Pattern Cognition, Pattern Recognition (CDT-19)

Preprint · December 2019

DOI: 10.13140/RG.2.2.15716.73600/1

1 author:

Luciano da F. Costa
University of São Paulo
734 PUBLICATIONS **13,156** CITATIONS

SEE PROFILE

Pattern *Cognition*, Pattern *Recognition* (CDT-19)

Luciano da Fontoura Costa
luciano@ifsc.usp.br

São Carlos Institute of Physics – DFCM/USP

v.1.1 : 15th May 2020
v.1.0 : 26th Dec. 2019

Abstract

Pattern recognition is much more disseminated and fundamental endeavor than usually realized. Arising from a need to achieve compact, organized and abstract representations of the entities in our world, it underlies not only daily tasks, but also constitutes the basis of language, knowledge, and modeling. In this work we revise the concept of patter recognition, its objectives and challenges, and present some simple but effective means to perform this important task, namely the $k-$neighbors and $k-$means methods.

'Tutta la nostra conoscenza ha origine nelle nostre percezioni.'

Leonardo da Vinci.

1 Introduction

Plants are amazing natural entities. Representing one of the earliest forms of life, plants have been fundamental for the development and maintenance of a myriad of other living forms, including animals and ourselves. Early in life, we learn to identify plants and trees, and progressively develop an ability to distinguish their varying types, and to know the respective names. This recognition involves taking into account several respective features, including the overall size, shape and color of the plants, as well as their parts (e.g. leaves, fruits, seeds, roots, etc). Then, when we observe the leaves of a given plant, we recognize that they share many features and often can hardly be distinguished one another. Our interaction with plants provides one of the most compelling and complete example of the important activity known as *pattern recognition*, which extends to many other types of living beings, objects, and even actions.

Indeed, the action of pattern recognition is much more ubiquitous than often realized. We recognize not only plants, but virtually every entity that can have implications to our lives. Each of the terms of our very *language*, for instance, entails its respective pattern recognition. This is self-evident in the case of substantives, but verbs and other types of words also involve recognition (actions, in the case of verbs). All other human abilities also relate to pattern recognition, including behavior, arts, and scientific modeling (e.g. [1]). Indeed, prediction can also be understood as being intrinsically related to pattern recognition, in the sense that often we make forecasts based on similarities with other previous instances of the phenomenon of interest. In brief, pattern recognition is at the very core of human activity.

Being able to identify the category of a given entity has been centrally important for us, as it allows one to identify if a fruit is edible, if an animal is friendly, if an action is potentially dangerous, among many other possibilities. Thus, in a sense, pattern recognition plays a critical role for living beings.

Figure 1 illustrates the basis of pattern recognition: given two objects, one needs to decide if they are of the same type or belong to different categories.

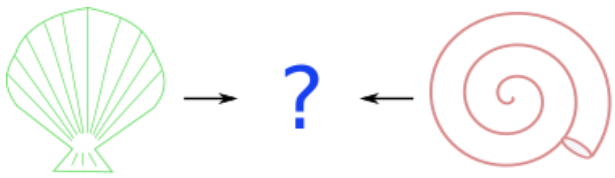

Figure 1: Given two objects, how can we decide if they are of the same type or belong to distinct categories? These issues pertain to the area of *pattern recognition*.

The understanding of the term *pattern recognition* can proceed by taking into account each of its two parts: *'pat-*

tern', which basically corresponds to any entity/object of any nature; and *'recognition'*, a slightly more elaborate concept that can actually be broken into two sub-parts: *'re'*, in the sense of repeating something, and *'cognition'*, which can be interpreted simply as knowing. So, recognition can be ultimately understood as knowing that some specific pattern has recurred. Therefore, cognition is a potential precondition to recognition. One important point here is that the recurrence of a pattern can take place because the same object is presented again, or because two similar objects are presented.

