

Article

A Two-Parameter Fractional Tsallis Decision Tree

Jazmín S. De la Cruz-García ^{1,†}, Juan Bory-Reyes ²  and Aldo Ramirez-Arellano ^{1,*,†} 

¹ SEPI-UPIICSA, Instituto Politécnico Nacional, Mexico City C.P. 08400, Mexico; susdcg@gmail.com

² SEPI-ESIME-ZACATENCO, Instituto Politécnico Nacional, Mexico City C.P. 07738, Mexico; juanboryreyes@yahoo.com

* Correspondence: aramirezar@ipn.mx; Tel.: +52-5528053125

† These authors contributed equally to this work.

Abstract: Decision trees are decision support data mining tools that create, as the name suggests, a tree-like model. The classical C4.5 decision tree, based on the Shannon entropy, is a simple algorithm to calculate the gain ratio and then split the attributes based on this entropy measure. Tsallis and Renyi entropies (instead of Shannon) can be employed to generate a decision tree with better results. In practice, the entropic index parameter of these entropies is tuned to outperform the classical decision trees. However, this process is carried out by testing a range of values for a given database, which is time-consuming and unfeasible for massive data. This paper introduces a decision tree based on a two-parameter fractional Tsallis entropy. We propose a constructionist approach to the representation of databases as complex networks that enable us an efficient computation of the parameters of this entropy using the box-covering algorithm and renormalization of the complex network. The experimental results support the conclusion that the two-parameter fractional Tsallis entropy is a more sensitive measure than parametric Renyi, Tsallis, and Gini index precedents for a decision tree classifier.

Keywords: decision trees; complex networks; two-parameter Tsallis entropy; Gini index



Citation: De la Cruz-García, J.S.; Bory-Reyes, J.; Ramirez-Arellano, A. A Two-Parameter Fractional Tsallis Decision Tree. *Entropy* **2022**, *24*, 572. <https://doi.org/10.3390/e24050572>

Academic Editor: Boštjan Brumen

Received: 28 February 2022

Accepted: 15 April 2022

Published: 19 April 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Entropy is a measure of the unpredictability of the state in physical systems that would be needed to specify the degree of disorder in full micro-structure of them. Claude Elwood Shannon [1] defined a measure of entropy to measure the amount of information in a digital system in the context of theory communication that has been applied in a variety of fields such as information theory, complex networks, and data mining techniques.

The most widely used form of the Shannon entropy is given by

$$I = \lim_{t \rightarrow -1} \frac{d}{dt} \sum_{i=1}^N p_i^{-t} = - \sum_{i=1}^N p_i \ln p_i, \quad (1)$$

where N is the number of possibilities p_i and $\sum_{k=1}^n p_k = 1$.

Two celebrated generalizations of Shannon entropy are Renyi [2] and Tsallis entropies [3]. Alfred Renyi proposed a universal formula to define a family of entropy measures given by the expression [2]

$$I_q^R = \frac{1}{1-q} \sum_{i=1}^N p_i^q, \quad (2)$$

where q denotes the order of moments.

Constantino Tsallis proposed the q -logarithm defined by

$$\ln_q(x) = \frac{x^{1-q} - 1}{1-q}, \quad (3)$$

to introduce a physical entropy given by [3]

$$I_q^T = - \sum_{i=1}^N p_i \ln_q p_i = \frac{1}{q-1} (1 - \sum_{i=1}^N p_i^q). \tag{4}$$

Tsallis entropy could be rewritten [4–6] as

$$I_q^T = \lim_{t \rightarrow -1} D_q^t \sum_{i=1}^N p_i^{-t}, \tag{5}$$

where D_q^t of a function f given by $D_q^t f(t) = \frac{f(qt) - f(t)}{(q-1)t}$, $t \neq 0$, stands for the Jackson [7] q -derivative, to reflect that it is an extension of Shannon entropy.

Renyi and Tsallis entropy measures depend on the parameter q , which describes their deviations from the standard Shannon entropy. Both entropies converge to Shannon entropy in the limit $q \rightarrow 1$. For complex network applications [8] and data mining techniques [9–17], the parameter q varies into a range of values. On the other hand, the computation of the entropic index q of the Tsallis entropy was implemented for physics applications in [18–25].

Shannon and Tsallis entropies can be obtained by the action of standard derivative or q -derivative, respectively, to the same generating function $\sum_{i=1}^N p_i^{-t}$ with respect to variable t and then letting $t \rightarrow -1$. This approach can be used to reveal different entropy measures based on the actions of appropriate fractional order differentiation operators [26–32].

The major goal of this work is to introduce a new decision tree based on a two-parameter fractional Tsallis entropy. This new kind of tree is tested on twelve databases for a classification task. The structure of the paper is as follows. Section 2 focuses attention on the notion of two-parameter fractional Tsallis entropy. In Section 3, two-parameter fractional Tsallis decision trees and a constructionist approach to the representation of the databases as a complex network are introduced. The basic facts on the box-covering algorithm of a complex network are reviewed. Finally, we compute an approximation set of parameters q , β , and α of the two-parameter fractional Tsallis entropy. Section 4 is concerned with the testing of two-parameter fractional Tsallis decision trees on twelve databases. Next, the approximations of q values are tested on Renyi and Tsallis entropies. Discussion of the findings of this study and concluding remarks are offered in Section 5.

2. Two-Parameter Fractional Tsallis Entropy

Based on the actions of fractional order differentiation operators, several entropy measures of fractional order are introduced in [26–30,33–43]. Following this approach in [22], the two-parameter fractional Tsallis entropy is introduced by merging two typical examples of fractional entropies.

The first fractional entropy of order $\delta \in (0, 1]$ is introduced as

$$I_\delta^1 = \lim_{t \rightarrow -1} \frac{d^\delta}{dt} \sum_{i=1}^N p_i^{-t} = - \sum_{i=1}^N p_i^\delta \ln p_i, \tag{6}$$

and the second one by

$$I_\delta^2 = \lim_{t \rightarrow -1} \frac{d}{dt} (-\infty D_t^{\delta-1} \sum_{i=1}^N e^{-t \ln p_i}) = \sum_{i=1}^N p_i (-\ln p_i)^\delta, \text{ where} \tag{7}$$

$$\frac{d^\delta f(t)}{dt} = \lim_{h \rightarrow 0} \frac{f^\delta(t+h) - f^\delta(t)}{(t+h)^\delta - h^\delta} \tag{8}$$

and

$${}_{-\infty}D_t^{\delta-1}f(t) = \frac{1}{\Gamma(1-\delta)} \int_{-\infty}^t \frac{f(t')}{(t-t')^\delta} dt', \tag{9}$$

where Γ denotes the gamma function.

Combining (6) with (7) yields a two-parameter fractional relative entropy as follows [31]:

$$I_\beta^\alpha = \sum_{i=1}^N p_i^\alpha (-\ln p_i)^\beta, \tag{10}$$

for $0 < \alpha, \beta$. The entropy (10) reduces to (6) when $\beta \rightarrow 1$ and reduces to (7) when $\alpha \rightarrow 1$, yielding the Shannon entropy when both parameters approach 1.

