ORIGINAL PAPER

AN ENHANCED CLASSIFICATION SYSTEM BASED ON KERNEL PRINCIPAL COMPONENT ANALYSIS AND DATA COMPLEXITY MEASURES

FATIH SAĞLAM[1], EMRE DÜNDER[1], MEHMET ALI CENGIZ[1]

Manuscript received: 23.09.2022; Accepted paper: 08.04.2023;
Published online: 30.06.2023.

Abstract. *Principal component analysis is commonly used as a pre-step before employing a classifier to avoid the negative effect of the dimensionality and multicollinearity. The performance of a classifier is severely affected by the deviations from the linearity of the data structure and noisy samples. In this paper, we propose a new classification system that overcomes the drawback of these crucial problems, simultaneously. Our proposal is relying on the kernel principal component analysis with a proper parameter selection approach with data complexity measures. According to the empirical results, F1, T2 and T3 in AUC, T3 in GMEAN and T2 and T3 in MCC performed better than classical and other complexity measures. Comparison of classifiers showed that Radial SVM performs better in AUC, and KNN performs better in GMEAN and MCC using KPCA with complexity measures. As a result, our proposed system produces better results in various classification algorithms with respect to classical approach.*

Keywords: *classification system; data complexity; kernel principal component analysis.*

1. INTRODUCTION

Classification systems have been used intensively to obtain the prediction of categorical responses for several areas such as bioinformatics [1-2], chemistry [3-4], medicine [5-7], ecology [8], and finance [9]. Due to the importance, the researchers aim to increase the accuracy of the classification results by improving the computational part of the classifiers.

During the classification process, it is an alternative way to employ PCA prior to the classifiers. The classifiers train the data using the compressed versions of the full features [10]. PCA is a preferable technique that extracts relevant components reflecting the variability of the full data [11]. It has attracted much attention to use PCA within the classification algorithms [12-17].

The classical PCA was designed to handle the linear projection on data [18]. In mostly cases, the data rely on a nonlinear structure. Classification systems are muchly sensitive to the structure of the data and it is necessary to use a method that can separate the data nonlinearly. Kernel PCA (KPCA) is a useful choice for such cases due to the nonlinear mapping capability [16,19]. KPCA was used to improve the classification performance of the models with several classifiers [20-23]. KPCA includes a kernel function and the proper selection of the parameters of this function should be carried out optimally [24-25]. However, the researchers

[1] Ondokuz Mayıs University, Faculty of Sciences, Department of Statistics, 55420 Samsun, Turkey.
E-mail: fatih.saglam@omu.edu.tr; emre.dunder@omu.edu.tr; macengiz@omu.edu.tr.

ignored to select the optimal parameters of the kernel functions within KPCA related to the classifiers.

Aforementioned researches using KPCA do not take into account of data complexity to make dataset more separatable. In this paper, we propose a novel approach to perform parameter selection in KPCA using the data complexity measures [26]. Our purpose is to select the kernel hyperparameters so as to achieve the maximal separation of the classes in the feature space. Hence, we aimed to tackle the dimensionality by separating classes harmonious with the distribution of the features. To the best of our knowledge, this study is the first attempt to present an objective way of the parameter selection in KPCA for the classification problems.

The paper is organized as follows: Section 2 gives a summary about theoretical background of KPCA. Section 3 introduces our proposed approach together with data complexities. Section 4 provides the experimental results that include the comparison of the traditional selection and our proposed way on several real data sets. Section 5 summarizes and discusses the obtained performance results.

2. KERNEL PRINCIPAL COMPONENT ANALYSIS

Let $\phi(\mathbf{x})$ be a nonlinear transformation of $\mathbf{x}$ with n data point and $\mathbf{x}_i$ a d-dimensional vector. This transformation maps $\mathbf{x}$ into a higher dimensional space ($\mathbb{R}^m$, $m > d$). Estimated standard principal components of $\phi(\mathbf{x})$ are nonlinear principal components of $\mathbf{x}$. Assuming $\sum_{i=1}^{n} \mathbf{x}_i = 0$ and $\sum_{i=1}^{n} \phi(\mathbf{x}_i) = 0$, we can write

$$S = \frac{1}{n}\sum_{i=1}^{n} \phi(\mathbf{x}_i)\phi(\mathbf{x}_i)^{\top}. \tag{1}$$

S is the $d \times d$ covariance matrix estimation of $\phi(\mathbf{x})$. Eigenvalues and eigen vectors of S satisfying

$$S\mathbf{V} = \lambda\mathbf{V}, \qquad \lambda \geq 0 \tag{2}$$

need to be calculated. From Equation (1) and (2), we can write

$$\frac{1}{n}\sum_{i=1}^{n} \phi(\mathbf{x}_i)\{\phi(\mathbf{x}_i)^{\top}\mathbf{V}\} = \lambda\mathbf{V} \tag{3}$$

asnd

$$V = \sum_{i}^{n} \alpha_i \phi(\mathrm{x}_i). \tag{4}$$

Substituting $\mathbf{V}$ in Equation (3) with Equation (4), we have

$$\frac{1}{n}\sum_{i=1}^{n} \phi(\mathbf{x}_i)\left\{\phi(\mathbf{x}_i)^{\top}\sum_{i}^{n} \alpha_i \phi(\mathbf{x}_i)\right\} = \lambda\sum_{i}^{n} \alpha_i \phi(\mathbf{x}_i). \tag{5}$$

We can define the function as

$$K(\mathbf{x}_i, \mathbf{x}_j) = \phi(\mathbf{x}_i)^\top \phi(\mathbf{x}_i). \tag{6}$$

Using Equation (5) and (6),

$$\begin{gathered} \frac{1}{n}\sum_{i=1}^{n} K(\mathrm{x}_i, \mathrm{x}_j)\sum_{j=1}^{n}\alpha_i K(\mathrm{x}_i, \mathrm{x}_j) = \lambda\sum_{i=1}^{n}\alpha_i K(\mathrm{x}_i, \mathrm{x}_j) \\ \mathrm{K}^2\alpha = \lambda n \mathrm{K}\alpha \\ \mathrm{K}\alpha = \lambda n \alpha \end{gathered} \tag{7}$$

Here α is n dimensional column vector consisting of α_i-s. Solving Equation (7) will allow as to calculate kernel principal components using

$$y(\mathbf{x}) = \phi(\mathbf{x})^\top \mathbf{V} = \sum_{i=1}^{n}\alpha_i K(\mathbf{x}, \mathbf{x}_i).$$

Kernel approach allows as to obtain principal components without calculating $\phi(\mathbf{x})$.

3. PARAMETER SELECTION BASED ON COMPLEXITY MEASURE

Kernels usually depend on a specific parameter. The selection of these parameters significantly affects the results. Therefore, correct parameter selection is important. We propose to use different complexity measures to determine the efficient kernel hyperparameter. The algorithm in Table 1 describes the approach for selecting optimal kernel hyperparameter. The aim of this approach is project dataset into higher dimensional space with lowest possible complexity.

Table 1. Algorithm for obtaining KPCA radial kernel matrix with automated σ selection

INPUT:	X_{train}, σ_{grid}, C (selected complexity measure)
1.	For σ in σ_{grid} X_{kpca} is calculated using σ with radial kernel C is calculated for X_{kpca} and mapped into C_{all}
2.	σ_{opt} for optimum C_{all} is determined.
OUTPUT:	X_{kpca} with σ_{opt}

3.1. COMPLEXITY MEASURES

Lorena et al. [27] categorizes complexity into six category: overlapping, linearity, neighborhood, network, dimensionality, and class imbalance measures. KPCA is used to handle all these categories except class imbalance. Therefore, we used the first five category.

The overlapping measures examines the separation of classes. The problem is assumed to be simpler when at least one variable has discriminative values. Neighborhood measures

examines neighbors of data points and shape of decision boundary. Linearity measures measures the linear separability of the dataset. The problem is assumed to be simpler if it is linearly separable [28]. Network measures extracts structural information from the data set using graph methods [27]. Dimensionality measures are the measurement of data sparsity. They try to find if there are areas with low density, which is assumed to be hard to classify [29]. Table 2 gives the complexity measures used in the study.

Figure 1 shows an example of KPCA components constructed using different hyper parameter selection methods on banana dataset. Let us take an example of linear classifier. In banana dataset, it is almost impossible to discriminate classes using a linear classifer without KPCA. When we look at F2 kernel components, it is safe to tell that it is easier to model compared to no KPCA. Of course, we just see the first two dimensions. Linear separability is hard to see from high dimensions. The main point is that it is important to select the right complexity measure and it requires the researcher to have a good understanding of the dataset.