Another key related aspect in pattern recognition is that objects identification involves the definition of *categories* (e.g. dogs, cats, horses, etc., in the case of vertebrates) into which the objects are grouped based on their relative similarity. This is ultimately necessary because we would be otherwise unable to label each possible entity in the real world. For instance, we would need to have names for all possible cats and dogs in this world, and so on. Abstracting similar entities into specific categories, while not undermining critical information characterizing the entities, can save a lot of memory and effort.

The definition of categories, which is one of the important and difficult steps in pattern recognition, should reflect not only the similarity between the considered objects, but also take into account the importance of specific features. For instance, details that can have important consequences but are small, and would therefore imply respective little discriminability, need to be emphasized while defining the respective categories.

Pattern recognition itself can be achieved in two main ways: (i) *supervised*, in the sense that examples/prototypes respective to each category are available; and (ii) *non-supervised*, in which the categories of patterns have to be inferred from scratch. Needless to say, the latter tends to be much more difficult than the former. In addition, there can be no supervised pattern recognition unless it is preceded by some unsupervised definition of categories.

Effective and versatile pattern recognition has represented a substantial challenge in science and technology, as a consequence of the lack of definitive means for selecting features for objects characterization, as well as the frequent existence of overlap between categories. It is therefore hardly surprising that the recent area of *deep learning* (e.g. [2]) has received so much attention. Here, massive data and computing resources, allied with new conceptual insights, is finally paving the way to effective automated pattern recognition as required by a wide range of applications and even theoretical developments. More and more most activities will involve or critically depend on automated pattern recognition, so that it is important to be acquainted with related concepts and methods.

In the present work, we develop a conceptual introduction to some of the main concepts and methods of pattern recognition. We start by discussing the necessary mapping from each entity/object into a respective mathematical, quantitative representation, giving rise to the concept of *feature vectors*. Because this allows one to derive models of the mapped object, allowing its understanding, here we call this preliminary state *pattern cognition*. The next stage in which more than one patterns are compared, grouped and assigned to categories constitutes the pattern recognition proper, which is discussed subsequently. Then, we introduce two of the main types of pattern recognition approaches: supervised, and unsupervised, and briefly present and illustrate simple respective methods, namely $k-$neighbors and $k-$means.

2 Pattern *Cognition*

Often, the entities to be recognized belong to the real-world, including birds, flowers, bikes, pencils, etc. Because the respective recognition takes place in our nervous system (biological pattern recognition), or in computer circuits (automated patter recognition), it is necessary to transfer quantitative information about the object to be recognized into these two types of systems. This can be achieved by the use of biological senses or artificial sensors and equipments (e.g. cameras, eyes, scales, rulers, voltmeters, etc.), which can operate on more than one modality (e.g. image, sound, smell, etc.).

These primary representation received by the pattern recognition system typically involve a large amount of data (e.g. an image containing millions of pixels), much of which tends to be redundant (e.g. the pixels associated to a limpid blue sky). It is therefore interesting to derive a new representation involving a smaller set of data but which retains as much as possible of the relevant information. The set of derived *features* or measurements f_i, $i = 1, 2, \ldots, M$, can be organized into a respective *feature vector* $\vec{f}$. Observe that *data* is not the same as *information*, the latter being more important.

Figure 2 illustrates the mapping of the scene of a real-world object (a sea snail) into a digital image in a computer, which is then transformed into a new, more compact and effective representation $\vec{f}$. In this case, possible features could include width, arc-length of the spiral, color, elongation, etc.

These two mapping stages culminate in a *representation* $\vec{f}$ of the original entity, containing respective relevant information. In a sense, the pattern recognition system can therefore be understood to contain knowledge, as provided by the several measurements in $\vec{f}$, about the real-world object, defining a respective *feature space* whose di-