Analogously, two extra-parameter-dependent Tsallis entropies are introduced [22]:

$$T_{\delta,q}^1 = - \sum_{i=1}^N p_i^\delta \ln_q p_i. \tag{11}$$

and

$$T_{\delta,q}^2 = \sum_{i=1}^N p_i (-\ln_q p_i)^\delta. \tag{12}$$

Combining these entropies and motivated by (10), we obtain the following two-parameter fractional Tsallis entropy [22]:

$$T_q^{\alpha,\beta} = \sum_{i=1}^N p_i^\alpha (-\ln_q p_i)^\beta, \tag{13}$$

for $0 < \alpha, \beta$.

Note that Tsallis entropy is recovered when $\lim_{\alpha,\beta \rightarrow 1} T_q^{\alpha,\beta}$. This implies that the non-extensibility of T_q [44] forces $T_q^{\alpha,\beta}$ to be so.

3. Parametric Decision Trees

A decision tree is a supervised data mining technique that creates a tree-like structure, where the non-leaf node tests a given attribute [45]. The outcome gives us the path to reach a leaf node, where the classification label is found. For example, let $(x = 3, y = 1)$ be a tuple to be classified by the decision tree of Figure 1. If we test $x = 1$, we must follow the left path to reach $y = 1$ and finally arrive at the leaf node with the classification label “a”.

In general, the cornerstone of the construction process of decision trees is the evaluation of all attributes to find the best node and the best split condition on this node to classify the tuple with the lower error rate. This evaluation is carried out by information gain on each attribute a [45]:

$$G_a = I(D) - I_c(D), \tag{14}$$

where $I(D)$ is the entropy of the database after being partitioned by the condition c of a given attribute a and $I_c(D)$ is the entropy induced by c . The tree’s construction needs to evaluate several partition conditions c on all attributes of the database, then chooses the pair of attribute–condition with the highest value. Once a pair is chosen, the process evaluates the partitioned database recursively using a different attribute–condition. The reader is referred to [45] for details on decision tree construction and computation of (14).

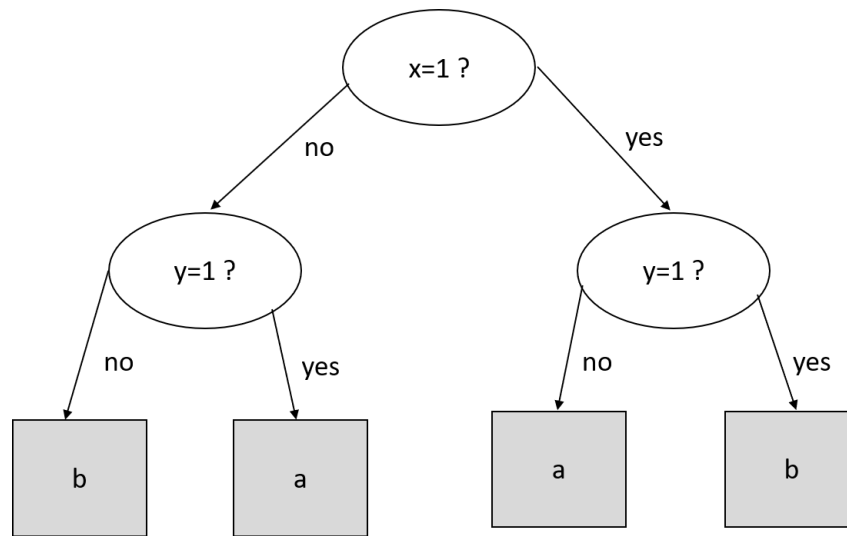


Figure 1. A decision tree to the classification task.

3.1. Renyi and Tsallis Decision Trees

In classical decision trees, I in (14) denotes Shannon entropy; however, other entropies such as Renyi or Tsallis can replace it. Thus, (14) can be written using Renyi entropy (2) as

$$G_{R,a} = I_q^R(D) - I_{q,c}^R(D), \tag{15}$$

and using Tsallis entropy (4) as follows:

$$G_{T,a} = I_q^T(D) - I_{q,c}^T(D). \tag{16}$$

The parametric decision trees generated by (15) or (16) have been studied in [9–14].

3.2. Two-Parameter Fractional Tsallis Decision Tree

Following a similar fashion, a two-parameter fractional decision tree can be induced by the information gain obtained by rewritten (14) using (13):

$$G_{T_q^{\alpha,\beta},a} = T_q^{\alpha,\beta}(D) - T_{q,c}^{\alpha,\beta}(D), \tag{17}$$

An alternative informativeness measure for constructing decision trees is the Gini index, or Gini coefficient, which is calculated by

$$Gini = 1 - \sum_{i=1}^N p_i^2. \tag{18}$$

The Gini index can be deduced from Tsallis entropy (4) using $q = 2$ [14]. On the other hand, the two-parameter fractional Tsallis entropy with $q = 2, \alpha = 1, \beta = 1$ reduces to the Gini index. Hence, Gini decision trees are a particular case of both Tsallis and two-parameter fractional Tsallis trees.

The main issue with Renyi and Tsallis decision trees is the estimation of q -value to obtain a better classification than the one produced by the classical decision trees. Trial and error is the accepted approach for this purpose. It consists of testing several values in a given interval, usually $[-10, 10]$, and comparing the classification rates. This approach becomes unfeasible in two-parameter fractional Tsallis decision trees as it is needed to tune $q, \alpha,$ and β . A representation of a database as a complex network is introduced to face this issue. This representation lets us compute α and β following the approach in [22], which is the basis for determining the fractional decision tree parameters.

3.3. Network's Construction

A network is a powerful tool to model the relationships among entities or parts of a system. When those relationships are complex, i.e., properties that cannot be found by examining single components, something emerges that is called a complex network. Thus, networks as a skeleton of complex systems [46] have attracted considerable attention in different areas of science [47–51]. Following this approach, a representation of the relationships among attributes (system entities) of a database (system) as a network is obtained.

The attribute's name will be concatenated before the value of a given row to distinguish the same value that might appear on different attributes. Consider the first record of the database shown on the top of Figure 2. The first node will be *NAME.BruceDickinson*, the second node will be *PHONE.54-76-90*, and the third node will be *ZIP.08510*. These nodes belong to the same record, so they must be connected; see dotted lines of the network in the middle of Figure 2. We next consider the second record; the nodes *Name.MichaelKiske* and *PHONE.87-34-67* will be added to the network. Note that the node *ZIP.08510* was added in the previous step. We may now add the links between these three nodes. This procedure is repeated for each record in the database.

