Research Article
Volume 13:4,2020
DOI: 10.37421/jcsb.2020.13.311

Journal of Computer Science & Systems Biology

ISSN: 0974-7230

Open Access

A Study of Using Custom-Clustering Algorithm for a New Treatment of COVID-19

Albert Figueras, Santiago Esteva*, and Josep Lluís De La Rosa

Vicorob Research Institute, University of Girona, Girona, Catalonia, Spain

Abstract

This work deals with the problem of knowing a group of people that adequately responds to a specific treatment in order to classify community in groups is the objective. In the process to make this classification, a lot of work is necessary to analyze the results from the cluster analysis and obtain the minimal parameters that define a specific group that can be classified as treatment target. Also is presented the algorithm called Custom-clustering to solve this problem.

Keywords: Classification • Clustering • COVID-19

Introduction

Classification of objects into meaningful groups is a central task in Science [1]. Cluster analysis is a statistical technique specialized to classify units into groups. Although cluster analysis is widely employed in other disciplines, it's used in Political Science, sociology, empirical research, then using factor analysis and other multivariate statistical techniques [2]. The principal aim of this paper is to present an algorithm to cluster analysis for medical treatment or political interest.Difference to neural network in AI, they try to adjust one Hypothesis to a neural net [3], adjusting weights to the examples, otherwise the clustering method give the true hypothesis from de experimental data. For a case to identify a chair, is necessary to deeply search in the NN to discover the weight for a two foot in a chair is 0, but in clustering there are two different areas 1 and 3 to 4 foot, at the end takes the same results from the same data.

Methodology Adopted

Multiple treatments are being experimented to reduce or cure the effects of COVID-19 in humans, a new method starting from a plasma with high-power immune agents, Grifolsenterprise produce Hyperimmune globulins designed to give a patientimmunity. From which it must be determined accurately the population that can be used with good results. One method to determine the limits of the set of population that is determined by certain characteristics that defines it, is the classification by clustering. In this work we present a classification algorithm using clustering function that in the first part of the algorithm adjusts the parameters to control the group at desired target specification, and the second part determines the minimum necessary characteristics to define it. Unlike the main PCA component analyzes, which determines a linear combination of the axes or parameters that determines the group, the resulting axes do not represent real values of characteristics of the subject, and neither to adjust the cluster to some specification.We will start from possible characteristics that define each subject, and through sampling we obtain a representation of the possible spectrum of people to whom the treatment has been applied, and from there we look for the population group that best accepts the treatment, it is called classification with labels, which can be: negative, in the case of harming the patient, positive otherwise, and neutral without effect on the subject. The objective is to determine apopulation group where the treatment is adequate and necessary, within desired margins age, harm, etc. and with sufficient representation, sampling to be able to consider the results valid and representative.

In many situations the superficial analysis of the results usually leads to say that they are not conclusive, because certain parameters indicate good performance in some range, and scattered in another, this does not mean that this parameter is not useful. It can be possible that said variable depends at the same time on another, so when representing the results in 2 dimensions it can be seen in some cases that the second variable displaces the results and discriminates the positives from the negatives according to their magnitude. Finding these relationships, in a multi-related and multidimensional space, is the objective of this work, and in the end to enclose the possible group among the indispensable parameters.For this, the maximum number of variables that could have information to define the group are collected, such as age, blood group, diseases suffered, and active, treatments, etc. This does not allow a representation by color results to visualize the position of the group, due to the multi-dimensionality of the data. We must analyze the data by the algorithm provided to enclose and reduce the variables to the minimum necessary.

Custom-Clustering algorithm

The name of Custom-clustering came from the possibility of the algorithm to close the group of individuals as desired. Difficulty to discover small group into the middle of the others big groups, and also depending on the sample density, is shown in the Figure 1. Is easy to imagine a group into 2 dimension representation and in colored results, but normally the data is multidimensional, with axes with information and others no. The work presented is to analyze these axes and eliminate the non-profitable, and find the group with a validating measure. We use cluster function provided from Matlab [4].

***Address for Correspondence:** Santiago Esteva, Vicorob Research Institute, University of Girona, Girona, Catalonia, Spain, Tel: +34 626411206; E-mail: santiesteva@gmail.com*

***Rec date:** 16 June 2020; **Acc date:** 07 July 2020; **Pub date:** 14 July 2020*

T=Cluster (Z,' MaxClust', N) constructs a maximum of N clusters using distance as a criterion.