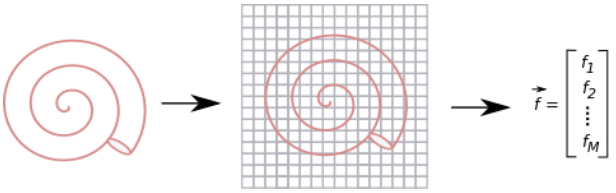

Figure 2: In order to be handled by an animal or computer, an object from the real world needs to be measured into a discrete space. In this particular example, a digital camera is used to transform the real-world scene of a snail into an image by sampling its colors, intensities, and geometry. The obtained image representation usually involves a large amount of data and contains much redundancy, so that more compact representations, capable of keeping as much as possible the original information while removing redundancies, can be subsequently obtained. The final product is the *feature vector* $\vec{f}$.

mension corresponds to the number of adopted features. It is them possible to develop a model of the object, providing some level of respective understanding. Thus, in a sense these mappings can be understood as a *'cognition'* stage, in which knowledge about some relevant properties of the object are derived. However, this system is as yet unable to identify recurrent presentations of this same object, as well as relationships between objects, i.e. it is not yet capable of pattern recognition.

Though conceptually simple, the mapping of an object into the respective feature vector $\vec{f}$ is by no means a straightforward task. The key issue here is to derive a small set of features that would be capable of establishing a bijective relationship between the real-world objects and the respective representations $\vec{f}$. Bijection is particularly important as it allows the inverse mapping to be defined, therefore establishing a unique correspondence between the entities and their respective representations.

If too few features are considered, it is likely that several objects will map into the same representation. For instance, in case we choose as feature the length and width of an object, several objects will map into the same representation $\vec{f}$, implying a non-injective mapping, which is non-invertible. It may also happen that some of the considered objects have not yet been mapped into respective feature vectors $\vec{f}$, implying in a non-surjective mapping, which is also non-invertible.

Invertibility is achieved if and only the map is injective and surjective. Yet, if too many features are selected aimed at achieving an invertible mapping, the representation will tend to become redundant and make it very difficult to populate and handle the respective feature space with samples of objects. Regarding redundancy, it is interesting to observe that, given a relatively simple entity, there is not much that can be said about it without engaging repetition.

Finding an invertible representation is typically a real challenge, as the choice is heavily influenced by the ability of each feature to convey information about specific properties of the object, as well as the existing types of objects and their distinguishing characteristics. As a second best alternative, one should aim at obtaining a as small as possible set of features, capable of providing enough information so as to solve a given pattern recognition problem with suitable performance.

Previous knowledge and familiarization with several types of features, the existing objects to be classified, as well as with a variety of pattern recognition methods represent some of the possible methodologies that can be used to select a suitable set of features. It is also possible to consider multivariate statistical methods such as principal component analysis – PCA (e.g. [3]) as a means to reduce the number of features through elimination of correlations between features, as well as automate feature selection methods.

All in all, the mapping of a single pattern into respective measurements allows this entity to be known in an objective, quantitative way, so that eventually respective models can be derived that not only help explaining how the entity behaves, but also allow predictions to be made about it. Though these achievements exhibit a high level of cognition, we still lack knowledge about how the mapped object is related to other co-existing entities. Deriving these relationships is a task well within the pattern *recognition* realm.

3 Pattern *Recognition*

Once an object has been mapped into respective feature vectores $\vec{f}$, it occupies a specific position in the associated feature space. More specifically, each object is mapped into the respective point given by vector $\vec{f}$ in the feature space. Because the components of these vectors correspond to features, which are measurements of the object properties, two similar objects tend to be mapped into two corresponding vectors that are also similar, implying the respective points in the feature space to be relatively near one another. The proximity between these two points can be quantified in terms of some distance measurement (observe that there is a virtually infinite number of possible distance definitions), such as the Euclidean distance between the $M-$dimensional respective feature vectors $\vec{f}_1$ and $\vec{f}_2$, i.e.

$$\delta_{1,2} = \sqrt{\sum_{i=1}^{M} (f_i^{(1)} - f_i^{(2)})^2} \qquad (1)$$

One important point that needs to be taken into account is that the concept of *distance* between patterns is a relative issue. Indeed, in case we have only two patterns,

in the absence of some preliminary imposed criterion, the Euclidean distance between them cannot be taken as being near or far. This can be realized by considering that we lack a reference for comparison, and that even the choice of the system of units can be arbitrary (e.g. inches or meters, etc.).