Table 2. Complexity measure used in the study

Category	Measure	Summary
Overlapping	Maximum Fisher's Discriminant Ratio (F1)	Largest discriminant ratio of all the variables.
	Volume of the overlapping region (F2)	The overlap of the distributions of the variables based on minimum and maximum values of variables for each class.
	Collective feature efficiency (F4)	A measurement of overlapping area of all variables.
Neighborhood	Fraction of borderline points (N1)	The percentage of vertexes incident to edges connecting examples of opposite classes in a Minimum Spanning Tree.
	Local Set Average Cardinality (LSC)	The set of points from the dataset whose distance of each example is smaller than the distance from the examples of the different class.
Linearity	Sum of the error distance by linear programming (L1)	The sum of the distances of incorrectly classified examples to a linear boundary used in linear SVM.
	Error rate of linear classifier (L2)	The error rate of the linear SVM classifier.
Network	Average Density of the network (Density)	The number of edges in the graph, divided by the maximum nur of edges between pairs of data points.
	Clustering coefficient (ClsCoef)	The average clustering tendency of the vertexes by the ratio of existent edges between its neighbors and the total number of edges that could possibly exist between them.
	Hubs score (Hubs)	A measure of the influence of each node of the graph
Dimensionality	Average number of points per dimension (T2)	The ratio between the number of examples and dimensionality of the dataset.
	Average number of points per PCA (T3)	The ratio between the number of examples and the number of PCA components needed to represent 95 percent variability as the base of data sparsity assessment.
	Ratio of the PCA Dimension to the Original (T4)	The proportion of relevant and the original dimensions for a dataset.

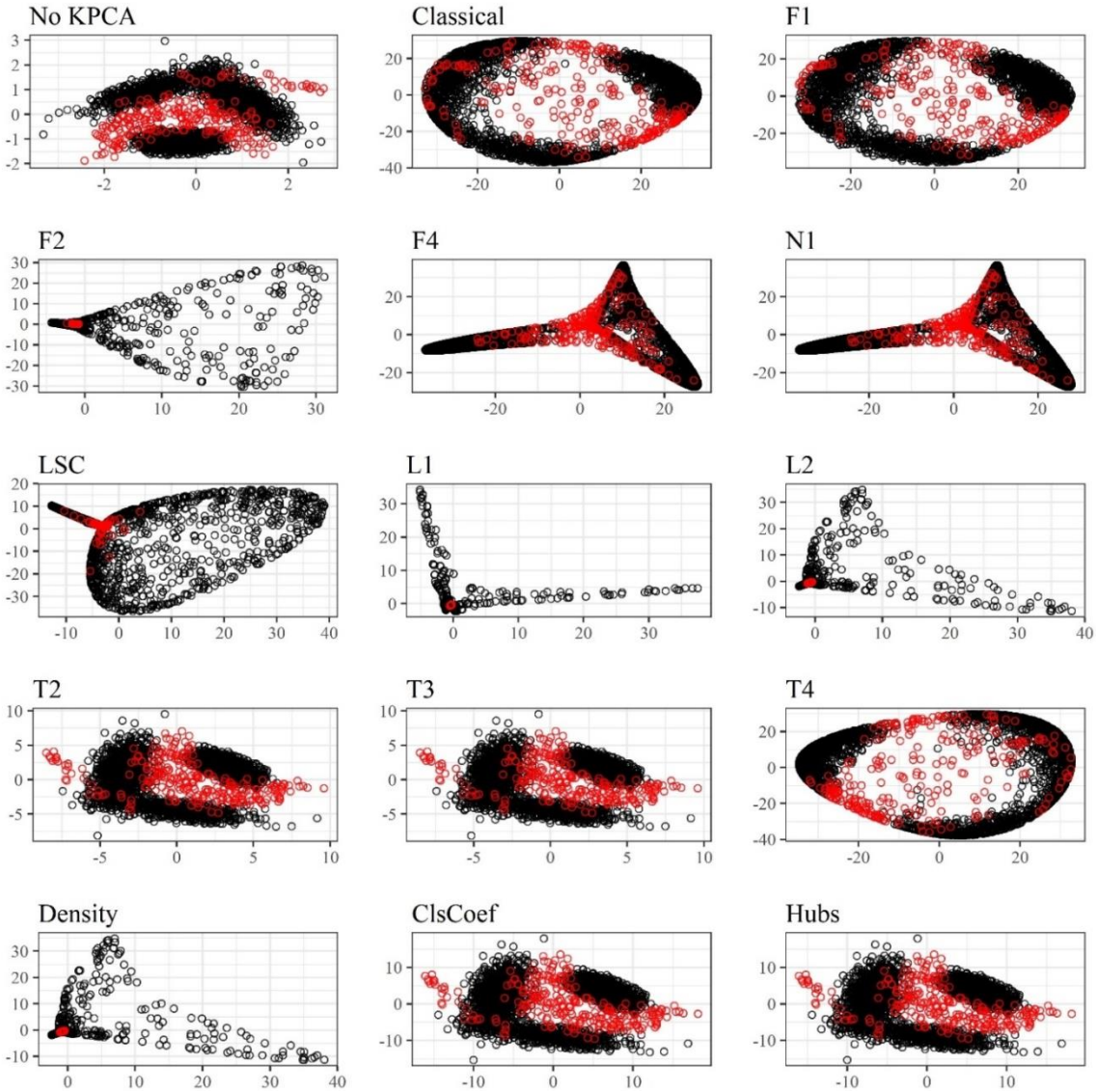

Figure 1. First two kernel principal components determined with different complexity measures for banana dataset.

4. EXPERIMENAL RESULTS

In this section we carried out extensive implementations on real data sets. On each experiment, we applied KPCA and obtained the principal component scores. Radial kernel function is used for KPCA and kernel hyperparameter is selected based on classical and proposed approaches. One of the parameter selection strategies in literature is chosen as classical approach [24]. In classical approach kernel hyperparameter is selected using

$$\sigma = 5 \cdot \underset{i}{\text{mean}}(d_i^{NN}).$$

Here, d_i^{NN} indicates distance from x_i to its nearest neighbor. Proposed approach used grid search where grid area is $\{2^{-9}, \dots, 2^{10}\}$. After obtaining the scores, we conducted 10 fold cross-validation for 10 times. The results were evaluated on test data with three well-known performance metrics: Area under the ROC curve (AUC), Matthews correlation coefficient

(MCC) and Geometric mean (G-Mean). All experiments were carried out using R software. The performance of our selection approach is examined using 13 benchmark data sets. The general properties of the data sets are given in Table 3. These datasets are easy to obtain from UCI and KEEL dataset repositories. Here, n is the number of samples, p is the number of features, and imbalance rate is $\frac{n^{majority}}{n^{minority}}$.

Table 3. Datasets used in the study

Datasets	n	p	Imbalance Rate
banana	2640	2	9.000
banknote	200	6	1.000
ecoli1	336	5	3.364
glass0	214	9	2.057
haberman	306	3	2.778
Ionosphere	351	32	1.786
iris0	150	4	2.000
kyphosis	81	3	3.765
newthyroid1	215	5	5.143
Vehicle1	846	18	2.881
Vehicle3	846	18	2.899
wisconsin	683	9	1.858
yeast4	1484	6	28.098

