

A new parametric information-gain criterion for tree-based machine learning algorithms

Diogo Costa¹, Vasco Costa², and Eugénio M. Rocha^{1,2}

¹Center for Research & Development in Mathematics and Applications, University of Aveiro, Aveiro, Portugal

²Department of Mathematics, University of Aveiro, Aveiro, Portugal

Corresponding author:

Diogo Costa¹

Email address: d.costa@ua.pt

COMPLETE RESULTS FOR PROPOSED METHOD

The best results are highlighted in bold in the tables. In the case of the introduced entropy estimation $\hat{E}$, it is considered the "best" when it is possible to converge the Bayesian search (within the set number of iterations) to a parametric combination that does not match any known formulation, that still provides performance equal or greater than classical entropies. Additionally, underlined values mean that the search did not manage to converge to a nonclassical parametrization, and hence $\hat{E}$ is only able to match the performance result using a classical entropy.

Table S1. Results obtained for the DT algorithm with parameters: $\{max_depth = 10\}$. Shown are the mean and standard deviation for the weighted F_1 -score over the 5 folds (for each dataset), computed for the proposed method $\hat{E}$, Gini, Shannon, Tsallis, and Rényi. The best parameters found for $\hat{E}$ are shown. The best values are in bold.

Dataset name	$\{\alpha; \beta; \gamma; \zeta_1; \zeta_2\}$	$\hat{E}$	Gini	Shannon	Tsallis ($k = 1, q = 2$)	Rényi ($\alpha = 2$)	LI (%)	HI (%)
BreastCancer	{2.0; 0.2; 1.4; 0.0; 0}	0.957 ± 0.005	0.941 ± 0.017	0.938 ± 0.031	0.943 ± 0.008	0.941 ± 0.017	1.463	1.985
Iris	{1.0; 1.4; 0.7; 1.0; 0}	0.941 ± 0.038	0.920 ± 0.046	0.941 ± 0.038	0.913 ± 0.041	0.920 ± 0.046	0	2.976
Spambase	{0.0; 0.7; 2.6; 0.0; 1}	0.923 ± 0.006	0.920 ± 0.004	0.920 ± 0.011	0.917 ± 0.010	0.920 ± 0.004	0.325	0.650
Statlog	{2.0; 2.0; 1.5; 0.0; 0}	0.941 ± 0.003	0.924 ± 0.003	0.911 ± 0.002	0.898 ± 0.002	0.924 ± 0.003	1.807	4.570
Wine Quality	{1.0; 2.6; 2.4; 0.0; 0}	0.547 ± 0.012	0.528 ± 0.016	0.489 ± 0.021	0.511 ± 0.016	0.528 ± 0.016	3.473	10.603
Students	{0.0; 1.2; 0.3; 0.0; 0}	0.733 ± 0.025	0.710 ± 0.008	0.715 ± 0.013	0.719 ± 0.016	0.710 ± 0.008	1.910	3.138
mfeat-morphological	{1.0; 2.9; 2.2; 0.0; 0}	0.700 ± 0.024	0.690 ± 0.018	0.671 ± 0.022	0.676 ± 0.019	0.690 ± 0.018	1.429	4.143
diabetes	{0.0; 0.8; 1.7; 1.0; 1}	0.730 ± 0.019	0.685 ± 0.020	0.700 ± 0.035	0.689 ± 0.039	0.685 ± 0.020	4.110	6.164
wdbc	{2.0; 1.5; 0.6; 1.0; 0}	0.947 ± 0.012	0.926 ± 0.026	0.942 ± 0.017	0.926 ± 0.036	0.926 ± 0.026	0.528	2.218
wilt	{1.0; 0.1; 0.3; 1.0; 0}	0.980 ± 0.005	0.977 ± 0.005	0.978 ± 0.007	0.973 ± 0.005	0.977 ± 0.005	0.204	0.714
Titanic	{1.0; 2.0; 1.8; 0.0; 0}	0.759 ± 0.012	0.759 ± 0.012	0.759 ± 0.012	0.759 ± 0.012	0.759 ± 0.012	0	0
dataset_31_credit-g	{0.0; 0.3; 1.1; 2.0; 1}	0.716 ± 0.019	0.687 ± 0.031	0.697 ± 0.027	0.699 ± 0.022	0.687 ± 0.031	2.374	4.050
phoneme	{2.0; 1.6; 0.0; 2.0; 0}	0.842 ± 0.004	0.842 ± 0.004	0.833 ± 0.013	0.830 ± 0.007	0.842 ± 0.004	0	1.425
synth_1	{0.0; 0.4; 0.5; 2.0; 0}	0.990 ± 0.020	0.950 ± 0.032	0.950 ± 0.032	0.950 ± 0.032	0.950 ± 0.032	4.040	4.040
synth_2	{2.0; 2.0; 2.3; 0.0; 0}	0.452 ± 0.102	0.327 ± 0.114	0.334 ± 0.051	0.368 ± 0.110	0.339 ± 0.101	18.584	27.655
synth_3	{0.0; 2.5; 0.3; 0.0; 0}	0.955 ± 0.024	0.953 ± 0.013	0.945 ± 0.014	0.937 ± 0.017	0.953 ± 0.013	0.209	1.885
synth_4	{1.0; 1.3; 0.5; 0.0; 0}	0.844 ± 0.024	0.828 ± 0.042	0.844 ± 0.024	0.844 ± 0.045	0.828 ± 0.042	0	1.896
synth_5	{0.0; 1.3; 2.7; 0.0; 0}	0.721 ± 0.034	0.681 ± 0.007	0.685 ± 0.014	0.704 ± 0.018	0.682 ± 0.006	2.358	5.548

Table S2. Results obtained for the DT algorithm with parameters: $\{max_depth = 50\}$. Shown are the mean and standard deviation for the weighted F_1 -score over the 5 folds (for each dataset), computed for the proposed method $\hat{E}$, Gini, Shannon, Tsallis, and Rényi. The best parameters found for $\hat{E}$ are shown. The best values are in bold. Underlined values mean that the search did not manage to converge to a nonclassical parametrization.