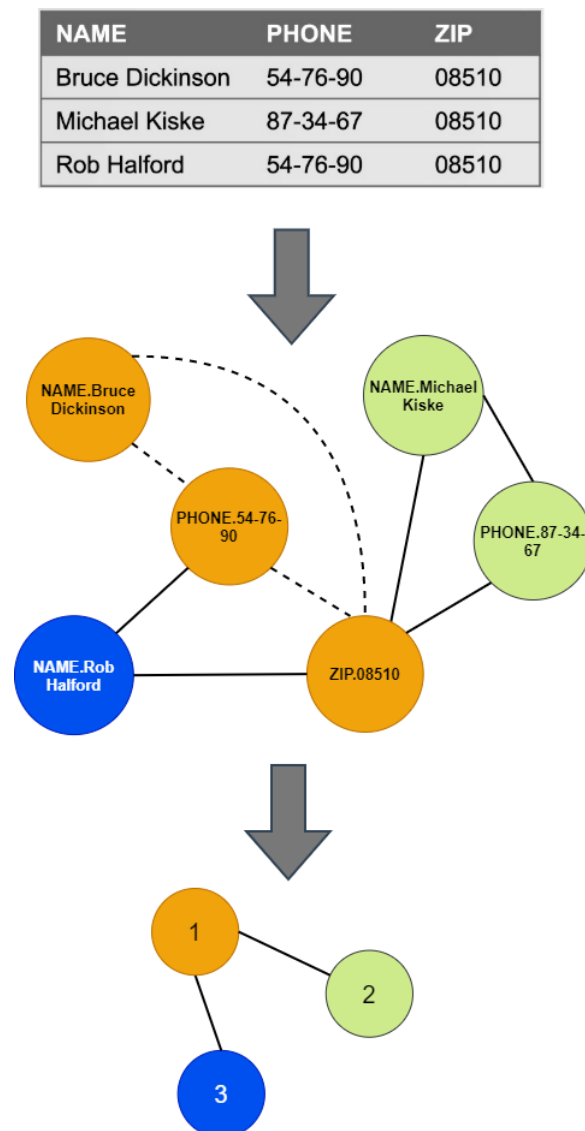


Figure 2. Network construction from a database. The nodes in the same color belong to the same box for $l = 2$.

The outcome is a complex network that exhibits non-simple topological features [52], which cannot be predicted by analyzing single nodes as occurs in random graphs or lattices [53].

3.4. Computation of Two-Parameter Fractional Tsallis Decision Tree Parameters

By the technique introduced in [22], the parameters α and β —on the network representation of the database—of the two-parameter fractional Tsallis decision tree are defined to be

$$\alpha_{l,i}^1 = 1 + \frac{|G_i| \text{innerdeg}(G_i)}{n \sum_{i=1}^{N_b} \text{innerdeg}(G_i)}, \tag{19}$$

$$\alpha_{l,i}^2 = 1 - \frac{|G_i| \text{innerdeg}(G_i)}{n \sum_{i=1}^{N_b} \text{innerdeg}(G_i)}, \tag{20}$$

where $|G_i|$ is the number of nodes in the box G_i obtained by the box-covering algorithm [54], n is the number of nodes of the network, and $\text{innerdeg}(G_i)$ is the average degree of the nodes of the G_i box. Similarly, two values of β are computed as follows [22]:

$$\beta_{l,i}^1 = 1 + \frac{\text{outerdeg}(G_i)l}{\Delta \sum_{i=1}^{N_b} \text{outerdeg}(G_i)}, \tag{21}$$

$$\beta_{l,i}^2 = 1 - \frac{\text{outerdeg}(G_i)l}{\Delta \sum_{i=1}^{N_b} \text{outerdeg}(G_i)}, \tag{22}$$

where $l > 2$ is the diameter of the box G_i , Δ is the diameter of the network, and $\text{outerdeg}(G_i)$ is the number of links among the boxes G_i . The computation of innerdeg and outerdeg will be explained later.

Inspired by the right-hand term of (19) and (20) (named α') with the fact that $\delta = \frac{lN_b(l)}{\Delta \sum_{l=2}^{\Delta} N_b(l)}$ is a normalized measure of the number of boxes to cover the network [20], an approximation of the q -value for the two-parameter fractional decision tree is introduced:

$$q_{\alpha'} = \frac{\delta}{\alpha'} = \frac{n l N_b(l) \sum_{i=1}^{N_b} \text{innerdeg}(G_i)}{\Delta |G_i| \text{innerdeg}(G_i) \sum_{l=2}^{\Delta} N_b(l)}. \tag{23}$$

Similarly, from the right hand of (21) and (22) (named β'), a second approximation of the q -value is given by

$$q_{\beta'} = \frac{\delta}{\beta'} = \frac{N_b(l) \sum_{i=1}^{N_b} \text{outerdeg}(G_i)}{\text{outerdeg}(G_i) \sum_{l=2}^{\Delta} N_b(l)}, \tag{24}$$

where $N_b(l)$ is the minimum number of boxes of diameter l to cover the network, n , Δ , $N_b(l)$, $|G_i|$.

The process to compute the minimum number of boxes N_b of diameter l to cover the network G is shown in Figure 3. A dual network (G') is created only with the nodes of the original network, Figure 3b. Then, the links in G' are added following the rule: two nodes i, j , in the dual network, are connected if the distance between l_{ij} is greater than or equal to l . In our example, $l = 3$, and node one is selected to start. Node one will be connected in G' with nodes five and six since their distance is four and three. The procedure is repeated with the remaining nodes to obtain the dual network shown in Figure 3b. Next, the nodes will be colored as follows: two directly connected nodes in G' must not have the same color. Finally, the nodes colored in G' are mapped to G ; see Figure 3c. The minimum number of boxes to cover the network given l equals the number of colors in G . In addition, the nodes in the same color belong to the same box. In practice, $l = [1, \Delta]$; thus, $N_b(l)$ of the example are shown in Table 1. For details of the box-covering algorithm, the reader is referred to [54].

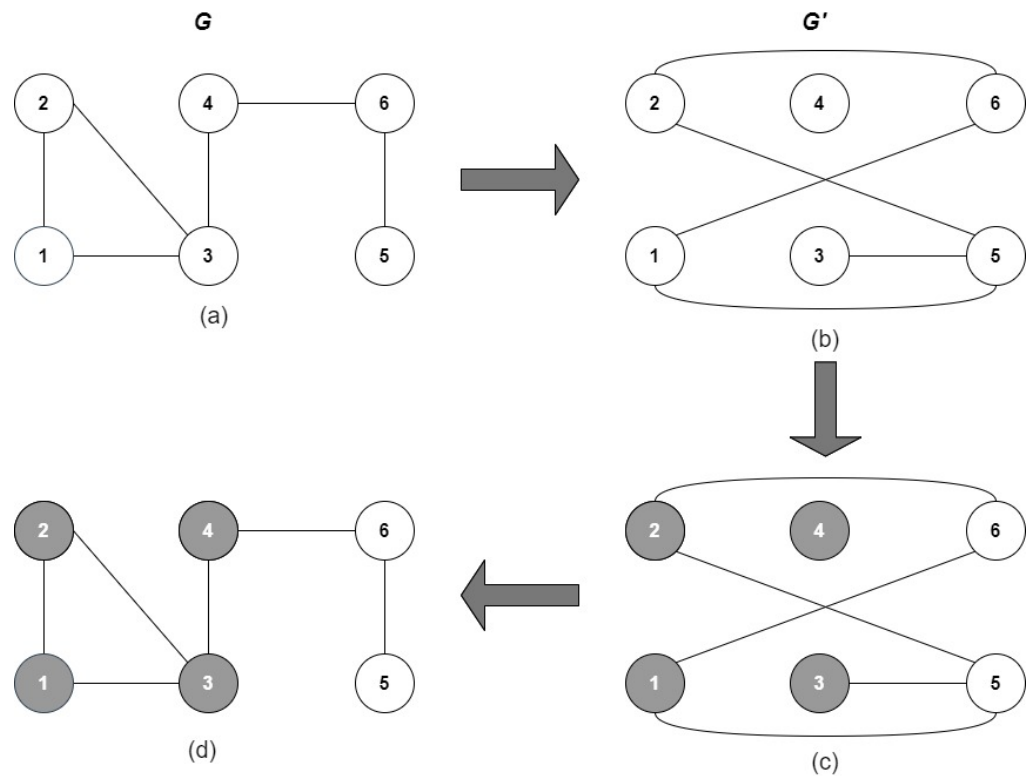


Figure 3. Box covering of a network for $l = 3$. (a) Original network. (b) Dual network. (c) Colouring process. (d) Mapping colours to the original network.