Table 4. AUC values for different parameter selection strategies

Datasets	Classifiers	Parameter Selection Strategy													
		classical	F1	F2	F4	N1	LSC	L1	L2	T2	T3	T4	Density	ClsCoef	Hubs
Banana	Naive Bayes	0.900	0.895	0.904	0.916	0.831	0.905	0.875	0.884	0.598	0.598	0.901	0.880	0.601	0.873
	Logistic regresion	0.960	0.954	0.755	0.853	0.769	0.844	0.713	0.726	0.585	0.585	0.962	0.697	0.584	0.794
	CART	0.883	0.870	0.648	0.845	0.868	0.817	0.645	0.656	0.867	0.867	0.877	0.657	0.871	0.862
	J48	0.872	0.868	0.696	0.823	0.859	0.782	0.574	0.591	0.882	0.882	0.862	0.557	0.881	0.871
	Linear SVM	0.960	0.952	0.809	0.927	0.709	0.879	0.786	0.794	0.491	0.514	0.962	0.766	0.507	0.605
	Radial SVM	0.951	0.954	0.622	0.931	0.955	0.938	0.457	0.491	0.960	0.960	0.949	0.435	0.960	0.961
	KNN	0.910	0.910	0.912	0.910	0.910	0.910	0.848	0.865	0.911	0.911	0.910	0.850	0.910	0.910
banknote	Naive Bayes	0.932	0.994	0.567	0.994	0.992	0.994	0.988	0.988	0.994	0.994	0.992	0.563	0.566	0.602
	Logistic regresion	0.003	0.002	0.004	0.002	0.002	0.002	0.004	0.004	0.002	0.002	0.002	0.004	0.004	0.009
	CART	0.786	0.942	0.553	0.942	0.942	0.942	0.946	0.946	0.942	0.942	0.969	0.548	0.550	0.562
	J48	0.794	0.954	0.542	0.953	0.953	0.954	0.955	0.955	0.954	0.954	0.956	0.539	0.539	0.550
	Linear SVM	0.003	0.999	0.014	0.999	0.999	0.999	0.997	0.997	0.999	0.999	0.997	0.004	0.004	0.009
	Radial SVM	0.003	0.999	0.006	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.005	0.005	0.014
	KNN	0.995	0.995	0.988	0.995	0.994	0.995	0.993	0.993	0.995	0.995	0.992	0.990	0.989	0.959
ecoli1	Naive Bayes	0.927	0.927	0.742	0.905	0.953	0.946	0.899	0.899	0.927	0.927	0.891	0.756	0.927	0.790
	Logistic regresion	0.951	0.928	0.872	0.892	0.954	0.952	0.728	0.728	0.928	0.928	0.834	0.848	0.928	0.877
	CART	0.883	0.883	0.667	0.880	0.905	0.896	0.795	0.795	0.883	0.883	0.890	0.642	0.883	0.680
	J48	0.868	0.879	0.608	0.817	0.900	0.901	0.759	0.759	0.879	0.879	0.776	0.645	0.879	0.694
	Linear SVM	0.941	0.927	0.128	0.932	0.949	0.949	0.574	0.569	0.927	0.927	0.931	0.419	0.927	0.479
	Radial SVM	0.924	0.896	0.129	0.922	0.940	0.945	0.873	0.873	0.896	0.896	0.916	0.407	0.896	0.476
	KNN	0.926	0.886	0.923	0.924	0.923	0.917	0.920	0.920	0.886	0.886	0.924	0.891	0.886	0.911
glass0	Naive Bayes	0.677	0.809	0.512	0.793	0.662	0.699	0.712	0.712	0.809	0.809	0.791	0.531	0.732	0.801
	Logistic regresion	0.731	0.844	0.800	0.684	0.758	0.732	0.717	0.717	0.844	0.844	0.692	0.691	0.714	0.841
	CART	0.732	0.806	0.512	0.815	0.722	0.748	0.747	0.747	0.806	0.806	0.816	0.538	0.771	0.808
	J48	0.720	0.797	0.516	0.752	0.701	0.711	0.725	0.725	0.797	0.797	0.760	0.539	0.717	0.794
	Linear SVM	0.746	0.836	0.589	0.806	0.762	0.742	0.726	0.726	0.836	0.836	0.808	0.561	0.715	0.829
	Radial SVM	0.614	0.833	0.607	0.847	0.646	0.638	0.630	0.630	0.833	0.833	0.854	0.577	0.723	0.834
	KNN	0.811	0.854	0.749	0.853	0.788	0.808	0.804	0.804	0.854	0.854	0.853	0.659	0.843	0.855
haberman	Naive Bayes	0.512	0.673	0.543	0.669	0.534	0.535	0.581	0.647	0.673	0.673	0.673	0.665	0.670	0.587
	Logistic regresion	0.605	0.668	0.610	0.560	0.535	0.559	0.606	0.569	0.668	0.668	0.669	0.671	0.664	0.607
	CART	0.533	0.622	0.538	0.637	0.503	0.517	0.498	0.530	0.622	0.622	0.621	0.608	0.625	0.514
	J48	0.530	0.600	0.509	0.534	0.523	0.515	0.529	0.540	0.600	0.600	0.598	0.592	0.599	0.520
	Linear SVM	0.578	0.628	0.550	0.566	0.508	0.534	0.560	0.575	0.600	0.619	0.625	0.602	0.620	0.573
	Radial SVM	0.569	0.670	0.571	0.658	0.498	0.534	0.529	0.461	0.670	0.670	0.670	0.641	0.667	0.536
	KNN	0.505	0.630	0.585	0.629	0.510	0.515	0.579	0.632	0.630	0.630	0.630	0.628	0.628	0.624
Ionosphere	Naive Bayes	0.517	0.953	0.749	0.953	0.913	0.949	0.934	0.934	0.953	0.953	0.938	0.585	0.756	0.943
	Logistic regresion	0.023	0.102	0.021	0.101	0.020	0.124	0.136	0.136	0.102	0.102	0.111	0.025	0.030	0.107
	CART	0.527	0.860	0.530	0.860	0.658	0.887	0.904	0.904	0.860	0.860	0.876	0.526	0.564	0.867
	J48	0.523	0.854	0.556	0.856	0.567	0.860	0.847	0.847	0.854	0.854	0.844	0.558	0.590	0.847
	Linear SVM	0.835	0.907	0.979	0.912	0.960	0.889	0.872	0.872	0.907	0.907	0.939	0.975	0.972	0.947
	Radial SVM	0.611	0.984	0.780	0.984	0.547	0.966	0.944	0.944	0.984	0.984	0.983	0.903	0.807	0.984

