of tools used to implement k–means and Voronoï diagram are based on reliability, flexibility and concordance with the method description.

- R software: It is a programming language and a free open source software environment mainly dedicated to statistical analyses, data science and graphical representations.
- RStudio is an interface facilitating the use of R. It is also free and freely distributed. The RStudio interface is divided into four windows: the console where you can execute commands; the environment where we can view the constructed objects; files and plots where you can install packages and view graphics; R script where you can keep the command lines as well as the comments.
- Installation of packages and import of libraries required.

We will first create a "mydata" data table which contains the x and y coordinates expressing the spatial position of the pharmaceutical sites.

Algorithm 3 Algorithm for creating a data table

```
Xvalue ← c(4, 1, 0, 20, 7, 7, 3, 10, 16, 1, 2, 11, 25, 1, 6,
  1)
Yvalue ← c(7, 9, 20, 11, 8, 10, 40, 25, 33, 21, 11, 8, 13,
  13, 12, 16)
mydata ←data.frame(x = Xvalue, y = Yvalue)
```

Consider the x and y coordinates of the following 16 sites below. We calculate the barycenters of the x and y coordinates for each cluster. We will first choose the number of cluster k using the curve we generate with the tools.

TABLE I. <X AND Y COORDINATES>

Order	X	Y
1	4	7
2	1	9
3	0	20
4	20	11
5	7	8
6	7	10
7	3	40
8	10	25
9	16	33
10	1	21
11	2	11
12	11	8
13	25	13
14	1	13
15	6	12
16	1	16

To classify sites into groups we need to know the number of groups k. The best choice of this number k makes it possible to obtain very homogeneous groups. The R software allows us to draw a curve with elbows. The numerical value of the elbow with the greatest inflection corresponds to the best k. Algorithm 4 allows to draw the curve in Fig. 5.

Algorithm 4 Algorithm for determining best k

{plot(1:10, t, type = "b", pch = 19, frame= FALSE, xlab = "Number of clusters (k)", ylab = "Intra-cluster sum of squares (wss)", main = "Optimal number of clusters k")}{}

We obtain the curve in Fig. 5 that presents the choice of the number of clusters.

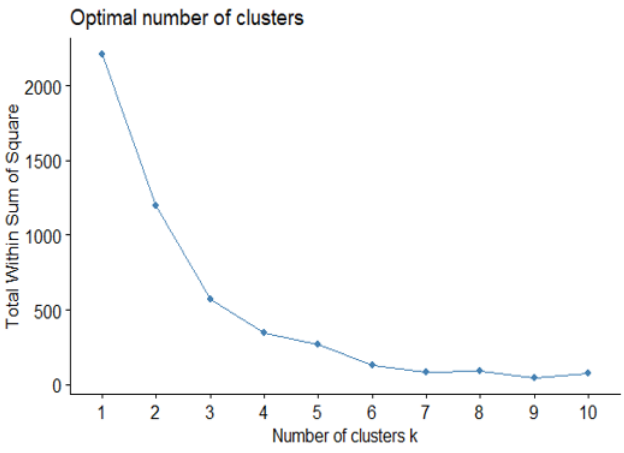

Fig. 5. Choice of the number of clusters.

The best number of clusters is obtained by choosing the elbow which presents an inflection. By observing we notice that k=3 and k=4 present inflections. We will work with his two cases and choose the best one in the end.

We will calculate the coordinates of the different centers of the k groups using the Algorithm 5.

Algorithm 5 Cluster center calculation algorithm

```
for (k in 1:10) {
kmeans result ← kmeans(data, centers = k)
t[k] ← sum(kmeans result$tot.withinss) }
 k=3
kmeans result ← kmeans(data, centers = k)
 cat(" Clusters centers : \n")
 print(kmeans_result$centers)
```

For k=3 the centroids of the clusters obtained have x and y coordinates in the Table II:

TABLE II. <X AND Y COORDINATES>

Order	X	Y
1	3.72	12.27
2	9.66	32.66
3	22.50	12.00

Algorithm 6 shows the representation of three groups.

Algorithm 6 Cluster graphing algorithm

fviz_cluster(kmeans_result, data = mydata, geom = "point", ellipse.type = "convex") + labs(x = "X-axes", y = "Y-axes", title = " K-means analysis")

We obtain three clusters grouped together in space as shown in Fig. 6:

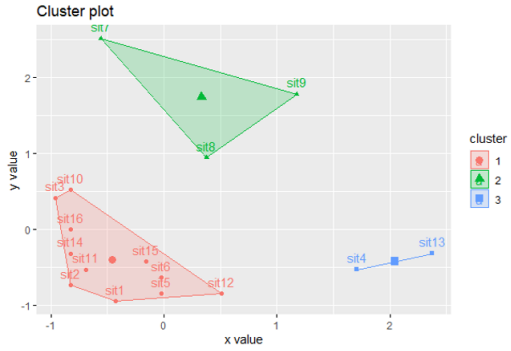

Fig. 6. Grouping into three kmeans clusters.