Dataset name	$\{\alpha; \beta; \gamma; \zeta_1; \zeta_2\}$	$\hat{E}$	Gini	Shannon	Tsallis ($k = 1, q = 2$)	Rényi ($\alpha = 2$)	LI (%)	HI (%)
BreastCancer	{2.0; 0.6; 1.4; 0.0; 0}	0.957 ± 0.005	0.941 ± 0.017	0.938 ± 0.031	0.943 ± 0.008	0.941 ± 0.017	1.463	1.985
Iris	{1.0; 1.9; 0.7; 1.0; 0}	0.941 ± 0.038	0.920 ± 0.046	0.941 ± 0.038	0.913 ± 0.041	0.920 ± 0.046	0	2.976
Spambase	{2.0; 1.6; 2.9; 1.0; 0}	0.936 ± 0.009	0.913 ± 0.005	0.919 ± 0.011	0.905 ± 0.012	0.913 ± 0.005	1.816	3.312
Statlog	{1.0; 1.2; 0.7; 0.0; 0}	0.997 ± 0.001	0.995 ± 0.000	0.997 ± 0.001	0.992 ± 0.001	0.995 ± 0.000	0	0.502
Wine Quality	{1.0; 2.6; 0.5; 0.0; 0}	0.604 ± 0.007	0.567 ± 0.010	0.523 ± 0.017	0.537 ± 0.019	0.568 ± 0.010	5.960	13.411
Students	{1.0; 1.7; 0.5; 0.0; 1}	0.717 ± 0.012	0.676 ± 0.010	0.691 ± 0.010	0.683 ± 0.011	0.676 ± 0.010	3.626	5.718
mfeat-morphological	{2.0; 0.1; 1.6; 0.0; 0}	0.697 ± 0.031	0.664 ± 0.035	0.650 ± 0.014	0.671 ± 0.019	0.663 ± 0.033	3.730	6.743
diabetes	{0.0; 0.4; 1.0; 1.0; 1}	0.745 ± 0.015	0.683 ± 0.023	0.702 ± 0.028	0.687 ± 0.038	0.683 ± 0.023	5.772	8.322
wdbc	{2.0; 2.2; 0.1; 0.0; 0}	0.950 ± 0.023	0.926 ± 0.026	0.942 ± 0.017	0.926 ± 0.036	0.926 ± 0.026	0.842	2.526
wilt	{1.0; 1.0; 0.2; 0.0; 0}	0.979 ± 0.003	0.977 ± 0.005	0.977 ± 0.006	0.973 ± 0.005	0.977 ± 0.005	0.204	0.613
Titanic	{1.0; 2.0; 1.8; 0.0; 0}	0.759 ± 0.012	0.759 ± 0.012	0.759 ± 0.012	0.759 ± 0.012	0.759 ± 0.012	0	0
dataset_31_credit-g	{2.0; 0.5; 1.8; 0.0; 0}	0.720 ± 0.036	0.689 ± 0.026	0.695 ± 0.025	0.699 ± 0.024	0.689 ± 0.026	2.917	4.306
phoneme	{2.0; 2.8; 0.8; 0.0; 0}	0.840 ± 0.009	0.836 ± 0.008	0.829 ± 0.010	0.822 ± 0.012	0.836 ± 0.008	0.476	2.143
synth_1	{0.0; 1.6; 1.5; 2.0; 0}	0.990 ± 0.020	0.950 ± 0.032	0.950 ± 0.032	0.950 ± 0.032	0.950 ± 0.032	4.040	4.040
synth_2	{2.0; 0.7; 1.4; 0.0; 0}	0.477 ± 0.158	0.329 ± 0.112	0.331 ± 0.048	0.368 ± 0.110	0.350 ± 0.093	22.851	31.027
synth_3	{0.0; 2.0; 1.0; 0.0; 0}	<u>0.953 ± 0.013</u>	0.953 ± 0.013	0.945 ± 0.014	0.940 ± 0.018	0.953 ± 0.013	0	1.157
synth_4	{1.0; 0.8; 0.3; 0.0; 0}	0.849 ± 0.029	0.828 ± 0.042	0.843 ± 0.023	0.844 ± 0.045	0.828 ± 0.042	0.589	2.473
synth_5	{2.0; 1.5; 1.2; 1.0; 0}	0.712 ± 0.030	0.669 ± 0.008	0.667 ± 0.020	0.697 ± 0.024	0.670 ± 0.008	2.107	6.320

Table S3. Results obtained for the DT algorithm with parameters: $\{max_depth = 100\}$. Shown are the mean and standard deviation for the weighted F_1 -score over the 5 folds (for each dataset), computed for the proposed method $\hat{E}$, Gini, Shannon, Tsallis, and Rényi. The best parameters found for $\hat{E}$ are shown. The best values are in bold. Underlined values mean that the search did not manage to converge to a nonclassical parametrization.