We need at least a third pattern, so that three distances can be now calculated, namely $\delta_{1,2}$, $\delta_{1,3}$ and $\delta_{2,3}$. The smallest of these will correspond to the most similar pair of objects, and also to derive proportions of similarity by considering ratios between the above distances.

Almost invariably, the entities of a same category will exhibit variations intrinsic to their conception (e.g. genetic content) as well as their history. For instance, a clam may have lived under harsher conditions, implying wearing of its shell. When such entities are transformed into respective feature vectors $\vec{f_i}$ and mapped into the respective feature space, they tend to give rise to a *cluster* of points. Figure 3 illustrates this tendency with respect to a hypothetical set of clam shells.

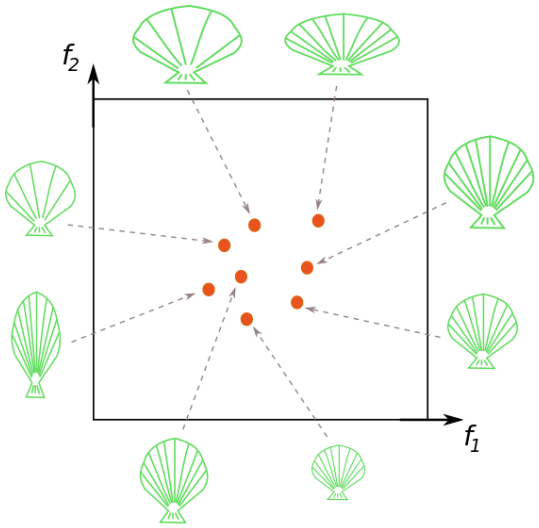

Figure 3: A set of hypothetical clam shells mapped, through some respective measurements f_1 and f_2, into a corresponding two dimensional feature space. The fact that the clams are similar but not identical implies in the formation of a cluster of points possessing a respective spatial dispersion. Total superimposition into a single point would have been obtained in case all clams were identical. Observe that the choice of other features would very likely yield different distribution of points in the feature space.

For analogous reasons, objects that are substantially different one another will tend to be mapped into distant respective points in the feature space. Now, if we have two groups of objects so that all objects in a same group are similar one another while objects from distinct groups are substantially different, we will obtain two well-separated clusters, corresponding to two respective distinct *categories*. Figure 4 illustrates this tendency with respect to hypothetical groups of clam shells and sea snails.

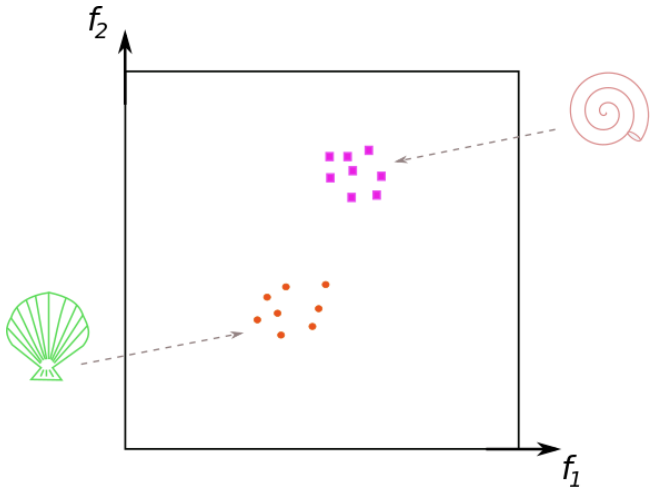

Figure 4: When two groups of individuals (clam shells and sea snails) that are similar within each group, but different between the groups are mapped into a two-dimensional feature space, two respective well-defined clusters appear, defining respective categories.

The simple principle described above lies at the very core of pattern recognition, which can therefore be understood as the endeavor of *finding and separating clusters in feature spaces* and assigning categories to given objects.