Table 1. The results of $N_b(l)$ and δ from the network of Figure 3, and the “pseudo matrix” of α' .

l	$N_b(l)$	δ	$\alpha'_{l,1}$	$\alpha'_{l,2}$	$\alpha'_{l,3}$
1	6	-	-	-	-
2	3	0.107	$\alpha'_{2,1}$	$\alpha'_{2,2}$	$\alpha'_{2,3}$
3	2	0.107	$\alpha'_{3,1}$	$\alpha'_{3,2}$	-
4	2	0.143	$\alpha'_{4,1}$	$\alpha'_{4,2}$	-
5	1	-	-	-	-

Now, we are ready to compute *innerdeg*. Two boxes were found following the previous example for for $l = 3$; see Figure 4a. The $innerdeg(G_1) = 2$ is the average link per node between the nodes of this box; for this reason, the link between nodes four and six is omitted in this computation. Similarly, $innerdeg(G_2) = 1$. The *outerdeg* is the degree of each node of the renormalized network; see the network of Figure 4b.

In our example, $outerdeg(G_1) = outerdeg(G_2) = 1$. The renormalization converts each box into a super node, preserving the connections between boxes. On the other hand, it is known that $N_b(1) = n$, and $N_b(\Delta + 1) = 1$; in the first case, each box contains a node, and in the second one, there is one box to cover the network that contains all nodes. For this reason, the *innerdeg* and *outerdeg* are not defined for $l = 1$ and $l = \Delta + 1$, respectively. This force to $l = [2, \Delta]$ as was stated in (19)–(24). Additionally, note that the right hand of (19) and (20) (α'), (21) and (22) (β') are “pseudo matrices”, where each row has $N_b(l)$ values; see Table 1. Consequently, $q_{\alpha'}$ and $q_{\beta'}$ are also “pseudo matrices”.

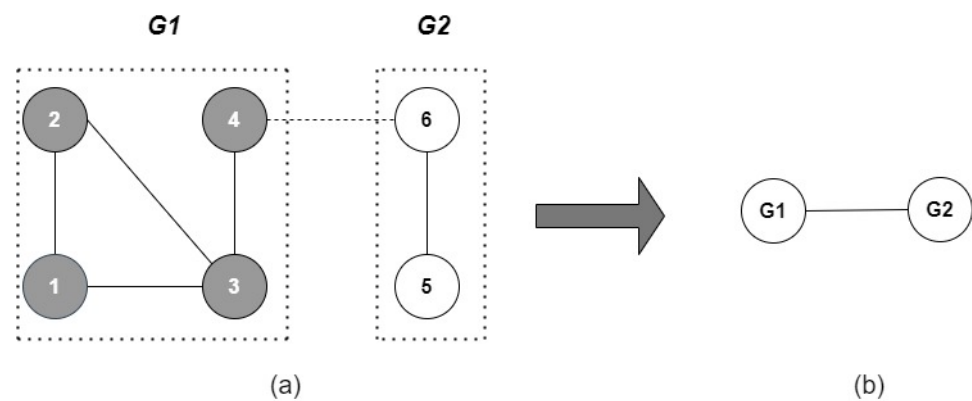


Figure 4. Renormalization of a network. (a) Grouping nodes into boxes. (b) Converting boxes into supernodes.

The network represents the relationships between attribute-value (nodes) of each record and the relationships between different database records. For example, the dotted lines in Figure 2 show the relationships between the first record's attribute value. Links of the node ZIP.08510 are the relationships between the three records, and the links of PHONE.54-76-90 are the relationships between the first and third one. The box-covering algorithm groups these relationships into boxes (network records). The network in the middle of Figure 2 shows that the three boxes (in orange, green, and blue) coincide with the number of records in the database. However, the attribute value of each box does not coincide entirely with records in the database since box-covering finds the minimum number of boxes with the maximum number of attributes where the boxes are mutually exclusive.

The nodes in each box (network record) are enough to differentiate the records in the database. For example, the first network record consists of name, phone, and zip values (nodes in orange). The second record in the database can be differentiated from the first by its name and phone (values of those attributes are the second network record in green). The third one can be distinguished from the two others by its name (the third network record in blue). The cost of differentiating the first network record (measured by *innerdeg*) is the highest; meanwhile, the lowest is for the third. Thus, α' measures the local differentiation cost for the network records.

On the other hand, β' measures the global differentiation cost (by *outerdeg*). For example, the global cost for the first network record is two, and one for the second and third; see the renormalized network (at the bottom of Figure 2). It means that the first network record needs to be differentiated from two network records, and the second and third only need to be distinguished from the first. Note that α' , β' for a given l relies on the topology network that captures the relationships of the records and their values. Finally, $q_{\alpha'}$ is the ratio between network records (normalized number of boxes δ) and the local differentiation cost; meanwhile, $q_{\beta'}$ is the ratio between network records and the global differentiation cost.

4. Methodology

Twelve databases (biological, technological, and social disciplines) from the UCI repository [55] were managed in the experiments; see Table 2. Their number of records, attributes, and classes are representative. Once a network was obtained from the database, q , α , and β parameters of the fractional Tsallis decision were approximated by the following four sets: $(\langle q_{\alpha'} \rangle, \langle \alpha^1 \rangle, \langle \beta^1 \rangle)$, $(\langle q_{\alpha'} \rangle, \langle \alpha^2 \rangle, \langle \beta^2 \rangle)$, $(\langle q_{\beta'} \rangle, \langle \alpha^1 \rangle, \langle \beta^1 \rangle)$, $(\langle q_{\beta'} \rangle, \langle \alpha^2 \rangle, \langle \beta^2 \rangle)$, where $\langle \rangle$ means the average value of the pseudo matrices obtained by (19)–(24).

The network can be obtained from a raw database or after being discretized. Since the classification—measured by the area under receiver operating characteristic curve (AU-ROC) and Matthews correlation coefficient (MCC)—was better using the approximations

computed on the networks from discretized databases, these approximations are only reported. The attribute discretization of a database can be found in [56]. The discretization technique is unsupervised and uses equal-frequency binning. The discretized databases were only used to obtain the networks so that the classification task was carried out using the original databases. The networks obtained from the discretized and non-discretized databases turned out to be different; see Figure 5.

Table 2. Database and network features. N = nominal, U = numerical, M = mixed.