The function permits to find a numbers of groups depending of the N parameter, and modifying it permits to adjust one desired group as fine as possible.

The difficulty to obtain at this time true data for use the algorithm in health problem make us to demonstrate their use into the problem of social interest, useful for commercial and electoral activities as a new paradigm came from the social media, which generates big amount of information from the users according its movements into the smartphones apps.

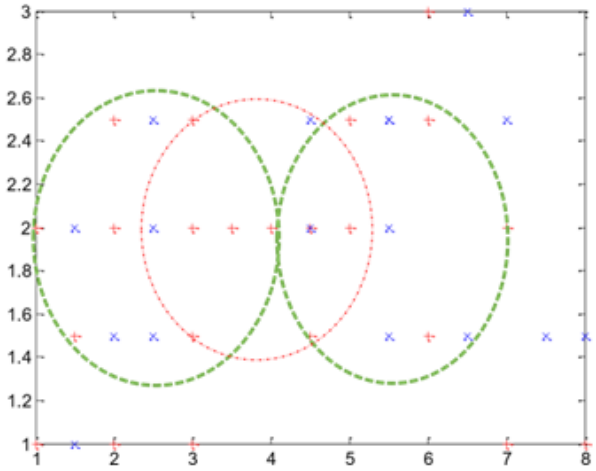

Figure 1. Difficulty to discover a small group into the middle of others.

Results and Discussion

The big job is to treat the information: analyze, search and classify, and after validate the results with correlations, variations and finally determine the maximum entropy to the variables and the results. In particular case of the politic interest finally is expected determine a groups of people defined by coordinates which confine these groups in a parameters segmenting in they different political ideas.

To do that is used the known tools as the clustering by density, there exist hundreds of types, and the correlation methods for validation, but previous work is essential, define segment of data, to configure data in order to find the parameters which determine minimum variables to define a specific group interested in a political party. Give the attributes expressed in words which means classification of the people and transform it into numerical values, representing correctly the graduation of these characteristics in a linear or nonlinear form, and represent different results orpolitical party in a series of numbers ordered by one criteria, is a preliminary job to use a numerical algorithm to find some results of this inquire.

Further job is analyze the results and eliminate one by one the extra dimension which have no relevant information for this founded group in order to determine the minimum number of dimensions maintaining the results, which are at the end the characteristics which defines this group. Different as using virtual variables determined by mathematical algorithms to define perpendicular axis to represent the clusters in a most differentiate expression, PCA form [5].

Proceed to solve the problem is the following diagram:

Loop 1:

- To decide the initial numbers of clusters N to be segmented the data. Apply Cluster function.
- Validate the discovered groups, measuring the false results in the area of the cluster. (result in % OK as efectivity and store the groups and results.)
- Repeat augmenting N+1.

With the optimal results in effectivity (Nop) or the obtained desired custom group apply following loop for desired groups:

Loop 2:

- Use -1 Axis, reduce one dimension,
- Value the results (clustering by Nop), if reduces the effectivity use this axe.
- Follow reducing axes.

As a result is the group desired (by efficiency or confined into desired parameters) defined by the minimal axes.

Example

This example is centered on the interest for the politicians to know where are the possible votes, segmenting by the age, and by some interest, familiar economy etc. and using inquires and polls try to figure out.

The Figure 2 represents the political interest of people of different ages and cities in a thousand habitants. Red is a progressive idea, and blue conservative.

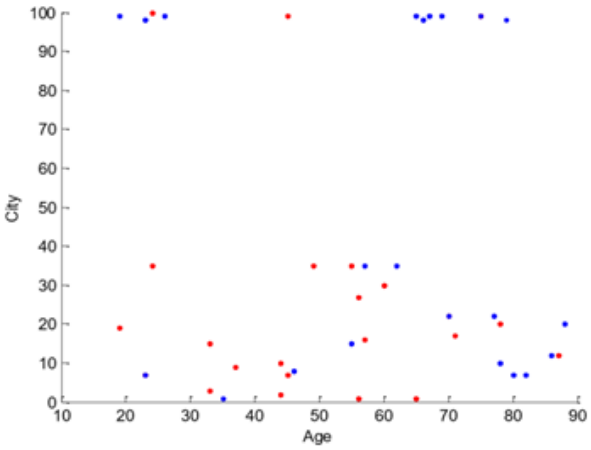

Figure 2. Representation of a people with their interest.

Representation

2D diagram represents easily the interest of the people to their political ideas. Using different colors as parties can be found groups of people with similar interest as groups of colors (Figure 3).