As with the task of deriving effective object representations, identifying cluster corresponds to another challenge, especially in case the involved clusters are not well-separated, which often turns happen. In addition, clusters can appear in several shapes and sizes, which further complicates their identification and separation. These difficulties have motivated the development of a large number of pattern recognition methods.

Generally speaking, categories can be more effectively identified when the individuals within each group are markedly similar, implying small *intra-cluster scattering*, while exhibiting strongly distinct properties between the groups, accounting for large *inter-cluster scattering*. Otherwise, there could be *overlapping* of points, i.e. individuals of one group could be found inside other groups.

We conclude this section by observing that the obtention of well-separated clusters, and therefore categories, in feature spaces depends critically on: (1) the entities of interest initially having intrinsic differentiating characteristics; (2) deriving pattern representations incorporating measurements that capture these variations; and (3) choosing effective features. Point (1) is particularly important because, in case the groups of patterns do not exhibit any reasonable intrinsic differences, objective assignment of categories is from the scratch doomed to failure.

4 Supervised Pattern Recognition

Though supervised pattern recognition can only be performed in case the respective categories have already been defined, we cover this type of pattern recognition first because it is generally simpler and easier than the non-supervised counterpart.

As implied by its own name, supervised patter recognition, or supervised classification, involves some previous knowledge about the groups and their properties. This knowledge can be of various types. For instance, we can have access to several objects whose category is know, so that they can be used in a training stage. Or, we can have prototypes of each class, and perhaps also know the intra-class variation of each group. Or yet, we may know statistical properties of each group. The basic idea is, starting from such prior information, to devise some means of identifying the category of new entities. So, supervised pattern recognition typically involves two subsequent stages: (a) training; and (b) recognition of new objects.

Among the large number of supervised approaches that have been developed (e.g. [4]), for simplicity's sake we only cover one that, despite being very simple, can often lead to good performance. We are speaking of the *nearest-neighbor* method (e.g. [4]). Given a set of objects with known respective categories, they can be mapped into feature vectors as previously discussed, defining a reference population of points in the feature space. As an example, we can consider the case in Figure 4, where hypothetical sets of clam shells sea snails were mapped into a t two-dimensional space by using some arbitrary measurements. We repeat that distribution of points in Figure 5 for the sake of convenience.

Now, given a new object of unknown category, we map it into the same feature space (e.g. the green or blue stars in Fig. 5), determine the k nearest points and take as the most likely category that having majority among the $k-$nearest known objects. Let's henceforth consider $k = 1$. In the case of the object mapped into the green star, we have that the nearest neighbor is a clam, so we take this as its most likely category. Observe that the same category would have been obtained even for larger values of k. Now, in the case of the object mapped into the blue star, it is sightly closer to a sea snail, which would be the result for $k = 1$. However, consideration of larger values of k would imply instability of this result.

This method extends immediately to larger number of groups and higher dimensional feature spaces. Also, when a large set of examples are provided during training, this method tends to minimize the chances of getting wrong classification results.

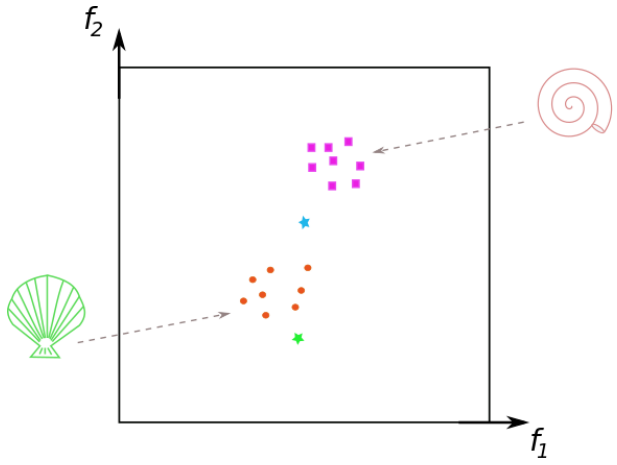

Figure 5: Two hypothetical groups of sea snails and clam shells have been mapped into a two-dimensional feature space during the training stage of the nearest-neighbor method. A new object of unknown category mapped into the green star can be understood as being a clam for several values of k. However, the most likely category for the object mapped into the blue star tends to be unstable.