Database	Records	Attributes	Type	Classes	Balanced	Nodes	Edges
Breast Cancer	699	9	N	2	No	737	1276
Car	1728	6	N	4	Yes	25	70
Cmc	1473	9	M	3	No	74	264
Glass	214	10	U	7	No	1159	1743
Haberman	306	3	U	2	No	94	395
Hayes	160	5	N	3	No	150	186
Image	2310	19	U	7	Yes	12,705	24,411
Letter	20,000	16	U	16	Yes	282	2700
Scale	625	4	N	3	No	23	90
Vehicle	946	18	U	4	Yes	1434	8064
Wine	178	13	U	3	No	1279	2239
Yeast	1484	9	M	10	No	1917	4907

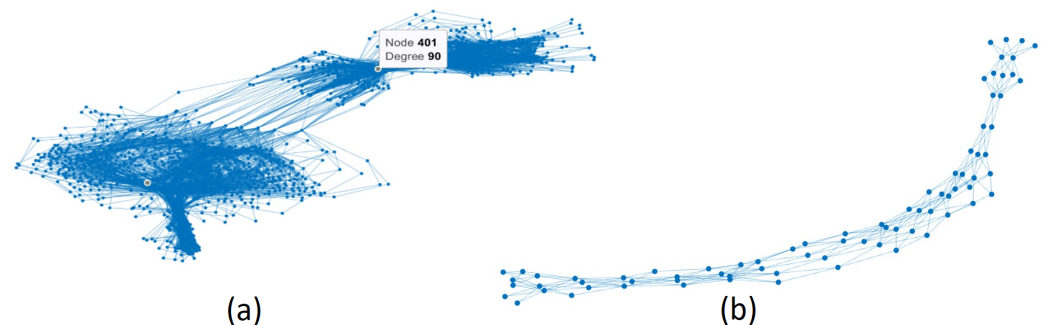


Figure 5. The networks from (a) non-discretized and (b) discretized vehicle database.

The classification task was performed by classical, Renyi, Tsallis, Gini, and the two-parameter fractional Tsallis decisions trees on each database. We used a 10-fold cross-validation repeated ten times to calculate the AUROC and MCC. The best value of the AUROC and MCC, produced by one of the four sets of parameters—used to approximate q , α , and β —of fractional Tsallis decisions trees, was chosen and compared with the classical and Gini decision trees. In the same way, $q_{\alpha'}$ or $q_{\beta'}$ was chosen for the q parameter of Renyi and Tsallis trees. Then, their AUROCs and MCCs were compared with those of the classical trees. It is known that decision trees could produce non-normal distributed AUROC and MCC measures [57]. Hence, the normality was verified by the Kolmogorov–Smirnov test. These measures were compared using a T or a U Mann–Whitney test, according to their normality [10,57–59].

5. Applications

The approximations of q , α , and β parameters computed on discretized databases are shown in Table 3. Table 4 shows the AUROC and MCC of classical and two-parameter fractional Tsallis decisions trees and the result of the statistical comparison. In addition, the values of the parameters of fractional Tsallis decisions trees are reported.

The two-parameter fractional Tsallis decision tree outperforms the AUROC and MCC of the classical trees for eight databases. The statistical result of both measures disagrees with Car and Haberman. The AUROC of the two-parameter fractional Tsallis tree was equal to the classical trees for Car, Image, Vehicle, and Yeast; meanwhile, for Haberman, Image, Vehicle, and Yeast, the MCC of both trees showed no difference.

Table 3. The parameters of the fractional Tsallis decision tree were obtained using the networks from discretized databases.

Database	$\langle q_{\alpha'} \rangle$	$\langle q_{\beta'} \rangle$	$\langle \alpha^1 \rangle$	$\langle \beta^1 \rangle$	$\langle \alpha^2 \rangle$	$\langle \beta^2 \rangle$
Breast Cancer	0.173	0.189	1.147	1.134	0.853	0.866
Car	0.303	0.347	1.137	1.120	0.863	0.880
Cmc	0.169	0.185	1.152	1.138	0.848	0.862
Glass	0.171	0.187	1.154	1.141	0.846	0.859
Haberman	0.344	0.420	1.333	1.273	0.667	0.727
Hayes	0.269	0.310	1.231	1.200	0.769	0.800
Image	0.117	0.123	1.056	1.054	0.944	0.946
Letter	0.155	0.165	1.05	1.047	0.950	0.953
Scale	0.352	0.421	1.217	1.182	0.783	0.818
Vehicle	0.092	0.096	1.106	1.101	0.894	0.899
Wine	0.119	0.127	1.147	1.138	0.853	0.862
Yeast	4.574	5.081	1.003	1.003	0.997	0.997

Tsallis entropy is a non-extensive measure [60] as well as a two-parameter fractional Tsallis entropy [22]. On the contrary, Shannon entropy is extensive. The super-extensive property is given by $q < 1$, and sub-extensive property by $q > 1$. Note that the approximations of the q parameter for all the databases, see Table 3, are < 1 except for Yeast. Thus, they can be considered candidates for being named super-extensive databases. We say that a database is super-extensive if $q < 1$ and its value produces a better classification (AUROC, MCC, or another measure) than the classical trees (based on Shannon entropy). Similarly, a database is sub-extensive if $q > 1$ and its value produces a better classification. Otherwise, the database is extensive since, in this case, the Shannon entropy (the cornerstone of classical trees) is a less complex measure than the two-parameter fractional Tsallis entropy; hence Shannon entropy must be preferred. The two-parameter fractional Tsallis trees produce classifications equal to or better than the classical trees. Following those conditions, based on MCC, Breast Cancer, Car, Cmc, Glass, Hayes, Letter, Scale, and Wine are super-extensive. Meanwhile, Haberman, Image, Vehicle, and Yeast can be classified as extensive.

The AUROC and MCC of Renyi and Tsallis decision trees are compared with the baseline of the classical ones. The $q_{\alpha'}$ and $q_{\beta'}$ were tested as the entropic index of both parametric decision trees. The parameters of Renyi (q_r) and Tsallis (q_t) that produce the better AUROCs and MCCs are reported in Table 5. The result shows that the AUROC of Renyi trees was better for Breast Cancer, Glass, Letter, and Yeast and worse for Cmc and Haberman than classical trees. The results are quite similar for MCC, where Car's classification outperforms the classical tree classification. On the contrary, the MCC of the Vehicle database was statistically less than that of the classical tree. The Tsallis AUROCs were better for Cmc, Glass, Haberman, Hayes, and Wine and worse for Yeast than those of classical trees. Additionally, the MCCs of Car, Cmc, Glass, and Scale were higher, and lower for Yeast, than the classical trees' MCCs. Based on MCC, Car, Cmc, Glass, and Scale are super-extensive, which is a subset of the classification obtained by two-parameter fractional Tsallis.

Table 4. The AUROC and MCC of classical (CT) and two-parameter fractional Tsallis decision trees (TFTT) and their parameters q, α, β . + means that AUROC or MCC is statistically greater than AUROC or MCC of CT.