To finish with the classification of the three groups, Algorithm 7 allows us to display a Table III which indicates the coordinates of each site and its group to which it belongs.

Algorithm 7 Cluster membership table algorithm

```
r ← data
rcluster < −as.factor(kmeans − resultcluster)
print(r)
```

TABLE III. <X, Y COORDINATES AND THEIR CLUSTER>

Order	X	Y	cluster
1	4	7	1
2	1	9	1
3	0	20	1
4	20	11	3
5	7	8	1
6	7	10	1
7	3	40	2
8	10	25	2
9	16	33	2
10	1	21	1
11	2	11	1
12	11	8	1
13	25	13	3
14	1	13	1
15	6	12	1
16	1	16	1

We will now apply the voronoï diagram to these three clusters. We manage to separate them into three cells. Each cluster covers its corresponding cell. We get Fig. 7 and 8.

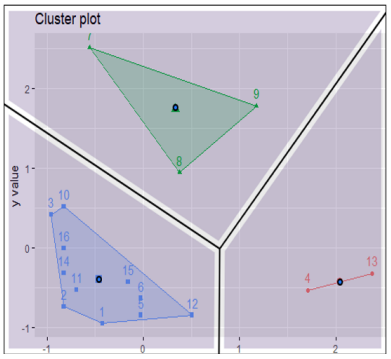

Fig. 7. Delineation of the area of each cluster with the voronoï diagram.

Fig. 8. Representation of each voronoï cell of each cluster.

For k=4 the centroids of the clusters obtained have x and y coordinates in the Table IV below:

TABLE IV. <X AND Y COORDINATES FOR K=4>

Order	X	Y
1	5.42	9.28
2	9.66	32.66
3	22.50	12.00
4	0.75	17.50

We obtain four clusters grouped in space as shown in Fig. 9.

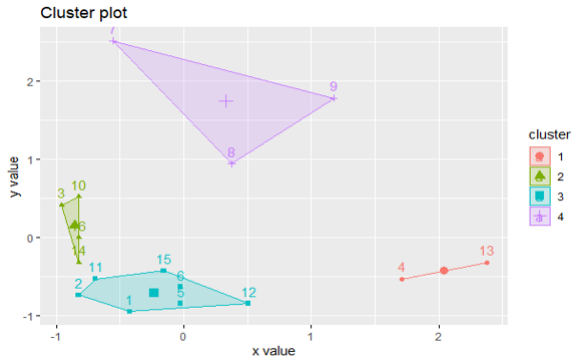

Fig. 9. Grouping into four k–means clusters.

We will now apply the voronoï diagram to these four clusters. We manage to separate them into four cells. Each cluster covers its corresponding cell. We get Fig. 10 and 11.

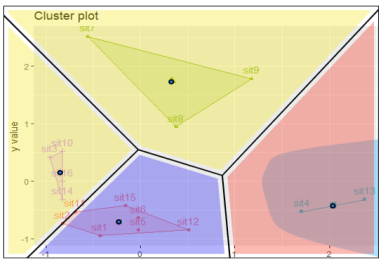

Fig. 10. Delineation of the area of each cluster with the voronoï diagram.

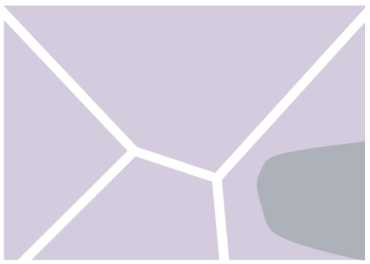

Fig. 11. Representation of each voronoï cell with its corresponding cluster.

V. DESCRIPTION OF THE PROCESS WITH THE CASE OF PHARMACIES

When a user searches for a pharmaceutical product, he can launch the search from his position. The platform retrieves its position and its product then searches in the cluster closest to its position, if no pharmacy contains the product then the search continues to the next cluster. When the product is found, the pharmacy containing the product is indicated.

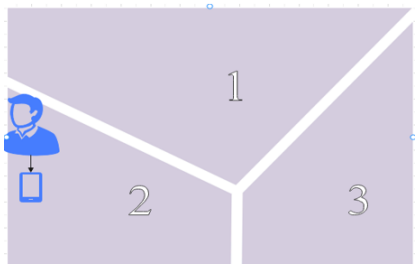

Fig. 12. Search for a pharmaceutical product.