5 The Confusion Matrix

There are several approaches to quantify the performance of a classifier, including concepts such as false positives and false negatives, ROC curves, etc.

A particularly straightforward approach consists in considering the so-called *confusion matrix*. Give a supervised classification problem with k categories, the respective confusion matrix has dimension $k \times k$. Confusion matrices are obtained by separating the training set into two (or more) groups, using one of them during traning and the other as a gold-standard (or ground-truth) to test the performance. An example of confusion matrix is as follows:

$$C = \begin{bmatrix} 45 & 10 & 15 \\ 5 & 50 & 9 \\ 0 & 5 & 90 \end{bmatrix} \tag{2}$$

Each element $C[i, j]$ indicates the number of cases which were assigned the label j while the original label was i (it is possible also to adopt the opposite interpretation). For instance, the fact that $C[2, 3] = 9$ in the above example indicates that 9 objects in the original category 2 were classified as belonging to category 3.

Observe that the sum along each line corresponds to the number of objects in that respective original category.

Ideally, the confusion matrix should be a diagonal matrix, therefor indicating that not misclassifications have taken place.

6 A Brief Digression

Though the above approach to clustering involving compact groups that are far away one another corresponds to the most often adopted approach, we would like to briefly discuss a complementary perspective (e.g. [5]). Consider the situation depicted in Figure 6.

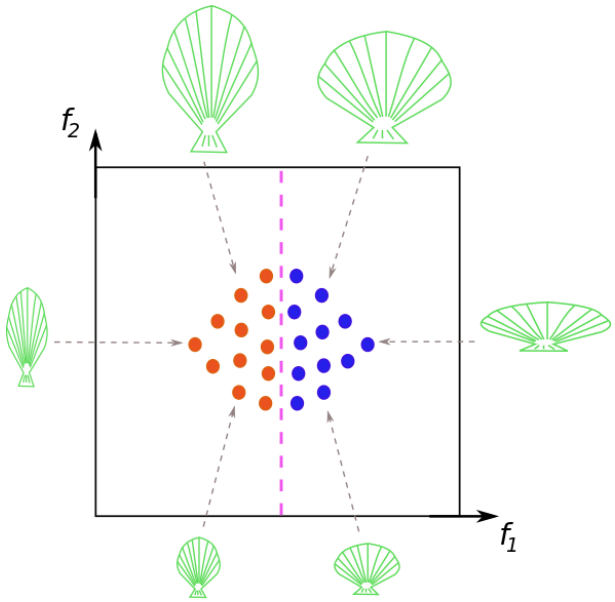

Figure 6: A situation in which two groups give rise to a single cluster, but in such a way that they can still be properly separated, e.g. by considering the linear decision boundary shown in magenta.

In this hypothetical case, we have two species of clams: one characterized by being taller than wider, and the other vice-versa. The individuals from each species are also assumed to appear with different diameters (i.e. the largest distance between any possible pair of its points). In case we choose f_1 to correspond to the ratio between the width and height of the shells, and f_2 to coincide with the their diameters, the distribution of points in the feature space could be similar to that shown in Figure 6.

Interestingly, though none of the two species give rise to a cluster, characterized by a compact, well-separated group of points, the two groups nevertheless resulted organized and separated in the feature space (no overlaps). Indeed, it is possible to perfectly separate the two groups by considering the linear boundary (also called *separatrix*) shown in magenta in Figure 6, and we can say that the two groups are linearly separable, as indeed also are the groups in Figure 4. However, we will not refer to these separable groups as clusters.