Dataset name	$\{\alpha; \beta; \gamma; \zeta_1; \zeta_2\}$	$\hat{E}$	Gini	Shannon	Tsallis ($k = 1, q = 2$)	Rényi ($\alpha = 2$)	LI (%)	HI (%)
BreastCancer	{2.0; 0.6; 1.4; 0.0; 0}	0.957 ± 0.005	0.941 ± 0.017	0.938 ± 0.031	0.943 ± 0.008	0.941 ± 0.017	1.463	1.985
Iris	{1.0; 1.9; 0.7; 1.0; 0}	0.941 ± 0.038	0.920 ± 0.046	0.941 ± 0.038	0.913 ± 0.041	0.920 ± 0.046	0	2.976
Spambase	{0.0; 0.3; 1.0; 0.0; 0}	0.927 ± 0.010	0.913 ± 0.005	0.919 ± 0.011	0.905 ± 0.012	0.913 ± 0.005	0.863	2.373
Statlog	{0.0; 2.6; 0.3; 0.0; 0}	0.997 ± 0.001	0.997 ± 0.001	0.997 ± 0.001	0.993 ± 0.001	0.997 ± 0.001	0	0.401
Wine Quality	{1.0; 2.8; 0.3; 0.0; 0}	0.595 ± 0.013	0.579 ± 0.012	0.529 ± 0.020	0.541 ± 0.021	0.579 ± 0.011	2.689	11.092
Students	{2.0; 1.8; 1.3; 1.0; 0}	0.717 ± 0.023	0.676 ± 0.010	0.689 ± 0.010	0.683 ± 0.011	0.676 ± 0.010	3.905	5.718
mfeat-morphological	{2.0; 0.7; 2.8; 0.0; 0}	0.696 ± 0.037	0.664 ± 0.035	0.643 ± 0.017	0.670 ± 0.020	0.663 ± 0.033	3.736	7.615
diabetes	{0.0; 0.4; 1.8; 0.0; 0}	0.726 ± 0.028	0.683 ± 0.023	0.702 ± 0.028	0.687 ± 0.038	0.683 ± 0.023	3.306	5.923
wdbc	{2.0; 0.7; 0.1; 1.0; 1}	0.951 ± 0.025	0.926 ± 0.026	0.942 ± 0.017	0.926 ± 0.036	0.926 ± 0.026	0.946	2.629
wilt	{1.0; 1.0; 0.2; 0.0; 0}	0.979 ± 0.003	0.977 ± 0.005	0.977 ± 0.006	0.973 ± 0.005	0.977 ± 0.005	0.204	0.613
Titanic	{1.0; 2.0; 1.8; 0.0; 0}	0.759 ± 0.012	0.759 ± 0.012	0.759 ± 0.012	0.759 ± 0.012	0.759 ± 0.012	0	0
dataset_31_credit-g	{2.0; 2.2; 1.6; 0.0; 0}	0.726 ± 0.026	0.689 ± 0.026	0.695 ± 0.025	0.699 ± 0.024	0.689 ± 0.026	3.719	5.096
phoneme	{2.0; 1.7; 0.9; 1.0; 0}	0.840 ± 0.007	0.836 ± 0.008	0.829 ± 0.010	0.822 ± 0.012	0.836 ± 0.008	0.476	2.143
synth_1	{0.0; 1.6; 1.5; 2.0; 0}	0.990 ± 0.020	0.950 ± 0.032	0.950 ± 0.032	0.950 ± 0.032	0.950 ± 0.032	4.040	4.040
synth_2	{2.0; 0.7; 1.4; 0.0; 0}	0.477 ± 0.158	0.329 ± 0.112	0.331 ± 0.048	0.368 ± 0.110	0.350 ± 0.093	22.851	31.027
synth_3	{0.0; 2.0; 1.0; 0.0; 0}	<u>0.953 ± 0.013</u>	0.953 ± 0.013	0.945 ± 0.014	0.940 ± 0.018	0.953 ± 0.013	0	1.157
synth_4	{0.0; 2.1; 0.7; 0.0; 0}	0.846 ± 0.022	0.828 ± 0.042	0.843 ± 0.023	0.844 ± 0.045	0.828 ± 0.042	0.236	2.128
synth_5	{0.0; 1.9; 1.3; 0.0; 0}	0.705 ± 0.018	0.669 ± 0.008	0.667 ± 0.020	0.697 ± 0.024	0.670 ± 0.008	1.135	5.390

Table S4. Results obtained for the DT algorithm with parameters: $\{max_depth = 200\}$. Shown are the mean and standard deviation for the weighted F_1 -score over the 5 folds (for each dataset), computed for the proposed method $\hat{E}$, Gini, Shannon, Tsallis, and Rényi. The best parameters found for $\hat{E}$ are shown. The best values are in bold. Underlined values mean that the search did not manage to converge to a nonclassical parametrization.

Dataset name	$\{\alpha; \beta; \gamma; \zeta_1; \zeta_2\}$	$\hat{E}$	Gini	Shannon	Tsallis ($k = 1, q = 2$)	Rényi ($\alpha = 2$)	LI (%)	HI (%)
BreastCancer	{2.0; 0.6; 1.4; 0.0; 0}	0.957 ± 0.005	0.941 ± 0.017	0.938 ± 0.031	0.943 ± 0.008	0.941 ± 0.017	1.463	1.985
Iris	{1.0; 1.9; 0.7; 1.0; 0}	0.941 ± 0.038	0.920 ± 0.046	0.941 ± 0.038	0.913 ± 0.041	0.920 ± 0.046	0	2.976
Spambase	{1.0; 2.6; 0.2; 0.0; 0}	0.931 ± 0.008	0.913 ± 0.005	0.919 ± 0.011	0.905 ± 0.012	0.913 ± 0.005	1.289	2.793
Statlog	{0.0; 2.7; 0.3; 0.0; 0}	0.997 ± 0.000	0.997 ± 0.001	0.997 ± 0.001	0.993 ± 0.001	0.997 ± 0.000	0	0.401
Wine Quality	{1.0; 1.6; 0.3; 0.0; 0}	0.604 ± 0.006	0.582 ± 0.011	0.530 ± 0.020	0.541 ± 0.021	0.583 ± 0.011	3.477	12.252
Students	{2.0; 2.0; 1.3; 1.0; 0}	0.717 ± 0.023	0.676 ± 0.010	0.689 ± 0.010	0.683 ± 0.011	0.676 ± 0.010	3.905	5.718
mfeat-morphological	{1.0; 2.2; 1.8; 0.0; 0}	0.683 ± 0.023	0.664 ± 0.035	0.643 ± 0.017	0.670 ± 0.020	0.663 ± 0.033	1.903	5.857
diabetes	{2.0; 2.9; 1.8; 0.0; 0}	0.742 ± 0.028	0.683 ± 0.023	0.702 ± 0.028	0.687 ± 0.038	0.683 ± 0.023	5.391	7.951
wdbc	{2.0; 0.7; 0.1; 1.0; 1}	0.951 ± 0.025	0.926 ± 0.026	0.942 ± 0.017	0.926 ± 0.036	0.926 ± 0.026	0.946	2.629
wilt	{1.0; 1.0; 0.2; 0.0; 0}	0.979 ± 0.003	0.977 ± 0.005	0.977 ± 0.006	0.973 ± 0.005	0.977 ± 0.005	0.204	0.613
Titanic	{1.0; 2.0; 1.8; 0.0; 0}	0.759 ± 0.012	0.759 ± 0.012	0.759 ± 0.012	0.759 ± 0.012	0.759 ± 0.012	0	0
dataset_31_credit-g	{2.0; 2.4; 1.9; 0.0; 0}	0.727 ± 0.020	0.689 ± 0.026	0.695 ± 0.025	0.699 ± 0.024	0.689 ± 0.026	3.851	5.227
phoneme	{1.0; 2.2; 1.5; 1.0; 0}	0.845 ± 0.007	0.836 ± 0.008	0.829 ± 0.010	0.822 ± 0.012	0.836 ± 0.008	1.065	2.722
synth_1	{0.0; 1.6; 1.5; 2.0; 0}	0.990 ± 0.020	0.950 ± 0.032	0.950 ± 0.032	0.950 ± 0.032	0.950 ± 0.032	4.040	4.040
synth_2	{2.0; 0.7; 1.4; 0.0; 0}	0.477 ± 0.158	0.329 ± 0.112	0.331 ± 0.048	0.368 ± 0.110	0.350 ± 0.093	22.851	31.027
synth_3	{0.0; 2.0; 1.0; 0.0; 0}	<u>0.953 ± 0.013</u>	0.953 ± 0.013	0.945 ± 0.014	0.940 ± 0.018	0.953 ± 0.013	0	1.157
synth_4	{2.0; 1.6; 0.8; 0.0; 0}	0.847 ± 0.032	0.828 ± 0.042	0.843 ± 0.023	0.844 ± 0.045	0.828 ± 0.042	0.354	2.243
synth_5	{1.0; 1.2; 1.6; 0.0; 0}	0.708 ± 0.024	0.669 ± 0.008	0.667 ± 0.020	0.697 ± 0.024	0.670 ± 0.008	1.554	5.791