In Fig. 12, the user using his telephone searches for doliprane. With regard to its position, the search begins in cluster 2 and ends if cluster 2 contains doliprane, otherwise the search continues in cluster 1 and ends if the latter contains doliprane, otherwise the search continues and ends in cluster 3. When the product is found, the platform returns the name of the pharmacy that contains it and its cluster.

VI. Conclusion and Perspectives

This paper carried out an unsupervised classification as k-means clustering of pharmaceutical sites in space to obtain different groups. Border Delimitation of each group has been done by voronoï diagram, to obtain definitely clusters. In this classification, data research could be firstly done with the closest cluster in serial case or in parallel on several clusters, simultaneously with map-reduce framework.

The future works will concern with analysis of new site insertion in the area that could change data structure, to lead to new clusters creation and to complete implementation of proposed algorithms in a software. Moreover, each database site or cluster will be also associated to a probability as Bayesian probability in order to start data research with cluster having the best probability.

Analysis could consider others criteria, as open sites in the schedule, the less expensive products in the sites, the existence of a path to connect the sites according to the initial position of users.

References

[1] Abdoulaye SERE, Jean Serge Dimitri OUATTARA, Didier BASSOLE, Jose Arthur OUEDRAOGO and Moubaric KABORE, “A Framework for Data Research in GIS Database using Meshing Techniques and the Map-Reduce Algorithm” International Journal of Advanced Computer Science and Applications(IJACSA), Vol. 12, No. 3, 2021.

[2] Karem, Fatma, Mounir Dhibi, and Arnaud Martin. ”Combinaison de classification supervisée, non-supervisée par la théorie des fonctions de croyance.” EGC. 2012.

[3] S. J. Owen, A survey of unstructured mesh generation technology, Proceedings, Seventh International Meshing Roundtable, Sandia National Laboratories, Albuquerque, NM, 1998, pp. 239–267.

[4] Kodinariya, Trupti M., and Prashant R. Makwana. ”Review on determining number of Cluster in K-Means Clustering.” International Journal 1.6 (2013): 90-95.

[5] Sinaga, Kristina P., and Miin-Shen Yang. ”Unsupervised K-means clustering algorithm.” IEEE access 8 (2020): 80716-80727.

[6] Kessler, Rémy, Juan Manuel Torres-Moreno, and Marc El-Bèze. ”Classification thématique de courriels avec apprentissage supervisé, semi-supervisé et non supervisé.” les actes de VSST (2004): 493-504.

[7] Botella, Arnaud. Génération de maillages non structurés volumiques de modèles géologiques pour la simulation de phénomènes physiques. Diss. Université de Lorraine, 2016.

[8] Koperski, Krzysztof, Jiawei Han, and Nebojsa Stefanovic. ”An efficient two-step method for classification of spatial data.” proceedings of International Symposium on Spatial Data Handling (SDH’98). 1998.

[9] Duchaine, François. Génération de maillage structuré par krigeage avec raffinement elliptique local. Diss. École de technologie supérieure, 2004.

[10] Gelb, Jérémy, and Philippe Apparicio. ”Apport de la classification floue c-means spatiale en géographie: Essai de taxinomie socio-résidentielle et environnementale à Lyon.” Cybergeo: European Journal of Geography (2021).

[11] Idbraim, Soufiane. Méthodes d’extraction de l’information spatiale et de classification en imagerie de télédétection: Applications à la cartographie thématique de la région d’Agadir (Maroc). Diss. Université de Toulouse, Université Toulouse III-Paul Sabatier, 2009.

[12] LEE, Der-Tsai and SCHACHTER, Bruce J. Two algorithms for constructing a Delaunay triangulation. International Journal of Computer and Information Sciences, 1980, vol. 9, no 3, p. 219-242.

[13] LEJEUNE DIRICHLET, G. Über die Reduction der positiven quadratischen Formen mit drei unbestimmten ganzen Zahlen. Journal für die reine und angewandte Mathematik (Crelles Journal), 1850, vol. 1850, no 40, p. 209-227.

[14] Fortune Steven. ”Voronoi diagrams and Delaunay triangulations.” Handbook of discrete and computational geometry. Chapman and Hall/CRC, 2017. 705-721.

[15] GREEN, Peter J. et SIBSON, Robin. Computing Dirichlet tessellations in the plane. The computer journal, 1978, vol. 21, no 2, p. 168-173.

[16] G. Voronoï, Nouvelles applications des paramètres continus à la théorie des formes quadratiques, J. Reine Angew. Math. 134, 198–287 (1908)

[17] M.Bastian, B.Q. Li, An efficient automatic mesh generator for quadrilateral elements implemented using C++, Finite Elements in Analysis and Design, Volume 39, Issue 9, 2003, Pages 905-930, ISSN 0168-874X