It is interesting to compare the two situations shown in Figure 4 and 6. Though the two groups are linearly separable in both cases, in case we did not know *a priori* that there are two groups, we would not be able to find them from the situation in Figure 6, while the separation would be obvious in the other case. On the other hand, the situation in Figure 6 has the potential for allowing enhanced resolution and discriminability between the points in each group, as a consequence of each group occupying a relatively larger area than it would be otherwise obtained in the case of two compact and distant groups.

So, in case we have *a priori* indication about the properties of the categories, as it happens in the training stage of supervised learning, it is possible to sacrifice the separability of the groups for enhanced resolution and characterization of objects inside each group. Observe that this approach would reduces the margin of error in the objects mapping, and could therefore result in overlap. In summary, the above discussion emphasizes that, given the limited resolution of measurements and feature space representations (especially in the case of discrete measurements), there exists a trade-off between margin of classification error and enhanced resolution. The decision of which should take priority depends critically on each specific problem and should be exerted with special care.

7 Unsupervised Pattern Recognition

Now, we proceed to the more challenging problem of unsupervised classification, sometimes also called *clustering*: nothing is known about the categories or their statistics properties, and we only have access to a set of observed feature vectors defining respective distribution of points in a feature space defined by the chosen measurements. An example can be immediately derived from Figure 4 by disregarding the respective known categories, giving rise to the distribution of undistinguishably marked points shown in Figure 7.

Though belonging to the 'easy' type of unsupervised classification, we can use this situation to introduce one of the simplest and best known clustering methods, namely $k-$means (e.g. [6]).

First, we have to choose how many categories will be taken into account. In this case, the obvious choice is $k = 2$, but other situations involving less separated groups can be more challenging. Then, we need two respective *'seeds'*, around which the respective categories can be organized. These can be either drawn randomly, or we may have some intuition or insight about the possible location of the clusters.

The $k-$means algorithm then identifies which of the feature space points are closer to each of the seeds, defining two preliminaries partitions of the feature space. The centers of mass of the points inside each of these partitions are then calculated and understood as new respective seeds. The method proceeds until the position of the

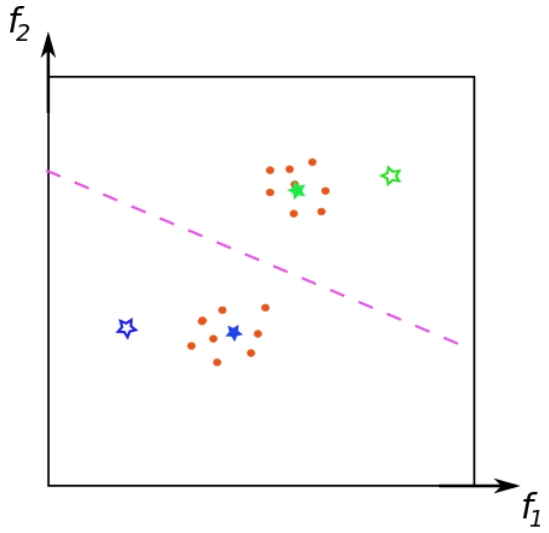

Figure 7: A set of points (orange discs) are to be separated into two categories by using the $k-$means algorithm. Two respective seeds are randomly assigned (e.g. the two empty stars), and the points closer to each of them are successively identified, while their centers of mass are then taken as new corresponding seeds. The method is terminated when the positions of the seeds become stable. A properly converging result is shown in terms of the two seeds represented as the blue and green stars, which define a respective Voronoi partitioning of the feature space, so that the border between the two Voronoi cells constitues a linear separatrix.

seeds become stable, which tends to happen in just a few steps. Hopefully, the resulting seeds will correspond to the sough categories, spanning respective regions so that each point inside these regions are closer to the corresponding seed, defining a Voronoi partitioning of the feature space. Figure 7 illustrates the two obtained categories. Observe that the Voronoi partition in this case presents a linear separation border.