Datasets	Classifiers	Parameter Selection Strategy													
		classical	F1	F2	F4	N1	LSC	L1	L2	T2	T3	T4	Density	ClsCoef	Hubs
	KNN	0.572	0.891	0.871	0.891	0.911	0.888	0.884	0.884	0.891	0.891	0.890	0.623	0.864	0.891
	Naive Bayes	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.522	0.979	0.548
	Logistic regresion	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.967	1.000	0.980
	CART	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.519	0.976	0.549
iris0	J48	0.986	0.988	0.988	0.988	0.988	0.988	0.988	0.988	0.988	0.988	0.985	0.551	0.973	0.565
	Linear SVM	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.894	0.990	0.934
	Radial SVM	0.993	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.910	0.990	0.890
	KNN	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.895	0.999	0.922
	Naive Bayes	0.625	0.847	0.473	0.756	0.716	0.675	0.566	0.566	0.847	0.847	0.798	0.690	0.694	0.752
	Logistic regresion	0.786	0.834	0.599	0.800	0.757	0.739	0.638	0.638	0.834	0.834	0.684	0.850	0.833	0.808
	CART	0.545	0.756	0.494	0.731	0.645	0.641	0.639	0.639	0.756	0.756	0.654	0.626	0.670	0.741
kyphosis	J48	0.586	0.646	0.502	0.672	0.603	0.604	0.618	0.618	0.646	0.646	0.664	0.600	0.643	0.644
	Linear SVM	0.709	0.824	0.400	0.831	0.727	0.618	0.519	0.485	0.824	0.814	0.737	0.532	0.660	0.750
	Radial SVM	0.583	0.831	0.387	0.850	0.750	0.678	0.712	0.695	0.831	0.831	0.820	0.554	0.675	0.812
	KNN	0.749	0.770	0.530	0.771	0.749	0.709	0.770	0.770	0.770	0.770	0.770	0.763	0.770	0.773
	Naive Bayes	0.515	0.960	0.771	0.959	0.959	0.960	0.959	0.959	0.960	0.960	0.961	0.740	0.551	0.669
	Logistic regresion	0.788	0.953	0.993	0.951	0.952	0.981	0.949	0.949	0.953	0.953	0.950	0.936	0.942	0.971
	CART	0.504	0.855	0.571	0.855	0.853	0.848	0.862	0.862	0.855	0.855	0.861	0.561	0.514	0.552
newthyroid1	J48	0.499	0.932	0.585	0.932	0.927	0.870	0.934	0.934	0.932	0.932	0.931	0.582	0.518	0.551
	Linear SVM	0.206	0.971	0.007	0.969	0.970	0.989	0.965	0.965	0.971	0.971	0.969	0.072	0.059	0.030
	Radial SVM	0.230	0.995	0.008	0.995	0.994	0.995	0.994	0.994	0.995	0.995	0.995	0.064	0.063	0.032
	KNN	0.604	0.995	0.978	0.995	0.995	0.994	0.995	0.995	0.995	0.995	0.995	0.924	0.839	0.911
	Naive Bayes	0.500	0.750	0.747	0.750	0.511	0.519	0.750	0.750	0.750	0.750	0.750	0.745	0.500	0.750
	Logistic regresion	0.500	0.009	0.022	0.009	0.141	0.059	0.009	0.009	0.009	0.009	0.009	0.016	0.064	0.009
	CART	0.500	0.810	0.730	0.809	0.512	0.532	0.810	0.810	0.810	0.810	0.810	0.803	0.536	0.810
Vehicle1	J48	0.500	0.782	0.718	0.782	0.517	0.530	0.782	0.782	0.782	0.782	0.782	0.776	0.550	0.782
	Linear SVM	0.500	0.991	0.978	0.991	0.126	0.203	0.991	0.991	0.991	0.991	0.991	0.977	0.056	0.991
	Radial SVM	0.500	0.885	0.817	0.883	0.118	0.140	0.885	0.885	0.885	0.885	0.885	0.874	0.059	0.885
	KNN	0.500	0.949	0.955	0.949	0.832	0.915	0.949	0.949	0.949	0.949	0.949	0.942	0.914	0.949
	Naive Bayes	0.500	0.715	0.617	0.715	0.710	0.715	0.715	0.715	0.715	0.715	0.715	0.534	0.530	0.517
	Logistic regresion	0.500	0.235	0.272	0.235	0.241	0.235	0.235	0.235	0.235	0.235	0.235	0.307	0.309	0.311
	CART	0.500	0.605	0.552	0.605	0.601	0.605	0.605	0.605	0.605	0.605	0.605	0.501	0.501	0.500
Vehicle3	J48	0.500	0.588	0.528	0.588	0.586	0.588	0.588	0.588	0.588	0.588	0.588	0.514	0.513	0.511
	Linear SVM	0.500	0.765	0.384	0.765	0.757	0.765	0.765	0.765	0.765	0.765	0.765	0.307	0.309	0.311
	Radial SVM	0.500	0.701	0.345	0.701	0.695	0.701	0.701	0.701	0.701	0.701	0.701	0.307	0.309	0.311
	KNN	0.500	0.703	0.706	0.703	0.697	0.703	0.703	0.703	0.703	0.703	0.703	0.743	0.732	0.713
	Naive Bayes	0.523	0.978	0.711	0.976	0.976	0.959	0.965	0.965	0.978	0.978	0.978	0.966	0.539	0.512
	Logistic regresion	0.953	0.973	0.984	0.967	0.984	0.946	0.945	0.945	0.973	0.973	0.973	0.944	0.980	0.977
	CART	0.508	0.975	0.511	0.975	0.731	0.967	0.971	0.971	0.975	0.975	0.975	0.971	0.534	0.509
wisconsin	J48	0.512	0.952	0.517	0.948	0.729	0.954	0.939	0.939	0.952	0.952	0.952	0.937	0.523	0.515
	Linear SVM	0.949	0.988	0.984	0.986	0.988	0.947	0.957	0.957	0.988	0.988	0.988	0.959	0.981	0.976
	Radial SVM	0.351	0.992	0.193	0.992	0.533	0.983	0.991	0.991	0.992	0.992	0.992	0.992	0.272	0.376
	KNN	0.846	0.985	0.758	0.985	0.959	0.985	0.985	0.985	0.985	0.985	0.985	0.985	0.943	0.851
	Naive Bayes	0.865	0.836	0.775	0.840	0.775	0.855	0.821	0.821	0.836	0.836	0.850	0.836	0.867	0.867
	Logistic regresion	0.898	0.836	0.649	0.750	0.645	0.839	0.585	0.585	0.836	0.836	0.720	0.836	0.859	0.859
	CART	0.777	0.518	0.511	0.706	0.511	0.733	0.572	0.572	0.518	0.518	0.718	0.518	0.626	0.626
yeast4	J48	0.674	0.500	0.545	0.606	0.543	0.634	0.597	0.597	0.500	0.500	0.595	0.500	0.500	0.500
	Linear SVM	0.801	0.538	0.536	0.852	0.540	0.761	0.577	0.586	0.433	0.469	0.844	0.501	0.504	0.501
	Radial SVM	0.816	0.528	0.577	0.825	0.582	0.798	0.749	0.749	0.524	0.545	0.823	0.520	0.632	0.632
	KNN	0.758	0.623	0.715	0.760	0.715	0.760	0.757	0.757	0.623	0.623	0.760	0.623	0.708	0.708

Table 5. GMEAN vales for different parameter selection strategies

Datasets	Classifiers	Parameter Selection Strategy													
		Classical	F1	F2	F4	N1	LSC	L1	L2	T2	T3	T4	Density	ClsCoef	Hubs
	Naive Bayes	0.634	0.605	0.861	0.852	0.468	0.857	0.854	0.857	0.000	0.000	0.685	0.862	0.003	0.470
	Logistic regresion	0.860	0.845	0.705	0.844	0.414	0.833	0.546	0.584	0.000	0.000	0.863	0.532	0.000	0.289
	CART	0.824	0.816	0.566	0.796	0.818	0.775	0.356	0.406	0.796	0.796	0.811	0.354	0.794	0.823
Banana	J48	0.824	0.817	0.643	0.821	0.822	0.785	0.475	0.500	0.801	0.801	0.827	0.470	0.802	0.833
	Linear SVM	0.853	0.827	0.019	0.750	0.134	0.556	0.002	0.010	0.000	0.004	0.853	0.000	0.005	0.156
	Radial SVM	0.853	0.850	0.540	0.841	0.840	0.833	0.432	0.457	0.825	0.825	0.850	0.437	0.826	0.831
	KNN	0.839	0.839	0.843	0.839	0.839	0.839	0.630	0.684	0.839	0.839	0.839	0.616	0.839	0.839
	Naive Bayes	0.984	0.977	0.918	0.976	0.974	0.977	0.970	0.970	0.977	0.977	0.968	0.918	0.918	0.871
	Logistic regresion	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	CART	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
banknote	J48	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	Linear SVM	0.300	1.000	0.744	0.999	0.995	1.000	0.976	0.979	1.000	1.000	0.986	0.742	0.729	0.694
	Radial SVM	0.243	0.991	0.557	0.989	0.988	0.991	0.992	0.989	0.991	0.991	0.987	0.541	0.556	0.541
	KNN	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	Naive Bayes	0.688	0.804	0.354	0.696	0.807	0.800	0.886	0.886	0.804	0.804	0.692	0.431	0.804	0.484
	Logistic regresion	0.847	0.784	0.234	0.825	0.831	0.805	0.694	0.694	0.784	0.784	0.814	0.245	0.784	0.331
	CART	0.792	0.789	0.483	0.782	0.810	0.802	0.743	0.743	0.789	0.789	0.784	0.341	0.789	0.432
ecoli1	J48	0.777	0.702	0.489	0.780	0.805	0.793	0.733	0.733	0.702	0.702	0.776	0.376	0.702	0.484
	Linear SVM	0.825	0.783	0.113	0.765	0.817	0.806	0.096	0.079	0.780	0.781	0.759	0.249	0.782	0.330
	Radial SVM	0.814	0.769	0.098	0.801	0.823	0.805	0.732	0.734	0.771	0.770	0.777	0.243	0.772	0.323
	KNN	0.861	0.763	0.839	0.853	0.846	0.837	0.852	0.852	0.763	0.763	0.852	0.762	0.763	0.791
	Naive Bayes	0.537	0.698	0.071	0.689	0.507	0.518	0.527	0.527	0.698	0.698	0.691	0.110	0.592	0.681
glass0	Logistic regresion	0.705	0.747	0.171	0.605	0.666	0.714	0.710	0.710	0.747	0.747	0.610	0.156	0.677	0.747
	CART	0.662	0.696	0.168	0.684	0.627	0.657	0.661	0.661	0.696	0.696	0.703	0.145	0.694	0.691