Database	CT_{AUROC}	$TFTT_{AUROC}$	CT_{MCC}	$TFTT_{MCC}$	q	α	β	Param. Set
Breast Cancer	0.959	0.964 +	0.889	0.967 +	0.173	0.853	0.866	$\langle q_{\alpha'} \rangle, \langle \alpha^2 \rangle, \langle \beta^2 \rangle$
Car	0.981	0.982	0.892	0.912 +	0.347	0.863	0.880	$\langle q_{\beta'} \rangle, \langle \alpha^2 \rangle, \langle \beta^2 \rangle$
Cmc	0.691	0.714 +	0.315	0.349 +	0.169	1.152	1.138	$\langle q_{\alpha'} \rangle, \langle \alpha^1 \rangle, \langle \beta^1 \rangle$
Glass	0.794	0.874 +	0.56	0.673 +	0.171	1.154	1.141	$\langle q_{\alpha'} \rangle, \langle \alpha^1 \rangle, \langle \beta^1 \rangle$
Haberman	0.579	0.610 +	0.18	0.156	0.344	1.333	1.273	$\langle q_{\alpha'} \rangle, \langle \alpha^1 \rangle, \langle \beta^1 \rangle$
Hayes	0.869	0.895 +	0.578	0.645 +	0.269	1.231	1.200	$\langle q_{\alpha'} \rangle, \langle \alpha^1 \rangle, \langle \beta^1 \rangle$
Image	0.994	0.992	0.982	0.978	0.123	1.056	1.054	$\langle q_{\beta'} \rangle, \langle \alpha^1 \rangle, \langle \beta^1 \rangle$
Letter	0.969	0.974 +	0.912	0.934 +	0.155	0.950	0.953	$\langle q_{\alpha'} \rangle, \langle \alpha^2 \rangle, \langle \beta^2 \rangle$
Scale	0.845	0.861 +	0.678	0.703 +	0.421	1.217	1.182	$\langle q_{\beta'} \rangle, \langle \alpha^1 \rangle, \langle \beta^1 \rangle$
Vehicle	0.762	0.755	0.395	0.387	0.092	0.894	0.899	$\langle q_{\alpha'} \rangle, \langle \alpha^2 \rangle, \langle \beta^2 \rangle$
Wine	0.968	0.977 +	0.933	0.957 +	0.119	1.147	1.138	$\langle q_{\alpha'} \rangle, \langle \alpha^1 \rangle, \langle \beta^1 \rangle$
Yeast	0.743	0.733	0.462	0.463	4.574	0.997	0.997	$\langle q_{\alpha'} \rangle, \langle \alpha^2 \rangle, \langle \beta^2 \rangle$

Table 5. AUROC and MCC of classical (CT), Renyi (RT), and Tsallis (TT) decision trees. + means that AUROC is statistically greater than AUROC or MCC of CT, and – means the opposite.

Database	CT_{AUROC}	RT_{AUROC}	TT_{AUROC}	CT_{MCC}	RT_{MCC}	TT_{MCC}	q_r	q_t
Breast Cancer	0.959	0.971 +	0.963	0.889	0.901 +	0.887	$\langle q_{\alpha'} \rangle = 0.173$	$\langle q_{\alpha'} \rangle = 0.173$
Car	0.981	0.983	0.982	0.892	0.906 +	0.912 +	$\langle q_{\alpha'} \rangle = 0.303$	$\langle q_{\beta'} \rangle = 0.347$
Cmc	0.691	0.676 –	0.712 +	0.315	0.256 –	0.35 +	$\langle q_{\beta'} \rangle = 0.185$	$\langle q_{\alpha'} \rangle = 0.169$
Glass	0.794	0.838 +	0.835 +	0.56	0.622 +	0.599 +	$\langle q_{\beta'} \rangle = 0.187$	$\langle q_{\alpha'} \rangle = 0.171$
Haberman	0.579	0.500 –	0.610 +	0.18	0.024 –	0.152	$\langle q_{\alpha'} \rangle = 0.344$	$\langle q_{\alpha'} \rangle = 0.334$
Hayes	0.869	0.869	0.895 +	0.578	0.579	0.587	$\langle q_{\alpha'} \rangle = 0.269$	$\langle q_{\alpha'} \rangle = 0.269$
Image	0.994	0.997	0.995	0.982	0.984	0.978	$\langle q_{\alpha'} \rangle = 0.117$	$\langle q_{\beta'} \rangle = 0.123$
Letter	0.969	0.980 +	0.967	0.912	0.939 +	0.913	$\langle q_{\beta'} \rangle = 0.165$	$\langle q_{\alpha'} \rangle = 0.155$
Scale	0.845	0.839	0.857	0.678	0.651	0.706 +	$\langle q_{\beta'} \rangle = 0.421$	$\langle q_{\beta'} \rangle = 0.421$
Vehicle	0.762	0.776	0.748	0.395	0.297 –	0.371	$\langle q_{\beta'} \rangle = 0.096$	$\langle q_{\alpha'} \rangle = 0.092$
Wine	0.968	0.963	0.976 +	0.933	0.923	0.924	$\langle q_{\alpha'} \rangle = 0.119$	$\langle q_{\alpha'} \rangle = 0.119$
Yeast	0.743	0.789 +	0.578 –	0.462	0.505 +	0.098 –	$\langle q_{\beta'} \rangle = 5.081$	$\langle q_{\alpha'} \rangle = 4.574$

Finally, the Gini and the two-parameter fractional Tsallis decisions trees are compared using AUROC and MCC. The results are shown in Table 6. These results indicate that two-parameter fractional Tsallis trees outperform AUROC of Gini trees in six databases, and MCC in ten. It underpins that Gini trees are a particular case of two-parameter fractional Tsallis trees with $q = 2$. In summarizing, two-parameter fractional Tsallis trees have better classifications than classical and Gini trees.

Table 6. AUROC and MCC of Gini decision trees (GT) and two-parameter fractional Tsallis decision trees (TFTT). + means that AUROC is statistically greater than AUROC or MCC of GT.

Database	GT_{AUROC}	$TFTT_{AUROC}$	GT_{MCC}	$TFTT_{MCC}$
Breast Cancer	0.963	0.964	0.888	0.967 +
Car	0.981	0.982	0.897	0.912 +
Cmc	0.58	0.714 +	0.357	0.349
Glass	0.712	0.874 +	0.437	0.673 +
Haberman	0.52	0.61 +	0.068	0.156 +
Hayes	0.871	0.895 +	0.655	0.645
Image	0.988	0.992	0.946	0.978 +
Letter	0.962	0.974 +	0.894	0.934 +
Scale	0.866	0.861	0.654	0.703 +
Vehicle	0.71	0.755 +	0.294	0.387 +
Wine	0.932	0.977	0.847	0.957 +
Yeast	0.728	0.733	0.414	0.463 +

6. Conclusions

This paper introduces two-parameter fractional Tsallis decision trees underpinned by fractional-order entropies. The three parameters of this new decision tree need to be tuned to produce better classifications than the classical ones. The trial and error approach is the standard method to adjust the entropic index for Renyi and Tsallis decision trees. However, it is unfeasible for two-parameter fractional Tsallis trees. From a database representation as a complex network, it was possible to determine a set of values for parameters q , α , and β based on this network. The experimental results on twelve databases show that the proposed values yield better classifications (AUROC, MCC) for eight of them, and for the four remaining, the classification was equal to that produced by classical trees.

Moreover, two values ($q_{\alpha\beta}$, $q_{\beta\alpha}$) were tested in Renyi and Tsallis decision trees. The results show that Renyi outperforms the classical trees in four (AUROC) and five (MCC) out of twelve databases. Similarly, Tsallis decision trees produced better classification for five (AUROC) and four (MCC) databases. The classification was worse in almost three and one databases for Renyi and Tsallis, respectively. The overall results of both parametric decision trees suggest that both outperform the classical trees in seven databases. All of the above is less favorable than what happened in eight databases analyzed with the two-parameter fractional Tsallis decision trees. In addition, the databases with a better classification using Tsallis decision trees are a subset of those for which two-parameter fractional Tsallis trees produced a better classification. It supports the conjecture that two-parameter fractional Tsallis entropy is a finer measure than the parametric entropies such as Renyi and Tsallis.