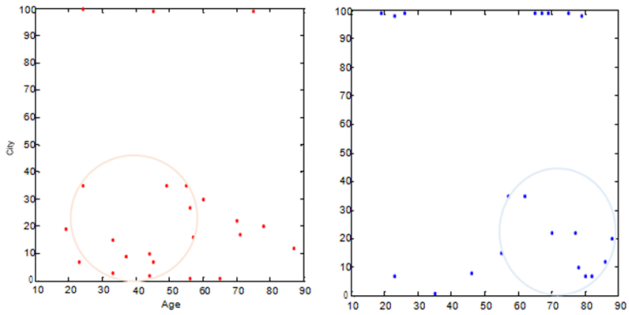

Figure 3. Classification of the people in two groups. With individual graphs the groups are moreidentifiable.

It is possible to change axis values and use the political parties as an axis with numerical value, assigned independently to each party, then the groups are distinguished by the zones not for colors, and using cluster function by density the groups finally are confined by the cluster function parameters.

The problem is when is multi-dimensional system, in such case is no possible a visual recognition the job is done by the cluster function which gives the possible groups, and then is necessary to evaluate these groups to validate the percentage of the elements to be in a desired result of the interest in their political party.

Sometimes are interest a group where the result percentage is divided in two political parties, in order to known the ideas competition for this group.To discover the interest of people is used an application try to be an entertainment moment, destressing interval and the method to be introduced is passing between users as the success of the MEME. In this case is necessary to introduce relaxing images as landscape, or beautiful actress, etc. images will be chosen by users and will define their interest.

Resulting of this proceed is determined a group, doing a poll in a restricted group of people, we find a subgroup of citizens in ages greater than 40 and less of 65 still working, and living in medium cities, resulting the divorced and single ones are more progressive than the others in couple which are more conservative, equal for man's and women's, and independently of the own car, holidays. Useful Information for the oriented publicity or the political message or action will be done to take interest of this defined group.

Robot applications

We work using this system to define the parameters imply to discover the slipping situation in a mobile robot in outdoor scenarios [6], in which the important variables help to warning to the robot a possible dangerous situation in order to prevent the blocking situation. From different measures of the robot, Volts applied to each wheel, I current consumed, velocity for each wheel and the robot inclination, and measuring situations in a slipping movements, in order to reduce the information to the most relevant to be able to proceed the control of the robot in real time, into their displacements to avoiding blocking situations while doing supervision of the street lights in the cities [7].

Using robot in different scenarios the objective is representing data marking clusters where the robot is in dangerous position and the wheels have no complete track on the soil. With these data the objective is to represent and determine the representative axis according the sensors, to identify areas where the robot must to avoid in their movements (Figure 4).

Figure 4. Mobile robot in outdoor scenarios.

Conclusion

Some different situations to prevent and avoid the non-stabilized position, looking into the data is possible to advise the robot is approximated to one identified cluster meaning a dangerous situation, and the control must to react to move the robot againstthis cluster, for example, changing velocity or inclination.

References

1. Ahlquist, John S, and Christian Breunig. "Model-based clustering and typologies in the social sciences." *Polit Anal* (2012): 92-112.
2. Tabachnick, Barbara G, and Linda S Fidell. "Using multivariate statistics. Allyn and Bacon." *Needham Heights, MA* (2001).
3. Esteva Santiago, Figueras Albert, and De la Rosa Josep Llluís. "Artificial Intelligence (AI) Based on Eccentricities". *J Comput Sci Syst Biol* (2020).
4. Release, MATLAB. "The mathworks." *Inc Natick Massachusetts, USA* 488 (2013).
5. Smith, Lindsay I. "A tutorial on principal components analysis." (2002).
6. Figueras, Albert, Santiago Esteva, XavierCufí, and De La Rosa Josep Llluís. "Applying AI to the motion control in robots. A sliding situation." *IFAC-Papers Online* 52 (2019): 393-396.
7. Figueras, Albert, JosepLluís De La Rosa, Santiago Esteva, and Xavier Cufí. "Robot team in the improvement of the neighborhood." In 2018 IEEE International Smart Cities Conference (ISC2). *IEEE* (2018).

How to cite this article: Albert Figueras, Esteva Santiago, and De La Rosa Josep Lluís. "A Study of Using Custom-Clustering Algorithm for a New Treatment of COVID-19". J Comput Sci Syst Biol 13 (2020) doi: 10.37421/jcsb.2020.13.311