Datasets	Classifiers	Parameter Selection Strategy													
		Classical	F1	F2	F4	N1	LSC	L1	L2	T2	T3	T4	Density	ClsCoef	Hubs
	J48	0.691	0.733	0.141	0.726	0.654	0.684	0.696	0.696	0.733	0.733	0.727	0.160	0.684	0.737
	Linear SVM	0.469	0.651	0.101	0.270	0.522	0.493	0.464	0.453	0.655	0.657	0.370	0.046	0.399	0.631
	Radial SVM	0.588	0.724	0.108	0.717	0.546	0.566	0.565	0.563	0.728	0.731	0.701	0.126	0.594	0.718
	KNN	0.723	0.755	0.591	0.746	0.675	0.702	0.699	0.699	0.755	0.755	0.741	0.318	0.740	0.754
	Naive Bayes	0.188	0.522	0.379	0.586	0.263	0.270	0.435	0.541	0.522	0.522	0.522	0.530	0.515	0.519
	Logistic regresion	0.121	0.476	0.121	0.461	0.121	0.141	0.202	0.445	0.476	0.476	0.476	0.450	0.467	0.252
	CART	0.154	0.472	0.264	0.487	0.100	0.162	0.245	0.383	0.472	0.472	0.472	0.432	0.468	0.254
haberman	J48	0.113	0.523	0.221	0.489	0.121	0.122	0.251	0.426	0.523	0.523	0.522	0.498	0.518	0.276
	Linear SVM	0.014	0.000	0.004	0.043	0.018	0.014	0.020	0.014	0.010	0.004	0.003	0.004	0.007	0.014
	Radial SVM	0.038	0.262	0.024	0.194	0.014	0.025	0.017	0.143	0.251	0.275	0.259	0.226	0.264	0.041
	KNN	0.003	0.484	0.212	0.479	0.033	0.085	0.251	0.482	0.484	0.484	0.484	0.464	0.480	0.415
	Naive Bayes	0.096	0.890	0.056	0.889	0.031	0.822	0.739	0.739	0.890	0.890	0.878	0.066	0.143	0.879
	Logistic regresion	0.020	0.083	0.031	0.083	0.037	0.110	0.125	0.125	0.083	0.083	0.066	0.030	0.034	0.072
	CART	0.179	0.836	0.247	0.834	0.499	0.872	0.888	0.888	0.836	0.836	0.847	0.208	0.301	0.838
Ionosphere	J48	0.185	0.847	0.276	0.848	0.331	0.851	0.843	0.843	0.847	0.847	0.843	0.310	0.338	0.838
	Linear SVM	0.178	0.845	0.904	0.849	0.897	0.817	0.790	0.792	0.844	0.844	0.864	0.875	0.897	0.881
	Radial SVM	0.279	0.942	0.272	0.941	0.229	0.888	0.840	0.842	0.942	0.941	0.942	0.247	0.338	0.943
	KNN	0.246	0.771	0.482	0.772	0.809	0.767	0.761	0.761	0.771	0.771	0.763	0.347	0.492	0.769
	Naive Bayes	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.115	0.965	0.163
	Logistic regresion	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.157	0.976	0.237
	CART	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.149	0.959	0.183
iris0	J48	0.985	0.987	0.987	0.987	0.987	0.987	0.987	0.987	0.987	0.987	0.984	0.222	0.962	0.247
	Linear SVM	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.116	0.965	0.218
	Radial SVM	0.904	0.989	0.989	0.989	0.989	0.989	0.989	0.989	0.989	0.989	0.987	0.145	0.961	0.194
	KNN	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.736	0.999	0.833
	Naive Bayes	0.204	0.469	0.048	0.471	0.363	0.235	0.209	0.209	0.469	0.469	0.544	0.272	0.291	0.318
	Logistic regresion	0.365	0.449	0.000	0.560	0.468	0.405	0.497	0.497	0.449	0.449	0.505	0.278	0.360	0.479
	CART	0.187	0.677	0.000	0.624	0.447	0.414	0.436	0.436	0.677	0.677	0.449	0.361	0.470	0.638
kyphosis	J48	0.301	0.540	0.000	0.493	0.369	0.361	0.431	0.431	0.540	0.540	0.442	0.353	0.436	0.530
	Linear SVM	0.312	0.043	0.000	0.151	0.122	0.061	0.044	0.048	0.037	0.043	0.147	0.032	0.068	0.036
	Radial SVM	0.063	0.240	0.000	0.339	0.229	0.152	0.075	0.060	0.253	0.245	0.219	0.149	0.196	0.321
	KNN	0.448	0.459	0.000	0.460	0.412	0.338	0.459	0.459	0.459	0.459	0.459	0.351	0.351	0.466
	Naive Bayes	0.010	0.820	0.268	0.819	0.822	0.895	0.828	0.828	0.820	0.820	0.824	0.234	0.150	0.193
	Logistic regresion	0.000	0.913	0.235	0.912	0.913	0.931	0.900	0.900	0.913	0.913	0.911	0.213	0.016	0.042
	CART	0.009	0.805	0.215	0.804	0.805	0.788	0.823	0.823	0.805	0.805	0.814	0.197	0.061	0.138
newthyroid1	J48	0.012	0.911	0.300	0.912	0.905	0.838	0.919	0.919	0.911	0.911	0.911	0.260	0.103	0.194
	Linear SVM	0.000	0.869	0.000	0.882	0.876	0.939	0.868	0.880	0.893	0.877	0.874	0.009	0.000	0.000
	Radial SVM	0.000	0.898	0.000	0.895	0.895	0.966	0.890	0.887	0.900	0.896	0.895	0.000	0.000	0.000
	KNN	0.081	0.917	0.785	0.917	0.913	0.905	0.910	0.910	0.917	0.917	0.917	0.697	0.285	0.514
	Naive Bayes	0.000	0.709	0.694	0.709	0.038	0.147	0.709	0.709	0.709	0.709	0.709	0.694	0.005	0.709
	Logistic regresion	0.000	0.026	0.034	0.026	0.001	0.000	0.026	0.026	0.026	0.026	0.026	0.025	0.000	0.026
	CART	0.000	0.745	0.700	0.744	0.088	0.210	0.745	0.745	0.745	0.745	0.745	0.731	0.184	0.745
Vehicle1	J48	0.000	0.772	0.702	0.772	0.086	0.194	0.772	0.772	0.772	0.772	0.772	0.757	0.239	0.772
	Linear SVM	0.000	0.922	0.851	0.919	0.017	0.034	0.921	0.921	0.919	0.919	0.919	0.902	0.000	0.922
	Radial SVM	0.000	0.313	0.109	0.314	0.007	0.024	0.320	0.316	0.316	0.320	0.322	0.310	0.000	0.316
	KNN	0.000	0.873	0.891	0.873	0.630	0.651	0.873	0.873	0.873	0.873	0.873	0.860	0.768	0.873
	Naive Bayes	0.000	0.672	0.505	0.672	0.657	0.672	0.672	0.672	0.672	0.672	0.672	0.011	0.011	0.005
	Logistic regresion	0.000	0.272	0.134	0.272	0.266	0.272	0.272	0.272	0.272	0.272	0.272	0.000	0.000	0.000
	CART	0.000	0.489	0.324	0.489	0.477	0.489	0.489	0.489	0.489	0.489	0.489	0.038	0.037	0.017
Vehicle3	J48	0.000	0.512	0.288	0.512	0.500	0.512	0.512	0.512	0.512	0.512	0.512	0.068	0.063	0.050
	Linear SVM	0.000	0.562	0.101	0.570	0.546	0.569	0.567	0.569	0.560	0.565	0.567	0.011	0.002	0.006
	Radial SVM	0.000	0.506	0.086	0.509	0.497	0.505	0.506	0.507	0.502	0.506	0.505	0.000	0.000	0.000
	KNN	0.000	0.593	0.559	0.593	0.581	0.593	0.593	0.593	0.593	0.593	0.593	0.588	0.555	0.518
	Naive Bayes	0.109	0.948	0.329	0.947	0.775	0.954	0.949	0.949	0.948	0.948	0.948	0.949	0.148	0.109
	Logistic regresion	0.780	0.936	0.958	0.932	0.958	0.907	0.912	0.912	0.936	0.936	0.936	0.913	0.921	0.838
	CART	0.089	0.974	0.109	0.974	0.614	0.964	0.964	0.964	0.974	0.974	0.974	0.965	0.187	0.089
wisconsin	J48	0.109	0.955	0.135	0.954	0.638	0.947	0.950	0.950	0.955	0.955	0.955	0.950	0.142	0.109
	Linear SVM	0.780	0.947	0.945	0.942	0.955	0.923	0.919	0.919	0.946	0.948	0.947	0.919	0.867	0.783
	Radial SVM	0.109	0.959	0.108	0.957	0.345	0.960	0.962	0.962	0.959	0.959	0.959	0.962	0.135	0.109
	KNN	0.174	0.965	0.563	0.966	0.905	0.967	0.965	0.965	0.965	0.965	0.965	0.966	0.813	0.291
	Naive Bayes	0.537	0.218	0.734	0.633	0.723	0.491	0.582	0.582	0.218	0.218	0.633	0.218	0.220	0.220
	Logistic regresion	0.276	0.000	0.330	0.508	0.331	0.315	0.444	0.444	0.000	0.000	0.501	0.000	0.000	0.000
	CART	0.408	0.000	0.045	0.447	0.049	0.351	0.204	0.204	0.000	0.000	0.451	0.000	0.052	0.052
yeast4	J48	0.345	0.000	0.132	0.506	0.124	0.237	0.294	0.294	0.000	0.000	0.510	0.000	0.000	0.000
	Linear SVM	0.000	0.000	0.004	0.026	0.000	0.049	0.000	0.000	0.012	0.000	0.024	0.000	0.000	0.009
	Radial SVM	0.269	0.000	0.339	0.177	0.350	0.309	0.128	0.126	0.000	0.000	0.206	0.000	0.000	0.009
	KNN	0.379	0.116	0.334	0.364	0.336	0.380	0.394	0.394	0.116	0.116	0.364	0.116	0.277	0.277