The approximate technique for the tree parameters introduced here is a valuable alternative for practitioners. Furthermore, the network classification based on the non-extensive properties of Tsallis and two-parameter fractional Tsallis entropies reveals that the relationships between the records and their attribute values (modeled by a network) are complex. Such complex relationships are better measured by two-parameter fractional Tsallis entropy, the cornerstone of the proposed decision tree.

The results pave the way for using the two-parameter Tsallis fractional entropy in other data mining techniques such as K-means, generic MST, Kruskal MST, and algorithms for dimension reduction in the future. Our research has the limitation that the databases used in the experiments are not large enough to reveal the reduction in time compared with the trial-and-error approach to set the tree parameters. However, we may conjecture that our method works in large databases, which will be the scope of future research.

Author Contributions: Conceptualization, A.R.-A.; formal analysis, J.S.D.I.C.-G.; investigation, J.S.D.I.C.-G. and J.B.-R.; methodology, J.S.D.I.C.-G.; supervision, A.R.-A.; writing—original draft, J.S.D.I.C.-G.; writing—review and editing, J.B.-R. and A.R.-A. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Secretaria de Investigación de Posgrado grant number SIP20220415.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data that support the findings of this study are available from the corresponding author upon reasonable request.

Acknowledgments: This work was partially supported by Secretaria de Investigación de Posgrado under Grant No. SIP20220415.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

MDPI	Multidisciplinary Digital Publishing Institute
DOAJ	Directory of open-access journals
TLA	Three-letter acronym
LD	Linear dichroism

References

- Shannon, C.E. A mathematical theory of communication. *Bell Syst. Tech. J.* **1948**, *27*, 379–423. [[CrossRef](#)]
- Rényi, A. On measures of entropy and information. In *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Contributions to the Theory of Statistics*; The Regents of the University of California, University of California Press: Berkeley, CA, USA, 1961; pp. 547–561.
- Tsallis, C. Possible generalization of Boltzmann-Gibbs statistics. *J. Stat. Phys.* **1988**, *52*, 479–487. [[CrossRef](#)]
- Abe, S. A note on the q-deformation-theoretic aspect of the generalized entropies in nonextensive physics. *Phys. Lett. A* **1997**, *224*, 326–330. [[CrossRef](#)]
- Johal, R.S. q calculus and entropy in nonextensive statistical physics. *Phys. Rev. E* **1998**, *58*, 4147. [[CrossRef](#)]
- Lavagno, A.; Swamy, P.N. q-Deformed structures and nonextensive-statistics: A comparative study. *Phys. A Stat. Mech. Appl.* **2002**, *305*, 310–315. [[CrossRef](#)]
- Jackson, D.O.; Fukuda, T.; Dunn, O.; Majors, E. On q-definite integrals. *Q. J. Pure Appl. Math.* **1910**, *41*, 193–203.
- Duan, S.; Wen, T.; Jiang, W. A new information dimension of complex network based on Rényi entropy. *Phys. A Stat. Mech. Appl.* **2019**, *516*, 529–542. [[CrossRef](#)]
- Maszczyk, T.; Duch, W. Comparison of Shannon, Rényi and Tsallis Entropy Used in Decision Trees. In *Artificial Intelligence and Soft Computing—ICAISC 2008*; Rutkowski, L., Tadeusiewicz, R., Zadeh, L.A., Zurada, J.M., Eds.; Springer: Berlin/Heidelberg, Germany, 2008; pp. 643–651.
- Ramirez-Arellano, A.; Bory-Reyes, J.; Hernandez-Simon, L.M. Statistical Entropy Measures in C4.5 Trees. *Int. J. Data Warehous. Min.* **2018**, *14*, 1–14. [[CrossRef](#)]
- Gajowniczek, K.; Orłowski, A.; Zabkowski, T. Entropy Based Trees to Support Decision Making for Customer Churn Management. *Acta Phys. Pol. A* **2016**, *129*, 971–979. [[CrossRef](#)]
- Lima, C.F.L.; de Assis, F.M.; Cleonilson Protásio, C.P. Decision Tree Based on Shannon, Rényi and Tsallis Entropies for Intrusion Tolerant Systems. In *Proceedings of the 2010 Fifth International Conference on Internet Monitoring and Protection*, Barcelona, Spain, 9–15 May 2010; pp. 117–122. [[CrossRef](#)]
- Wang, Y.; Song, C.; Xia, S.T. Improving decision trees by Tsallis Entropy Information Metric method. In *Proceedings of the 2016 International Joint Conference on Neural Networks (IJCNN)*, Vancouver, BC, Canada, 24–29 July 2016; pp. 4729–4734. [[CrossRef](#)]
- Wang, Y.; Xia, S.T.; Wu, J. A less-greedy two-term Tsallis Entropy Information Metric approach for decision tree classification. *Knowl.-Based Syst.* **2017**, *120*, 34–42. [[CrossRef](#)]
- Sharma, S.; Bassi, I. Efficacy of Tsallis Entropy in Clustering Categorical Data. In *Proceedings of the 2019 IEEE Bombay Section Signature Conference (IBSSC)*, Mumbai, India, 26–28 July 2019; pp. 1–5. [[CrossRef](#)]
- Zhang, L.; Cao, Q.; Lee, J. A novel ant-based clustering algorithm using Rényi entropy. *Appl. Soft Comput.* **2013**, *13*, 2643–2657. [[CrossRef](#)]
- Wang, Y.; Xia, S.T. Unifying attribute splitting criteria of decision trees by Tsallis entropy. In *Proceedings of the 2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, New Orleans, LA, USA, 5–9 March 2017; pp. 2507–2511. [[CrossRef](#)]