Table S5. Results obtained for the RF algorithm with parameters: $\{max_depth = 10; n_estimators = 10\}$. Shown are the mean and standard deviation for the weighted F_1 -score over the 5 folds (for each dataset), computed for the proposed method $\hat{E}$, Gini, Shannon, Tsallis, and Rényi. The best parameters found for $\hat{E}$ are shown. The best values are in bold. Underlined values mean that the search did not manage to converge to a nonclassical parametrization.

Dataset name	$\{\alpha; \beta; \gamma; \zeta_1; \zeta_2\}$	$\hat{E}$	Gini	Shannon	Tsallis ($k = 1, q = 2$)	Rényi ($\alpha = 2$)	LI (%)	HI (%)
BreastCancer	{2.0; 1.2; 0.7; 1.0; 1}	0.971 ± 0.009	0.960 ± 0.014	0.955 ± 0.011	0.956 ± 0.008	0.961 ± 0.009	1.030	1.648
Iris	{0.0; 0.9; 0.7; 0.0; 0}	0.946 ± 0.027	0.932 ± 0.044	0.946 ± 0.017	0.912 ± 0.070	0.932 ± 0.044	0	3.594
Spambase	{0.0; 1.2; 2.9; 0.0; 0}	0.866 ± 0.010	0.851 ± 0.030	0.850 ± 0.015	0.842 ± 0.029	0.835 ± 0.035	1.732	3.580
Statlog	{0.0; 1.2; 2.9; 0.0; 0}	0.827 ± 0.034	0.773 ± 0.037	0.721 ± 0.009	0.759 ± 0.023	0.785 ± 0.053	5.079	12.817
Wine Quality	{0.0; 1.2; 2.7; 0.0; 0}	0.526 ± 0.013	0.450 ± 0.017	0.330 ± 0.031	0.411 ± 0.016	0.454 ± 0.019	13.688	37.262
Students	{2.0; 2.4; 2.0; 0.0; 0}	0.677 ± 0.021	0.639 ± 0.029	0.649 ± 0.012	0.648 ± 0.018	0.658 ± 0.012	2.806	5.613
mfeat-morphological	{1.0; 1.1; 2.4; 0.0; 0}	0.649 ± 0.038	0.625 ± 0.020	0.531 ± 0.081	0.541 ± 0.058	0.559 ± 0.037	3.698	18.182
diabetes	{0.0; 1.5; 1.5; 0.0; 0}	0.665 ± 0.053	0.619 ± 0.037	0.626 ± 0.041	0.621 ± 0.042	0.652 ± 0.050	1.955	6.917
wdbc	{2.0; 0.0; 0.0; 0.0; 1}	<u>0.963 ± 0.022</u>	0.941 ± 0.026	0.939 ± 0.020	0.956 ± 0.023	0.963 ± 0.022	0	1.572
wilt	{2.0; 0.7; 0.6; 1.0; 0}	0.922 ± 0.003	0.920 ± 0.001	0.920 ± 0.001	0.920 ± 0.000	0.920 ± 0.001	0.217	0.217
Titanic	{1.0; 0.3; 2.0; 0.0; 0}	0.703 ± 0.028	0.648 ± 0.049	0.620 ± 0.061	0.669 ± 0.070	0.657 ± 0.044	4.836	11.807
dataset_31_credit-g	{2.0; 0.7; 1.2; 0.0; 0}	0.669 ± 0.049	0.659 ± 0.018	0.623 ± 0.033	0.604 ± 0.008	0.636 ± 0.025	1.495	9.716
phoneme	{0.0; 0.6; 2.0; 0.0; 1}	0.816 ± 0.010	0.798 ± 0.028	0.800 ± 0.016	0.787 ± 0.010	0.806 ± 0.022	1.225	3.554
synth_1	{1.0; 1.6; 0.9; 1.0; 0}	0.980 ± 0.040	0.950 ± 0.045	0.930 ± 0.025	0.895 ± 0.111	0.950 ± 0.064	3.061	8.673
synth_2	{0.0; 2.0; 1.0; 0.0; 0}	<u>0.445 ± 0.122</u>	0.445 ± 0.122	0.349 ± 0.099	0.392 ± 0.105	0.424 ± 0.169	0	19.770
synth_3	{0.0; 1.0; 0.7; 0.0; 0}	0.952 ± 0.019	0.941 ± 0.028	0.945 ± 0.034	0.934 ± 0.028	0.926 ± 0.018	0.735	2.731
synth_4	{2.0; 1.2; 2.9; 0.0; 0}	0.801 ± 0.044	0.756 ± 0.023	0.754 ± 0.072	0.768 ± 0.040	0.773 ± 0.050	3.496	5.868
synth_5	{0.0; 1.1; 0.6; 0.0; 0}	0.659 ± 0.016	0.601 ± 0.033	0.607 ± 0.015	0.631 ± 0.040	0.620 ± 0.024	4.249	8.801

Table S6. Results obtained for the RF algorithm with parameters:

$\{max_depth = 10; n_estimators = 25\}$. Shown are the mean and standard deviation for the weighted F_1 -score over the 5 folds (for each dataset), computed for the proposed method $\hat{E}$, Gini, Shannon, Tsallis, and Rényi. The best parameters found for $\hat{E}$ are shown. The best values are in bold.