Table 6. MCC values for different parameter selection strategies

Datasets	Classifiers	Parameter Selection Strategy													
		Classical	F1	F2	F4	N1	LSC	L1	L2	T2	T3	T4	Density	ClsCoef	Hubs
	Naive Bayes	0.426	0.415	0.698	0.567	0.342	0.543	0.577	0.607	0.000	0.000	0.467	0.586	0.002	0.378
	Logistic regresion	0.791	0.779	0.498	0.716	0.352	0.690	0.334	0.374	0.000	0.000	0.794	0.322	0.000	0.253
	CART	0.749	0.727	0.466	0.702	0.732	0.668	0.255	0.303	0.707	0.707	0.728	0.243	0.706	0.733
Banana	J48	0.741	0.727	0.442	0.698	0.738	0.633	0.288	0.310	0.719	0.719	0.733	0.282	0.719	0.740
	Linear SVM	0.791	0.773	0.018	0.709	0.124	0.508	0.001	0.008	0.000	0.000	0.798	0.001	0.002	0.142
	Radial SVM	0.800	0.796	0.434	0.777	0.779	0.761	0.320	0.346	0.758	0.758	0.796	0.328	0.759	0.770
	KNN	0.762	0.762	0.724	0.763	0.762	0.760	0.524	0.577	0.762	0.762	0.762	0.511	0.762	0.762

Datasets	Classifiers	Parameter Selection Strategy													
		Classical	F1	F2	F4	N1	LSC	L1	L2	T2	T3	T4	Density	ClsCoef	Hubs
banknote	Naive Bayes	0.662	0.945	0.115	0.945	0.940	0.945	0.918	0.918	0.945	0.945	0.956	0.109	0.115	0.122
	Logistic regresion	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	CART	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	J48	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	Linear SVM	0.560	0.982	0.162	0.982	0.982	0.982	0.959	0.961	0.982	0.982	0.952	0.174	0.178	0.213
	Radial SVM	0.654	0.968	0.199	0.966	0.959	0.964	0.938	0.936	0.971	0.966	0.928	0.193	0.188	0.197
	KNN	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
ecoli1	Naive Bayes	0.511	0.604	0.247	0.480	0.660	0.640	0.710	0.710	0.604	0.604	0.459	0.282	0.604	0.343
	Logistic regresion	0.728	0.618	0.153	0.678	0.704	0.668	0.484	0.484	0.618	0.618	0.660	0.202	0.618	0.269
	CART	0.621	0.638	0.338	0.593	0.651	0.650	0.581	0.581	0.638	0.638	0.593	0.264	0.638	0.330
	J48	0.652	0.597	0.335	0.616	0.678	0.664	0.544	0.544	0.597	0.597	0.598	0.275	0.597	0.363
	Linear SVM	0.725	0.627	0.281	0.666	0.699	0.678	0.074	0.060	0.625	0.627	0.647	0.169	0.628	0.219
	Radial SVM	0.704	0.644	0.262	0.681	0.718	0.691	0.577	0.582	0.649	0.648	0.657	0.189	0.649	0.241
	KNN	0.740	0.587	0.709	0.728	0.720	0.709	0.727	0.727	0.587	0.587	0.726	0.604	0.587	0.651
glass0	Naive Bayes	0.345	0.445	0.037	0.429	0.322	0.338	0.343	0.343	0.445	0.445	0.429	0.056	0.391	0.437
	Logistic regresion	0.443	0.533	0.135	0.354	0.421	0.451	0.442	0.442	0.533	0.533	0.362	0.111	0.399	0.534
	CART	0.393	0.466	0.042	0.433	0.353	0.375	0.374	0.374	0.466	0.466	0.459	0.046	0.450	0.450
	J48	0.430	0.488	0.036	0.478	0.389	0.407	0.432	0.432	0.488	0.488	0.489	0.089	0.421	0.491
	Linear SVM	0.304	0.402	0.058	0.121	0.339	0.322	0.305	0.286	0.409	0.404	0.192	0.026	0.207	0.396
	Radial SVM	0.195	0.488	0.066	0.517	0.182	0.175	0.163	0.161	0.494	0.496	0.491	0.097	0.280	0.483
	KNN	0.472	0.541	0.341	0.524	0.401	0.425	0.420	0.420	0.541	0.541	0.515	0.212	0.511	0.539
haberman	Naive Bayes	0.036	0.223	0.097	0.260	0.068	0.069	0.143	0.295	0.223	0.223	0.223	0.223	0.221	0.127
	Logistic regresion	0.045	0.203	0.045	0.121	0.045	0.054	0.016	0.026	0.203	0.203	0.202	0.192	0.198	0.026
	CART	0.022	0.136	0.021	0.158	0.009	0.028	0.022	0.073	0.136	0.136	0.138	0.114	0.140	0.020
	J48	0.018	0.240	0.022	0.151	0.035	0.036	0.034	0.076	0.240	0.240	0.236	0.220	0.236	0.001
	Linear SVM	0.022	0.015	0.018	0.020	0.010	0.015	0.024	0.006	0.006	0.011	0.009	0.007	0.005	0.021
	Radial SVM	0.008	0.129	0.020	0.080	0.020	0.016	0.017	0.005	0.119	0.137	0.132	0.120	0.132	0.014
	KNN	0.039	0.190	0.042	0.181	0.014	0.016	0.073	0.184	0.190	0.190	0.190	0.174	0.187	0.140
Ionosphere	Naive Bayes	0.078	0.786	0.046	0.783	0.025	0.727	0.659	0.659	0.786	0.786	0.776	0.053	0.124	0.776
	Logistic regresion	0.154	0.761	0.857	0.760	0.858	0.706	0.677	0.677	0.761	0.761	0.790	0.805	0.849	0.795
	CART	0.131	0.697	0.125	0.693	0.382	0.762	0.793	0.793	0.697	0.697	0.716	0.116	0.179	0.702
	J48	0.134	0.714	0.173	0.715	0.215	0.732	0.731	0.731	0.714	0.714	0.713	0.203	0.242	0.699
	Linear SVM	0.089	0.761	0.791	0.768	0.790	0.712	0.665	0.667	0.762	0.761	0.783	0.747	0.786	0.809
	Radial SVM	0.093	0.884	0.197	0.883	0.072	0.807	0.738	0.742	0.885	0.883	0.895	0.188	0.270	0.888
	KNN	0.175	0.677	0.376	0.677	0.733	0.672	0.665	0.665	0.677	0.677	0.669	0.244	0.386	0.675
iris0	Naive Bayes	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.070	0.960	0.119
	Logistic regresion	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.134	0.973	0.211
	CART	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.059	0.952	0.103
	J48	0.979	0.981	0.981	0.981	0.981	0.981	0.981	0.981	0.981	0.981	0.978	0.159	0.951	0.181
	Linear SVM	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.093	0.960	0.192
	Radial SVM	0.872	0.985	0.985	0.985	0.985	0.985	0.985	0.985	0.985	0.985	0.982	0.113	0.952	0.154
	KNN	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.692	0.999	0.796
kyphosis	Naive Bayes	0.086	0.332	0.001	0.310	0.232	0.122	0.075	0.075	0.332	0.332	0.336	0.187	0.185	0.199
	Logistic regresion	0.239	0.330	0.000	0.398	0.306	0.246	0.248	0.248	0.330	0.330	0.289	0.198	0.266	0.345
	CART	0.049	0.488	0.009	0.418	0.276	0.285	0.285	0.285	0.488	0.488	0.267	0.252	0.322	0.457
	J48	0.167	0.373	0.007	0.304	0.173	0.219	0.234	0.234	0.373	0.373	0.267	0.241	0.299	0.356
	Linear SVM	0.166	0.006	0.077	0.108	0.067	0.021	0.007	0.008	0.003	0.007	0.088	0.069	0.018	0.005
	Radial SVM	0.018	0.153	0.076	0.277	0.153	0.103	0.014	0.000	0.179	0.165	0.163	0.018	0.094	0.234
	KNN	0.309	0.304	0.008	0.308	0.272	0.218	0.302	0.302	0.304	0.304	0.304	0.239	0.240	0.312
newthyroid1	Naive Bayes	0.007	0.664	0.244	0.662	0.666	0.704	0.665	0.665	0.664	0.664	0.672	0.205	0.071	0.130
	Logistic regresion	0.000	0.816	0.221	0.813	0.819	0.890	0.809	0.809	0.816	0.816	0.812	0.201	0.015	0.040
	CART	0.005	0.739	0.175	0.739	0.737	0.687	0.757	0.757	0.739	0.739	0.748	0.148	0.029	0.099
	J48	0.007	0.863	0.190	0.864	0.859	0.757	0.870	0.870	0.863	0.863	0.863	0.176	0.042	0.101
	Linear SVM	0.000	0.802	0.260	0.811	0.816	0.903	0.802	0.816	0.828	0.817	0.805	0.236	0.003	0.089
	Radial SVM	0.000	0.863	0.248	0.862	0.860	0.927	0.855	0.853	0.866	0.861	0.859	0.239	0.002	0.084
	KNN	0.064	0.892	0.736	0.892	0.888	0.878	0.885	0.885	0.892	0.892	0.892	0.655	0.255	0.474
Vehicle1	Naive Bayes	0.000	0.412	0.417	0.412	0.023	0.077	0.412	0.412	0.412	0.412	0.412	0.403	0.003	0.412
	Logistic regresion	0.000	0.858	0.731	0.856	0.057	0.291	0.858	0.858	0.858	0.858	0.858	0.841	0.072	0.858
	CART	0.000	0.540	0.375	0.537	0.044	0.121	0.540	0.540	0.540	0.540	0.540	0.530	0.145	0.540
	J48	0.000	0.556	0.412	0.556	0.059	0.114	0.556	0.556	0.556	0.556	0.556	0.544	0.190	0.556
	Linear SVM	0.000	0.873	0.779	0.870	0.008	0.026	0.871	0.871	0.868	0.868	0.870	0.850	0.024	0.873
	Radial SVM	0.000	0.166	0.056	0.166	0.046	0.127	0.170	0.168	0.168	0.170	0.171	0.163	0.015	0.167
	KNN	0.000	0.771	0.784	0.771	0.516	0.560	0.771	0.771	0.771	0.771	0.771	0.759	0.643	0.771
Vehicle3	Naive Bayes	0.000	0.313	0.241	0.313	0.305	0.313	0.313	0.313	0.313	0.313	0.313	0.011	0.009	0.011
	Logistic regresion	0.000	0.334	0.094	0.334	0.327	0.334	0.334	0.334	0.334	0.334	0.334	0.000	0.000	0.000
	CART	0.000	0.176	0.151	0.176	0.170	0.176	0.176	0.176	0.176	0.176	0.176	0.006	0.007	0.009
	J48	0.000	0.171	0.085	0.171	0.166	0.171	0.171	0.171	0.171	0.171	0.171	0.009	0.010	0.006
	Linear SVM	0.000	0.298	0.111	0.307	0.292	0.303	0.306	0.301	0.298	0.306	0.303	0.093	0.085	0.065
	Radial SVM	0.000	0.259	0.120	0.263	0.255	0.259	0.258	0.262	0.255	0.260	0.260	0.076	0.068	0.054
	KNN	0.000	0.260	0.253	0.260	0.254	0.260	0.260	0.260	0.260	0.260	0.260	0.277	0.258	0.245
wisconsin	Naive Bayes	0.088	0.876	0.272	0.874	0.690	0.886	0.880	0.880	0.876	0.876	0.876	0.878	0.124	0.088
	Logistic regresion	0.597	0.879	0.897	0.874	0.908	0.836	0.845	0.845	0.879	0.879	0.879	0.846	0.818	0.679
	CART	0.069	0.939	0.077	0.938	0.537	0.922	0.922	0.922	0.939	0.939	0.939	0.923	0.144	0.069
	J48	0.087	0.912	0.100	0.912	0.563	0.897	0.903	0.903	0.912	0.912	0.912	0.903	0.113	0.084
	Linear SVM	0.597	0.899	0.865	0.892	0.897	0.860	0.858	0.858	0.898	0.900	0.899	0.859	0.724	0.601
	Radial SVM	0.088	0.915	0.056	0.913	0.277	0.914	0.916	0.917	0.916	0.916	0.915	0.917	0.112	0.088
	KNN	0.152	0.932	0.457	0.932	0.844	0.934	0.932	0.932	0.932	0.932	0.932	0.932	0.740	0.251
yeast4	Naive Bayes	0.272	0.138	0.202	0.221	0.201	0.342	0.213	0.213	0.138	0.138	0.224	0.138	0.127	0.127
	Logistic regresion	0.227	0.002	0.195	0.320	0.195	0.226	0.164	0.164	0.002	0.002	0.302	0.002	0.003	0.003