18. Tsallis, C.; Tirnakli, U. Non-additive entropy and nonextensive statistical mechanics – Some central concepts and recent applications. *J. Phys. Conf. Ser.* **2010**, *201*, 012001. [[CrossRef](#)]
19. Tsallis, C. *Introduction to Non-Extensive Statistical Mechanics: Approaching a Complex World*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2009.
20. Ramirez-Arellano, A.; Hernández-Simón, L.M.; Bory-Reyes, J. A box-covering Tsallis information dimension and non-extensive property of complex networks. *Chaos Solitons Fractals* **2020**, *132*, 109590. [[CrossRef](#)]
21. Ramirez-Arellano, A.; Sigarreta-Almira, J.M.; Bory-Reyes, J. Fractional information dimensions of complex networks. *Chaos Interdiscip. J. Nonlinear Sci.* **2020**, *30*, 093125. [[CrossRef](#)]
22. Ramirez-Arellano, A.; Hernández-Simón, L.M.; Bory-Reyes, J. Two-parameter fractional Tsallis information dimensions of complex networks. *Chaos Solitons Fractals* **2021**, *150*, 111113. [[CrossRef](#)]
23. Ramírez-Reyes, A.; Hernández-Montoya, A.R.; Herrera-Corral, G.; Domínguez-Jiménez, I. Determining the Entropic Index q of Tsallis Entropy in Images through Redundancy. *Entropy* **2016**, *18*, 299. [[CrossRef](#)]
24. Chen, X.; Zhou, J.; Liao, Z.; Liu, S.; Zhang, Y. A Novel Method to Rank Influential Nodes in Complex Networks Based on Tsallis Entropy. *Entropy* **2020**, *22*, 848. [[CrossRef](#)]
25. Zhang, Q.; Li, M.; Deng, Y. A new structure entropy of complex networks based on non-extensive statistical mechanics. *Int. J. Mod. Phys. C* **2016**, *27*, 1650118. [[CrossRef](#)]
26. Shafee, F. Lambert function and a new non-extensive form of entropy. *IMA J. Appl. Math.* **2007**, *72*, 785–800. [[CrossRef](#)]
27. Ubriaco, M.R. Entropies based on fractional calculus. *Phys. Lett. A* **2009**, *373*, 2516–2519. [[CrossRef](#)]
28. Ubriaco, M.R. A simple mathematical model for anomalous diffusion via Fisher’s information theory. *Phys. Lett. A* **2009**, *373*, 4017–4021. [[CrossRef](#)]
29. Karci, A. Fractional order entropy: New perspectives. *Optik* **2016**, *127*, 9172–9177. [[CrossRef](#)]
30. Karci, A. Notes on the published article “Fractional order entropy: New perspectives” by Ali KARCI, *Optik-International Journal for Light and Electron Optics*, Volume 127, Issue 20, October 2016, Pages 9172–9177. *Optik* **2018**, *171*, 107–108. [[CrossRef](#)]
31. Radhakrishnan, C.; Chinnarasu, R.; Jambulingam, S. A Fractional Entropy in Fractal Phase Space: Properties and Characterization. *Int. J. Stat. Mech.* **2014**, *2014*, 460364. [[CrossRef](#)]
32. Ferreira, R.A.C.; Tenreiro Machado, J. An Entropy Formulation Based on the Generalized Liouville Fractional Derivative. *Entropy* **2019**, *21*, 638. [[CrossRef](#)]
33. Machado, J.T. Entropy analysis of integer and fractional dynamical systems. *Nonlinear Dyn.* **2010**, *62*, 371–378.
34. Machado, J.T. Fractional order generalized information. *Entropy* **2014**, *16*, 2350–2361. [[CrossRef](#)]
35. Wang, Q.A. Extensive Generalization of Statistical Mechanics Based on Incomplete Information Theory. *Entropy* **2003**, *5*, 220–232. [[CrossRef](#)]
36. Wang, Q.A. Incomplete statistics: Nonextensive generalizations of statistical mechanics. *Chaos Solitons Fractals* **2001**, *12*, 1431–1437. [[CrossRef](#)]
37. Kaniadakis, G. Maximum entropy principle and power-law tailed distributions. *Eur. Phys. J. B* **2009**, *70*, 3–13. [[CrossRef](#)]
38. Tsallis, C. An introduction to nonadditive entropies and a thermostatistical approach to inanimate and living matter. *Contemp. Phys.* **2014**, *55*, 179–197. [[CrossRef](#)]
39. Kapitaniak, T.; Mohammadi, S.A.; Mekhilef, S.; Alsaadi, F.E.; Hayat, T.; Pham, V.T. A New Chaotic System with Stable Equilibrium: Entropy Analysis, Parameter Estimation, and Circuit Design. *Entropy* **2018**, *20*, 670. [[CrossRef](#)] [[PubMed](#)]
40. Jalab, H.A.; Subramaniam, T.; Ibrahim, R.W.; Kahtan, H.; Noor, N.F.M. New Texture Descriptor Based on Modified Fractional Entropy for Digital Image Splicing Forgery Detection. *Entropy* **2019**, *21*, 371. [[CrossRef](#)] [[PubMed](#)]
41. Ibrahim, R.W.; Jalab, H.A.; Gani, A. Entropy solution of fractional dynamic cloud computing system associated with finite boundary condition. *Bound. Value Probl.* **2016**, *2016*, 94. [[CrossRef](#)]
42. He, S.; Sun, K.; Wu, X. Fractional symbolic network entropy analysis for the fractional-order chaotic systems. *Phys. Scr.* **2020**, *95*, 035220. [[CrossRef](#)]
43. Machado, J.T.; Lopes, A.M. Fractional Rényi entropy. *Eur. Phys. J. Plus* **2019**, *134*, 217. [[CrossRef](#)]
44. Beck, C. Generalized information and entropy measures in physics. *Contemp. Phys.* **2009**, *50*, 495–510. [[CrossRef](#)]
45. Han, J.; Kamber, M.; Pei, J. *Data Mining: Concepts and Techniques*, 3rd ed.; Morgan Kaufmann Publishers Inc.: San Francisco, CA, USA, 2011.
46. Hilpert, J.C.; Marchand, G.C. Complex Systems Research in Educational Psychology: Aligning Theory and Method. *Educ. Psychol.* **2018**, *53*, 185–202. [[CrossRef](#)]
47. Karuza, E.A.; Thompson-Schill, S.L.; Bassett, D.S. Local Patterns to Global Architectures: Influences of Network Topology on Human Learning. *Trends Cogn. Sci.* **2016**, *20*, 629–640. [[CrossRef](#)]
48. Ramirez-Arellano, A. Students learning pathways in higher blended education: An analysis of complex networks perspective. *Comput. Educ.* **2019**, *141*, 103634. [[CrossRef](#)]
49. Zhao, X.; Wang, X.Y. An Approach to Compute Fractal Dimension of Color Images. *Fractals* **2017**, *25*, 1750007. [[CrossRef](#)]
50. Stanisz, T.; Kwapien, J.; Drozd, S. Linguistic data mining with complex networks: A stylometric-oriented approach. *Inf. Sci.* **2019**, *482*, 301–320. [[CrossRef](#)]
51. RamirezArellano, A. Classification of Literary Works: Fractality and Complexity of the Narrative, Essay, and Research Article. *Entropy* **2020**, *22*, 904. [[CrossRef](#)] [[PubMed](#)]

52. Kim, J.; Wilhelm, T. What is a complex graph? *Phys. A Stat. Mech. Appl.* **2008**, *387*, 2637–2652. [[CrossRef](#)]
53. van Steen, M. *Graph Theory and Complex Networks: An Introduction*; Cambridge University Press: Cambridge, MA, USA, 2010.
54. Song, C.; Gallos, L.K.; Havlin, S.; Makse, H.A. How to calculate the fractal dimension of a complex network: The box covering algorithm. *J. Stat. Mech. Theory Exp.* **2007**, *2007*, P03006. [[CrossRef](#)]
55. Dua, D.; Graff, C. UCI Machine Learning Repository. 2017. Available online: <https://archive.ics.uci.edu/ml/index.php> (accessed on 22 February 2022).
56. Yang, Y.; Webb, G.I. Proportional k-Interval Discretization for Naive-Bayes Classifiers. In *European Conference on Machine Learning (ECML 2001)*; Springer: Berlin/Heidelberg, Germany, 2001; Volume 2167, pp. 564–575.
57. Demšar, J. Statistical Comparisons of Classifiers over Multiple Data Sets. *J. Mach. Learn. Res.* **2006**, *7*, 1–30.
58. Sprent, P.; Smeeton, N.C. *Applied Nonparametric Statistical Methods*, 3rd ed.; Texts in Statistical Science; Chapman & Hall/CRC: Boca Raton, FL, USA, 2001.
59. Montgomery, D.C.; Runger, G.C. *Applied Statistics and Probability for Engineers*; John Wiley & Sons: Hoboken, NJ, USA, 2003.
60. Tsallis, C. Entropic nonextensivity: A possible measure of complexity. *Chaos Solitons Fractals* **2002**, *13*, 371–391. [[CrossRef](#)]