Dataset name	$\{\alpha; \beta; \gamma; \zeta_1; \zeta_2\}$	$\hat{E}$	Gini	Shannon	Tsallis ($k = 1, q = 2$)	Rényi ($\alpha = 2$)	LI (%)	HI (%)
BreastCancer	{0.0; 0.1; 0.5; 1.0; 0}	0.973 ± 0.012	0.964 ± 0.016	0.967 ± 0.009	0.967 ± 0.009	0.961 ± 0.012	0.617	1.233
Iris	{0.0; 0.5; 2.9; 0.0; 0}	0.960 ± 0.013	0.932 ± 0.057	0.939 ± 0.041	0.940 ± 0.025	0.946 ± 0.017	1.458	2.917
Spambase	{1.0; 2.0; 1.3; 0.0; 0}	0.890 ± 0.017	0.874 ± 0.013	0.886 ± 0.015	0.879 ± 0.026	0.879 ± 0.014	0.449	1.798
Statlog	{1.0; 2.8; 1.2; 0.0; 0}	0.847 ± 0.060	0.747 ± 0.034	0.723 ± 0.014	0.758 ± 0.021	0.747 ± 0.021	10.508	14.640
Wine Quality	{0.0; 1.2; 2.9; 0.0; 0}	0.540 ± 0.017	0.441 ± 0.017	0.342 ± 0.039	0.379 ± 0.046	0.439 ± 0.021	18.333	36.667
Students	{0.0; 2.8; 1.1; 0.0; 0}	0.669 ± 0.015	0.651 ± 0.008	0.645 ± 0.014	0.648 ± 0.011	0.652 ± 0.024	2.541	3.587
mfeat-morphological	{2.0; 0.7; 2.9; 0.0; 0}	0.662 ± 0.023	0.631 ± 0.043	0.610 ± 0.023	0.624 ± 0.017	0.648 ± 0.027	2.115	7.855
diabetes	{2.0; 2.0; 1.5; 1.0; 0}	0.669 ± 0.048	0.653 ± 0.019	0.633 ± 0.031	0.648 ± 0.022	0.634 ± 0.028	2.392	5.381
wdbc	{0.0; 0.6; 2.8; 0.0; 0}	0.963 ± 0.015	0.959 ± 0.015	0.958 ± 0.007	0.950 ± 0.016	0.952 ± 0.019	0.415	1.350
wilt	{0.0; 1.3; 1.4; 0.0; 1}	0.920 ± 0.001	0.920 ± 0.000	0.920 ± 0.001	0.920 ± 0.001	0.920 ± 0.001	0	0
Titanic	{2.0; 0.1; 2.0; 0.0; 0}	0.736 ± 0.030	0.697 ± 0.040	0.690 ± 0.022	0.687 ± 0.019	0.661 ± 0.062	5.299	10.190
dataset_31_credit-g	{1.0; 0.9; 2.8; 0.0; 0}	0.613 ± 0.025	0.603 ± 0.023	0.595 ± 0.010	0.604 ± 0.019	0.596 ± 0.015	1.468	2.936
phoneme	{2.0; 0.2; 1.2; 1.0; 0}	0.828 ± 0.013	0.816 ± 0.006	0.821 ± 0.014	0.812 ± 0.010	0.818 ± 0.008	0.845	1.932
synth_1	{0.0; 2.7; 1.6; 0.0; 0}	1.000 ± 0.000	0.970 ± 0.025	0.940 ± 0.038	0.950 ± 0.032	0.940 ± 0.038	3.000	6.000
synth_2	{2.0; 2.2; 2.1; 1.0; 0}	0.447 ± 0.105	0.340 ± 0.131	0.441 ± 0.205	0.338 ± 0.120	0.340 ± 0.123	1.342	24.385
synth_3	{1.0; 2.7; 0.6; 0.0; 0}	0.961 ± 0.007	0.956 ± 0.023	0.950 ± 0.016	0.950 ± 0.016	0.943 ± 0.026	0.520	1.873
synth_4	{2.0; 1.2; 1.3; 0.0; 0}	0.826 ± 0.059	0.788 ± 0.023	0.812 ± 0.043	0.787 ± 0.019	0.807 ± 0.057	1.695	4.722
synth_5	{2.0; 0.0; 0.8; 1.0; 0}	0.704 ± 0.051	0.665 ± 0.039	0.657 ± 0.024	0.649 ± 0.028	0.673 ± 0.057	4.403	7.812

Table S7. Results obtained for the RF algorithm with parameters:

$\{max_depth = 10; n_estimators = 50\}$. Shown are the mean and standard deviation for the weighted F_1 -score over the 5 folds (for each dataset), computed for the proposed method $\hat{E}$, Gini, Shannon, Tsallis, and Rényi. The best parameters found for $\hat{E}$ are shown. The best values are in bold.