Datasets	Classifiers	Parameter Selection Strategy													
		Classical	**F1**	**F2**	**F4**	**N1**	**LSC**	**L1**	**L2**	**T2**	**T3**	**T4**	**Density**	**ClsCoef**	**Hubs**
	CART	0.298	0.002	0.035	0.305	0.038	0.244	0.146	0.146	0.002	0.002	0.310	0.002	0.024	0.024
	J48	0.239	0.000	0.081	0.342	0.075	0.153	0.161	0.161	0.000	0.000	0.343	0.000	0.000	0.000
	Linear SVM	0.000	0.000	0.003	0.021	0.000	0.041	0.000	0.000	0.005	0.000	0.023	0.001	0.000	0.003
	Radial SVM	0.242	0.000	0.223	0.137	0.239	0.265	0.092	0.091	0.000	0.000	0.155	0.000	0.001	0.004
	KNN	0.268	0.072	0.219	0.257	0.221	0.256	0.277	0.277	0.072	0.072	0.258	0.072	0.171	0.171

The key point of the classification process is determined by the kernel hyperparameter selection approaches. We utilize 7 different classifiers and 13 different data complexity measures which shown in Table 3, 4, and 5. Table 3, 4, and 5 gives AUC, GMEAN and MCC values for each case. Different performance measures may find different approaches better. Some cases did not require any kernel hyperparameter selection as predictions were completely correct. However, there were cases which kernel hyperparameter selection is utmost importance. Datasets Haberman, ionosphere, and kyphosis are among these. Number of wins for parameter selection strategies in each classifier and each performance mesure are counted and is given in Table 7. As an example, it should be read like this: F1 parameter selection strategy performed best 5 times using CART classifier in AUC for all datasets. These wins are based on Table 4, 5 and 6. Win count is increased 1 for each tie also. So total numbers can be more than 13, which is the number of datasets. In total wins, F1, T2, and T3 measures are performed better for AUC criteria. The same goes for T3 for GMEAN and T2 and T3 for MCC criteria. Table 8 gives total win numbers according to classifiers in a similar way to Table 7. Number of wins are the count of how many times a classifier performed best in each performance measure and each parameter selection strategy. An example of how to read numbers is: CART algorithm performed best 3 times using classical parameter selection strategy in AUC for all datasets. Total wins are total number of wins of all parameter selectino strategies. Based on total wins, Radial SVM performed better with all KPCA strategies in AUC. KNN performed better than other classifiers in GMEAN and MCC.

Radial SVM wins 56 times and is the best one in AUC performance. KNN performed better than other classifiers in GMEAN and MCC.