Though relatively simple, this method often provides good performance. The main difficulty is that the positions of the original seeds can have a strong effect on the final result. For instance, two seeds may converge to the same cluster. One possibility to improve the result is to perform the algorithm several times with respect to many initial seeds positions, and to disregard the cases resulting in seeds superimposition. Another limitation is that the number of categories needs to be properly estimated.

8 Concluding Remarks

As a means of understanding and communicating about the world, human beings have mapped the several of its entities into abstract concepts or categories, which ultimately played an important role on the development of language and science. In the present work, we discussed how the task of pattern recognition has been fundamental in this process.

We have seen that pattern recognition typically involves two subsequent stages. First, each of the entities to be recognized is transformed into a quantitative representation in terms of a feature vector, a stage that we called *pattern cognition*. Then, the feature vectors corresponding to several objects are analyzed in order to identify their respective categories, constituting the patter recognition proper. This can proceed in two main manners, one proceeding from scratch, without any information about the possible categories or their statistics (unsupervised pattern recognition), while the other is based on a training stage involving samples or prototypes with known categories. We have described two simple but relatively effective respective methods, namely $k-$neighbors (supervised) and $k-$means (unsupervised).

The reader should keep in mind that, as a consequence of the central importance of pattern recognition in science and technology, many other interesting and effective concepts and methods exist, including the application of principal component analysis (e.g. [3]) as a means to reduce the dimensionality of the feature space, as well as a large number of alternative interesting methods for supervised and unsupervised pattern recognition.

Acknowledgments.

Luciano da F. Costa thanks CNPq (grant no. 307085/2018-0) for sponsorship. This work has benefited from FAPESP grant 15/22308-2.

References

[1] L. da F. Costa. Modeling: The human approach to science. Researchgate, 2019. https://www.researchgate.net/publication/333389500_Modeling_The_Human_Approach_to_Science_CDT-8. Online; accessed 03-June-2019.

[2] H. F. de Arruda, A. Benatti, C. H. Comin, and L. da F. Costa. Learning deep learning. Researchgate, 2019. https://www.researchgate.net/publication/335798012_Learning_Deep_Learning_CDT-15. Online; accessed 22-Dec-2019.

[3] F. Gewers, G. R. Ferreira, H. F. Arruda, F. N. Silva, C. H. Comin, D. R. Amancio, and L. da F. Costa. Principal component analysis: A natural approach to data exploration. Researchgate, 2019. https://www.researchgate.net/publication/324454887_Principal_Component_Analysis_A_Natural_Approach_to_Data_Exploration. Online; accessed 25-Dec-2019.

[4] D. R. Amancio, C. H. Comin, D. Casanova, G. Travieso, O. M. Bruno, F. A. Rodrigues, and L. da F. Costa. A systematic comparison of supervised classifiers. *PLOS One*, 2014.

[5] C. H. Comin, F. Nascimento, and L. da F. Costa. A framework for evaluating complex networks measurements. *Europhysics Letters*, 110:68002, 2015.

[6] M. Z. Rodriguez, C. H. Comin, D. Casanova, O. M. Bruno, D. R. Amancio, L. da F. Costa, and F. A. Rodrigues. Clustering algorithms: A comparative approach. *PLOS One*, 14, 2019.

Costa's Didactic Texts – CDTs

CDTs intend to be a halfway point between a formal scientific article and a dissemination text in the sense that they: (i) explain and illustrate concepts in a more informal, graphical and accessible way than the typical scientific article; and (ii) provide more in-depth mathematical developments than a more traditional dissemination work.

It is hoped that CDTs can also integrate new insights and analogies concerning the reported concepts and methods. We hope these characteristics will contribute to making CDTs interesting both to beginners as well as to more senior researchers.

Each CDT focuses on a limited set of interrelated concepts. Though attempting to be relatively self-contained, CDTs also aim at being relatively short. Links to related material are provided in order to complement the covered subjects.

Observe that CDTs, which come with absolutely no warranty, are non distributable and for non-commercial use only.

The complete set of CDTs can be found at: `https://www.researchgate.net/project/Costas-Didactic-Texts-CDTs`.