Dataset name	$\{\alpha; \beta; \gamma; \zeta_1; \zeta_2\}$	$\hat{E}$	Gini	Shannon	Tsallis ($k = 1, q = 2$)	Rényi ($\alpha = 2$)	LI (%)	HI (%)
BreastCancer	{0.0; 0.3; 1.5; 1.0; 0}	0.973 ± 0.009	0.960 ± 0.016	0.971 ± 0.012	0.969 ± 0.007	0.967 ± 0.007	0.206	1.336
Iris	{2.0; 0.2; 0.2; 1.0; 1}	0.953 ± 0.016	0.918 ± 0.047	0.939 ± 0.033	0.926 ± 0.033	0.933 ± 0.043	1.469	3.673
Spambase	{2.0; 2.6; 1.7; 0.0; 0}	0.881 ± 0.011	0.877 ± 0.009	0.872 ± 0.019	0.880 ± 0.007	0.869 ± 0.016	0.114	1.362
Statlog	{0.0; 1.1; 1.8; 0.0; 0}	0.836 ± 0.018	0.753 ± 0.021	0.737 ± 0.009	0.758 ± 0.013	0.759 ± 0.018	9.211	11.842
Wine Quality	{2.0; 1.2; 2.3; 0.0; 0}	0.552 ± 0.011	0.452 ± 0.018	0.338 ± 0.026	0.407 ± 0.017	0.446 ± 0.017	18.116	38.768
Students	{0.0; 1.4; 2.3; 0.0; 0}	0.659 ± 0.011	0.650 ± 0.013	0.648 ± 0.016	0.642 ± 0.010	0.655 ± 0.012	0.607	2.580
mfeat-morphological	{2.0; 1.6; 1.6; 1.0; 0}	0.681 ± 0.029	0.651 ± 0.022	0.629 ± 0.027	0.637 ± 0.021	0.651 ± 0.031	4.405	7.636
diabetes	{1.0; 1.2; 2.9; 0.0; 0}	0.653 ± 0.035	0.632 ± 0.035	0.612 ± 0.019	0.633 ± 0.036	0.650 ± 0.035	0.459	6.279
wdbc	{0.0; 0.9; 2.8; 0.0; 1}	0.964 ± 0.019	0.945 ± 0.019	0.948 ± 0.021	0.950 ± 0.019	0.954 ± 0.013	1.037	1.971
wilt	{1.0; 2.0; 1.8; 0.0; 1}	0.920 ± 0.001	0.920 ± 0.001	0.920 ± 0.001	0.920 ± 0.001	0.920 ± 0.001	0	0
Titanic	{2.0; 1.6; 0.0; 1.0; 0}	0.700 ± 0.039	0.683 ± 0.024	0.684 ± 0.024	0.684 ± 0.022	0.681 ± 0.023	2.286	2.714
dataset_31_credit-g	{0.0; 2.0; 2.6; 0.0; 0}	0.600 ± 0.022	0.594 ± 0.015	0.590 ± 0.010	0.584 ± 0.009	0.587 ± 0.013	1.000	2.667
phoneme	{0.0; 1.2; 2.9; 0.0; 0}	0.832 ± 0.010	0.810 ± 0.010	0.820 ± 0.017	0.822 ± 0.009	0.815 ± 0.005	1.202	2.644
synth_1	{2.0; 1.4; 0.6; 0.0; 1}	1.000 ± 0.000	0.960 ± 0.020	0.970 ± 0.040	0.960 ± 0.058	0.950 ± 0.032	3.000	5.000
synth_2	{0.0; 1.6; 1.6; 0.0; 0}	0.477 ± 0.139	0.353 ± 0.105	0.302 ± 0.069	0.359 ± 0.102	0.407 ± 0.161	14.675	36.688
synth_3	{2.0; 0.2; 0.9; 1.0; 0}	0.960 ± 0.016	0.952 ± 0.022	0.952 ± 0.019	0.958 ± 0.021	0.952 ± 0.023	0.208	0.833
synth_4	{2.0; 0.6; 1.0; 0.0; 0}	0.836 ± 0.037	0.825 ± 0.041	0.807 ± 0.053	0.828 ± 0.060	0.822 ± 0.044	0.957	3.469
synth_5	{1.0; 2.1; 0.7; 0.0; 0}	0.719 ± 0.028	0.689 ± 0.018	0.653 ± 0.042	0.683 ± 0.050	0.687 ± 0.030	4.172	9.179

Table S8. Results obtained for the RF algorithm with parameters:

$\{max_depth = 25; n_estimators = 10\}$. Shown are the mean and standard deviation for the weighted F_1 -score over the 5 folds (for each dataset), computed for the proposed method $\hat{E}$, Gini, Shannon, Tsallis, and Rényi. The best parameters found for $\hat{E}$ are shown. The best values are in bold.