Table 7. Number of wins for parameter selection strategies

Performance measure	Classifier	Parameter selection strategy													
		Classical	F1	F2	F4	N1	LSC	L1	L2	T2	T3	T4	Density	ClsCoef	Hubs
AUC	CART	3	5	1	3	2	2	5	5	5	5	6	0	0	1
	J48	1	4	1	4	1	5	3	3	5	5	2	0	0	0
	KNN	3	3	3	2	3	1	2	3	3	3	4	1	0	2
	Linear SVM	1	7	2	5	2	4	3	3	6	6	5	0	0	1
	Logistic regresion	4	2	3	1	2	1	2	2	2	2	2	2	1	1
	Naive Bayes	1	8	1	5	2	3	3	3	8	8	6	0	1	2
	Radial SVM	0	5	1	3	2	3	2	2	5	5	3	0	0	1
	Total	13	**34**	12	23	14	19	20	21	**34**	**34**	28	3	2	8
GMEAN	CART	3	6	2	4	3	3	6	6	6	6	7	1	1	2
	J48	1	6	2	4	3	4	4	4	6	6	4	1	1	3
	KNN	3	6	4	4	3	4	4	4	6	6	5	1	1	2
	Linear SVM	4	3	2	3	2	4	1	1	2	3	1	0	0	0
	Logistic regresion	3	4	4	5	2	4	4	4	4	4	4	1	1	2
	Naive Bayes	2	5	2	3	1	4	4	4	5	5	4	1	0	1
	Radial SVM	1	1	1	3	3	2	2	2	1	3	1	0	0	1
	Total	17	31	17	26	17	25	25	25	30	**33**	26	5	4	11
MCC	CART	3	7	2	4	3	3	6	6	7	7	6	1	1	2
	J48	2	6	2	4	3	4	4	4	6	6	4	1	1	2
	KNN	3	5	3	4	3	3	3	3	5	5	4	2	1	2
	Linear SVM	3	3	1	3	1	4	2	1	3	3	2	0	0	1
	Logistic regresion	3	5	2	5	4	4	4	4	5	5	5	1	1	3
	Naive Bayes	1	4	3	2	1	5	3	4	4	4	4	0	0	0
	Radial SVM	1	1	1	4	2	3	1	2	2	2	2	0	0	0
	Total	16	31	14	26	17	26	23	24	**32**	**32**	27	5	4	10

Table 8. Number of wins for classifiers

Performance measure	Classifier	Parameter selection strategy													Total Wins
		Classical	F1	F2	F4	N1	LSC	L1	L2	T2	T3	T4	Density	ClsCoef	Hubs	
AUC	CART	3	1	1	1	1	1	1	1	1	1	0	0	0	13
	J48	2	0	0	0	0	0	0	0	0	0	0	0	0	2
	KNN	5	3	5	3	4	4	5	5	3	3	2	3	4	5	54
	Linear SVM	5	4	4	6	5	3	3	3	4	4	5	2	2	1	51
	Logistic regresion	10	4	6	1	4	4	2	1	3	3	2	6	4	3	53
	Naive Bayes	3	4	2	2	2	2	3	4	4	4	3	2	2	1	38
	Radial SVM	2	6	1	7	3	5	4	4	6	6	7	1	1	3	**56**
GMEAN	CART	4	4	2	4	2	3	3	3	4	4	3	2	2	2	42
	J48	3	2	1	1	1	1	2	2	2	2	2	1	2	1	23
	KNN	8	4	7	5	6	5	3	3	5	5	5	7	8	8	**79**
	Linear SVM	4	3	2	2	2	2	2	2	3	3	2	2	1	1	31
	Logistic regresion	6	2	3	2	5	3	4	4	2	2	3	1	2	2	41
	Naive Bayes	5	4	5	5	4	5	6	6	4	4	5	3	1	1	58
	Radial SVM	3	2	0	1	1	2	0	0	1	1	1	0	0	1	13
MCC	CART	4	3	1	3	1	2	2	2	3	3	2	1	1	1	29
	J48	2	1	0	1	0	0	0	0	1	1	2	0	1	0	9
	KNN	8	3	7	4	4	4	6	6	4	4	4	6	7	7	**74**
	Linear SVM	4	3	2	3	2	2	3	3	3	3	3	1	0	2	34
	Logistic regresion	5	2	3	2	6	3	3	3	2	2	2	1	2	1	37
	Naive Bayes	4	2	2	2	2	3	3	3	2	2	3	3	0	0	31
	Radial SVM	3	3	2	2	2	3	0	0	2	2	1	1	2	2	25

5. CONCLUSION

KPCA is a form of PCA which can identify nonlinear structure of dataset. The kernel hyperparameters in KPCA play a major role in the emergence of this structure. Therefore, it is important to choose these parameters correctly. We suggested that these parameters can be estimated accurately using measures of complexity. We summarized the results we obtained from 7 different classifiers and 13 different data sets in Table 7. Measures F1, T2 and T3 were found to be more successful than other measures. Also, classical parameter selection clearley performed worse than all complexity measure approaches. Therefore, we recommend this approach to researchers who will use KPCA.

Declaration of Competing Interest: *We have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.*

REFERENCES

[1] Zhang, B.L., *Australasian Joint Conference on Artificial Intelligence*, Berlin, Heidelberg, pp. 719-728, 2006.
[2] Garcia-Diaz, P., Sánchez-Berriel, I., Martínez-Rojas, J.A., Diez-Pascual, A.M., *Genomics*, **112**(2), 1916, 2020.
[3] Kelly, J.G., Angelov, P.P., Trevisan, J., Vlachopoulou, A., Paraskevaidis, E., Martin-Hirsch, P.L., Martin, F.L., *Analytical and bioanalytical chemistry*, **398**(5), 2191, 2010.
[4] Dess, B.W., Small, G.W., *Journal of Environmental Radioactivity*, **217**, 1, 2020.
[5] Kastorini, C.M., Papadakis, G., Milionis, H.J., Kalantzi, K., Puddu, P.E., Nikolaou, V., Vemmos, K.N., Goudevenos, J.A., Panagiotakos, D.B., *Artificial intelligence in medicine*, **59**(3), 175, 2013.

[6] Falchetti, M., Prediger, R.D., Zanotto-Filho, A., *Computers in Biology and Medicine*, **124**, 1, 2020.
[7] Hameed, N., Shabut, A.M., Ghosh, M.K., Hossain, M.A., *Expert Systems with Applications*, **141**, 1, 2020.
[8] Tuda, M., Luna-Maldonado, A. I., *Ecological Informatics*, **60**, 1, 2020.
[9] Chen, J.H., Tsai, Y.C., *Financial Innovation*, **6**(1), 1, 2020.
[10] Olivieri, A.C., *Introduction to Multivariate Calibration*, Springer, Berlin, Germany, pp. 73-86, 2018.
[11] Zheng, W., Zou, C., Zhao, L., *Neural Processing Letters*, **22**(1), 49, 2005.
[12] Hu, Y.C., *Computers & Industrial Engineering*, **51**(1), 117, 2006.
[13] Polat, K., Gunes, S., *Expert systems with Applications*, **34**(1), 773, 2008.
[14] Pamukcu, E., Bozdogan, H., Calik, S., *Computational and mathematical methods in medicine*, **2015**, 1, 2015.
[15] Arivudainambi, D., KA, V.K., Visu, P. *Computer Communications*, **147**, 50, 2019.
[16] Lu, H., Meng, Y., Yan, K., Gao, Z., *Cognitive Systems Research*, **53**, 111, 2019.
[17] Zhang, X., He, L., Zhang, J., Whiting, M.D., Karkee, M., Zhang, Q., *Biosystems Engineering*, **193**, 247, 2020.
[18] Biehl, M., Kastner, M., Lange, M., Villmann, T., *Advances in Self-Organizing Maps*, Springer, Berlin, Heidelberg, pp. 23-33, 2013.
[19] Scholkopf, B., Smola, A., Muller, K.R., Kernel principal component analysis, *International conference on artificial neural networks*, Berlin, Heidelberg, pp. 583-588, 1997.
[20] Kaieda, K., Abe, S., *International Journal of Approximate Reasoning*, **37**(3), 189, 2004.
[21] Cheng, L., Ding, Y., Hao, K., Hu, Y., *Neurocomputing*, **83**, 1, 2012.
[22] Karg, M., Jenke, R., Seiberl, W., Kuhnlenz, K., Schwirtz, A., Buss, M., *3rd International Conference on Affective Computing and Intelligent Interaction and Workshops*, pp. 1-6, 2009.
[23] Kuang, F., Xu, W., Zhang, S., *Applied Soft Computing*, **18**, 178, 2014.
[24] Wang, Q., *arXiv* arXiv, **1207**, 3538, 2014.
[25] Alam, M.A., Fukumizu, K., *Journal of Computer Science*, **10**(7), 1139, 2014.
[26] Lorena, A.C., Maciel, A.I., de Miranda, P.B., Costa, I.G., Prudêncio, R.B., *Machine Learning*, **107**(1), 209, 2018.
[27] Lorena, A.C., Garcia, L.P., Lehmann, J., Souto, M.C., Ho, T.K., *ACM Computing Surveys (CSUR)*, **52**(5), 1, 2019.
[28] Orriols-Puig, A., Macia, N., Ho, T.K., *Universitat Ramon Llull*, **196**, 1, 2010.
[29] Lorena, A.C., Costa, I.G., Spolaor, N., De Souto, M. C., *Neurocomputing*, **75**(1), 33, 2012.