Dataset name	$\{\alpha; \beta; \gamma; \zeta_1; \zeta_2\}$	$\hat{E}$	Gini	Shannon	Tsallis ($k = 1, q = 2$)	Rényi ($\alpha = 2$)	LI (%)	HI (%)
BreastCancer	{1.0; 2.8; 0.7; 1.0; 1}	0.967 ± 0.015	0.961 ± 0.009	0.954 ± 0.010	0.958 ± 0.020	0.963 ± 0.010	0.414	1.344
Iris	{2.0; 2.9; 0.0; 2.0; 1}	0.953 ± 0.035	0.912 ± 0.070	0.939 ± 0.026	0.939 ± 0.026	0.939 ± 0.034	1.469	4.302
Spambase	{1.0; 1.1; 2.1; 0.0; 1}	0.872 ± 0.017	0.840 ± 0.025	0.858 ± 0.012	0.843 ± 0.038	0.834 ± 0.027	1.606	4.358
Statlog	{2.0; 2.2; 1.2; 0.0; 0}	0.919 ± 0.034	0.836 ± 0.043	0.774 ± 0.022	0.783 ± 0.048	0.814 ± 0.019	9.032	15.778
Wine Quality	{0.0; 1.2; 2.0; 0.0; 0}	0.603 ± 0.005	0.480 ± 0.023	0.403 ± 0.062	0.454 ± 0.020	0.476 ± 0.008	20.398	33.167
Students	{2.0; 2.8; 1.6; 1.0; 0}	0.676 ± 0.021	0.647 ± 0.019	0.650 ± 0.015	0.654 ± 0.028	0.643 ± 0.024	3.254	4.882
mfeat-morphological	{2.0; 1.4; 2.7; 1.0; 0}	0.638 ± 0.022	0.605 ± 0.035	0.588 ± 0.028	0.614 ± 0.022	0.595 ± 0.040	3.762	7.837
diabetes	{0.0; 1.7; 2.9; 0.0; 0}	0.671 ± 0.031	0.618 ± 0.028	0.602 ± 0.019	0.635 ± 0.035	0.632 ± 0.038	5.365	10.283
wdbc	{1.0; 2.4; 0.1; 1.0; 1}	0.954 ± 0.033	0.941 ± 0.018	0.947 ± 0.018	0.932 ± 0.009	0.952 ± 0.015	0.210	2.306
wilt	{0.0; 2.8; 1.0; 0.0; 0}	0.924 ± 0.005	0.923 ± 0.002	0.924 ± 0.004	0.923 ± 0.002	0.924 ± 0.004	0	0.108
Titanic	{1.0; 2.0; 1.8; 0.0; 1}	0.733 ± 0.035	0.662 ± 0.069	0.697 ± 0.042	0.681 ± 0.023	0.665 ± 0.061	4.911	9.686
dataset_31_credit-g	{0.0; 1.2; 2.9; 0.0; 0}	0.692 ± 0.052	0.651 ± 0.041	0.616 ± 0.018	0.635 ± 0.039	0.669 ± 0.030	3.324	10.983
phoneme	{1.0; 1.6; 1.5; 0.0; 0}	0.811 ± 0.013	0.784 ± 0.020	0.785 ± 0.018	0.790 ± 0.021	0.777 ± 0.026	2.589	4.192
synth_1	{2.0; 0.9; 0.4; 0.0; 1}	0.990 ± 0.020	0.950 ± 0.032	0.929 ± 0.052	0.919 ± 0.069	0.950 ± 0.045	4.040	7.172
synth_2	{0.0; 2.6; 0.6; 0.0; 0}	0.427 ± 0.070	0.362 ± 0.064	0.373 ± 0.095	0.339 ± 0.159	0.309 ± 0.166	12.646	27.635
synth_3	{0.0; 0.2; 1.7; 0.0; 1}	0.947 ± 0.030	0.938 ± 0.034	0.935 ± 0.016	0.930 ± 0.014	0.933 ± 0.027	0.950	1.795
synth_4	{2.0; 0.8; 1.6; 0.0; 0}	0.793 ± 0.050	0.759 ± 0.032	0.755 ± 0.058	0.775 ± 0.055	0.750 ± 0.020	2.270	5.422
synth_5	{2.0; 1.0; 2.9; 0.0; 0}	0.665 ± 0.031	0.565 ± 0.033	0.597 ± 0.039	0.587 ± 0.059	0.631 ± 0.037	5.113	15.038

Table S9. Results obtained for the RF algorithm with parameters:

$\{max_depth = 25; n_estimators = 25\}$. Shown are the mean and standard deviation for the weighted F_1 -score over the 5 folds (for each dataset), computed for the proposed method $\hat{E}$, Gini, Shannon, Tsallis, and Rényi. The best parameters found for $\hat{E}$ are shown. The best values are in bold.

Dataset name	$\{\alpha; \beta; \gamma; \zeta_1; \zeta_2\}$	$\hat{E}$	Gini	Shannon	Tsallis ($k = 1, q = 2$)	Rényi ($\alpha = 2$)	LI (%)	HI (%)
BreastCancer	{0.0; 1.2; 1.3; 0.0; 1}	0.973 ± 0.009	0.967 ± 0.015	0.966 ± 0.008	0.963 ± 0.012	0.963 ± 0.015	0.617	1.028
Iris	{2.0; 1.1; 1.7; 0.0; 1}	0.953 ± 0.034	0.919 ± 0.051	0.918 ± 0.047	0.918 ± 0.060	0.953 ± 0.017	0	3.673
Spambase	{0.0; 1.2; 0.2; 0.0; 0}	0.889 ± 0.008	0.863 ± 0.022	0.874 ± 0.012	0.875 ± 0.015	0.868 ± 0.008	1.575	2.925
Statlog	{2.0; 1.5; 0.7; 0.0; 0}	0.901 ± 0.058	0.854 ± 0.037	0.790 ± 0.019	0.827 ± 0.022	0.853 ± 0.026	5.216	12.320
Wine Quality	{0.0; 1.2; 1.9; 0.0; 0}	0.620 ± 0.011	0.479 ± 0.009	0.370 ± 0.041	0.416 ± 0.039	0.471 ± 0.013	22.742	40.323
Students	{1.0; 2.0; 1.1; 0.0; 0}	0.673 ± 0.020	0.653 ± 0.022	0.658 ± 0.009	0.658 ± 0.012	0.661 ± 0.021	1.783	2.972
mfeat-morphological	{0.0; 1.4; 1.4; 0.0; 0}	0.651 ± 0.016	0.625 ± 0.016	0.603 ± 0.008	0.638 ± 0.032	0.603 ± 0.045	1.997	7.373
diabetes	{0.0; 0.6; 0.9; 0.0; 0}	0.691 ± 0.027	0.664 ± 0.022	0.641 ± 0.034	0.642 ± 0.040	0.657 ± 0.036	3.907	7.236
wdbc	{1.0; 0.3; 0.3; 0.0; 0}	0.963 ± 0.025	0.956 ± 0.019	0.963 ± 0.010	0.959 ± 0.021	0.954 ± 0.019	0	0.935
wilt	{2.0; 2.7; 0.9; 1.0; 0}	0.923 ± 0.002	0.923 ± 0.002	0.922 ± 0.002	0.922 ± 0.002	0.923 ± 0.002	0	0.108
Titanic	{0.0; 2.3; 1.5; 0.0; 0}	0.726 ± 0.037	0.684 ± 0.026	0.680 ± 0.022	0.681 ± 0.022	0.690 ± 0.022	4.959	6.336
dataset_31_credit-g	{1.0; 2.6; 2.1; 0.0; 0}	0.622 ± 0.034	0.595 ± 0.013	0.614 ± 0.026	0.597 ± 0.022	0.608 ± 0.027	1.286	4.341
phoneme	{2.0; 2.3; 0.0; 0.0; 0}	0.829 ± 0.005	0.807 ± 0.011	0.814 ± 0.010	0.812 ± 0.017	0.820 ± 0.017	1.086	2.654
synth_1	{0.0; 0.6; 1.8; 0.0; 1}	0.990 ± 0.020	0.970 ± 0.025	0.970 ± 0.025	0.960 ± 0.038	0.970 ± 0.025	2.020	3.030
synth_2	{2.0; 1.5; 0.0; 0.0; 0}	0.467 ± 0.097	0.425 ± 0.100	0.339 ± 0.112	0.443 ± 0.153	0.376 ± 0.100	5.139	27.409
synth_3	{2.0; 0.1; 0.0; 1.0; 1}	0.959 ± 0.018	0.952 ± 0.022	0.949 ± 0.019	0.951 ± 0.017	0.944 ± 0.018	0.730	1.564
synth_4	{1.0; 2.3; 1.8; 0.0; 0}	0.833 ± 0.040	0.804 ± 0.036	0.805 ± 0.026	0.809 ± 0.043	0.800 ± 0.045	2.881	3.962
synth_5	{2.0; 1.2; 1.8; 1.0; 0}	0.694 ± 0.035	0.664 ± 0.035	0.673 ± 0.049	0.670 ± 0.062	0.643 ± 0.028	3.026	7.349

Table S10. Results obtained for the RF algorithm with parameters:

$\{max_depth = 25; n_estimators = 50\}$. Shown are the mean and standard deviation for the weighted F_1 -score over the 5 folds (for each dataset), computed for the proposed method $\hat{E}$, Gini, Shannon, Tsallis, and Rényi. The best parameters found for $\hat{E}$ are shown. The best values are in bold. Underlined values mean that the search did not manage to converge to a nonclassical parametrization.

Dataset name	$\{\alpha; \beta; \gamma; \zeta_1; \zeta_2\}$	$\hat{E}$	Gini	Shannon	Tsallis ($k = 1, q = 2$)	Rényi ($\alpha = 2$)	LI (%)	HI (%)
BreastCancer	{1.0; 0.8; 2.9; 0.0; 1}	0.971 ± 0.013	0.964 ± 0.009	0.966 ± 0.017	0.963 ± 0.009	0.960 ± 0.011	0.515	1.133
Iris	{1.0; 1.7; 1.2; 1.0; 0}	0.953 ± 0.016	0.933 ± 0.030	0.939 ± 0.041	0.946 ± 0.027	0.939 ± 0.033	0.735	2.099
Spambase	{0.0; 1.2; 2.9; 0.0; 0}	0.881 ± 0.006	0.873 ± 0.016	0.871 ± 0.009	0.859 ± 0.015	0.877 ± 0.011	0.454	2.497
Statlog	{2.0; 2.8; 0.6; 1.0; 0}	0.928 ± 0.036	0.824 ± 0.027	0.787 ± 0.022	0.811 ± 0.033	0.830 ± 0.020	10.560	15.194
Wine Quality	{2.0; 2.0; 2.0; 0.0; 0}	0.634 ± 0.012	0.479 ± 0.016	0.377 ± 0.034	0.450 ± 0.016	0.486 ± 0.010	23.344	40.536
Students	{1.0; 2.1; 0.6; 0.0; 0}	0.669 ± 0.013	0.649 ± 0.013	0.654 ± 0.012	0.659 ± 0.013	0.666 ± 0.014	0.448	2.990
mfeat-morphological	{1.0; 2.0; 2.5; 0.0; 0}	0.667 ± 0.020	0.643 ± 0.021	0.610 ± 0.021	0.639 ± 0.019	0.643 ± 0.023	3.598	8.546
diabetes	{0.0; 2.8; 2.0; 0.0; 0}	0.671 ± 0.037	0.642 ± 0.025	0.646 ± 0.028	0.637 ± 0.015	0.654 ± 0.037	2.534	5.067
wdbc	{1.0; 0.8; 1.6; 0.0; 0}	0.961 ± 0.012	0.957 ± 0.027	0.959 ± 0.025	0.949 ± 0.028	0.959 ± 0.024	0.208	1.249
wilt	{1.0; 0.5; 0.9; 1.0; 0}	0.923 ± 0.002	0.922 ± 0.002	0.923 ± 0.002	0.923 ± 0.002	0.922 ± 0.002	0	0.108
Titanic	{2.0; 0.7; 0.0; 0.0; 0}	0.709 ± 0.025	0.686 ± 0.021	0.690 ± 0.022	0.685 ± 0.023	0.688 ± 0.018	2.680	3.385
dataset_31_credit-g	{1.0; 1.2; 0.2; 0.0; 0}	0.615 ± 0.017	0.598 ± 0.008	0.598 ± 0.013	0.602 ± 0.013	0.602 ± 0.018	2.114	2.764
phoneme	{2.0; 0.6; 0.8; 1.0; 0}	0.832 ± 0.012	0.820 ± 0.007	0.827 ± 0.009	0.817 ± 0.007	0.817 ± 0.005	0.601	1.803
synth_1	{2.0; 1.4; 1.1; 1.0; 1}	1.000 ± 0.000	0.960 ± 0.038	0.960 ± 0.038	0.950 ± 0.000	0.990 ± 0.020	1.000	5.000
synth_2	{2.0; 0.0; 0.0; 0.0; 0}	0.453 ± 0.126	0.370 ± 0.159	0.318 ± 0.127	0.453 ± 0.126	0.367 ± 0.161	0	26.389
synth_3	{2.0; 1.1; 0.0; 2.0; 0}	0.958 ± 0.023	0.954 ± 0.022	0.954 ± 0.025	0.945 ± 0.023	0.944 ± 0.027	0.418	1.461
synth_4	{2.0; 1.0; 0.7; 1.0; 0}	0.835 ± 0.040	0.818 ± 0.037	0.823 ± 0.043	0.817 ± 0.045	0.821 ± 0.052	1.437	2.156
synth_5	{2.0; 1.8; 0.0; 2.0; 0}	0.721 ± 0.016	0.688 ± 0.023	0.685 ± 0.020	0.693 ± 0.045	0.697 ± 0.015	3.329	4.993