ISSN 2712-0554
Original Research

Heritage and Sustainable Development
Vol. 5, No. 2, August 2023, pp.183-198
https://doi.org/10.37868/hsd.v5i2.236

Dynamic System linear models and Bayes classifier for time series classification in promoting sustainability

Abdulsattar Abdullah Hamad [1*], Faris Maher Ahmed [2], Mamoon Fattah Khalf [3], M. Lellis Thivagar[4]

[1, 2, 3] Department of physics, College of Education, University of Samarra, Iraq
[4] School of Mathematics, Madurai Kamaraj University, Madurai, Tamil Nadu, India

*Corresponding author E-mail: a.a.hamad@uosamarra.edu.iq

Received Apr. 20, 2023
Revised Aug. 15, 2023
Accepted Aug. 27, 2023

Abstract

Research purpose: The current work introduces a novel method for time series discriminant analysis (DA). Proposing a version for the Bayes classifier employing Dynamic Linear Models, which we denote by BCDLM This article explores the application of DLMs and the Bayes Classifier in time series classification to promote application in sustainability across diverse sectors.

Method: This paper presents some computer simulation studies in which we generate four different scenarios corresponding to time series observations from various Dynamic Linear Models (DLMs). In Discriminant Analysis, we investigated strategies for estimating variance in models and compared the performance of the BCDLM with other common classifiers. Such datasets are composed of real-time series (data from SONY AIBO Robot and spectrometry of coffee types) and pseudo-time series (data from Swedish leaves adapted for time series). We also point out that algorithm was used to determine training and test sets in real-world applications.

Results: Considering the real-time series examined in this paper, The results obtained indicate that the parametric approach developed represents a promising alternative for this class of DA problems, with observations of time series in a situation that is quite difficult in practice when we have series with large sizes with respect to the number of observations in the classes, even though more thorough studies are required.

Conclusions: It concludes that the BCDLM performed comparably to the results of the classifiers 1NN, RDA, NBND and NBK and superior to the methods LDA and QDA. This offers a powerful combination for time series classification, enabling accurate predictions and informed decision-making in areas such as energy consumption, waste management, and resource allocation.

Keywords: Discriminant Analysis, Application in Sustainability, Naive Bayes, time series, Dynamic Linear Models, classifiers

1. Introduction

The problems addressed in Discriminant Analysis (DA) are characterized by the observation of a set of variables on the objects of study, which have different characteristics or behaviors, with the aim of associating them with previously defined classes. When addressing these problems, a procedure must be developed to classify objects, which is called a declassifier. The set of variables is called a feature vector, and, in practice, we have observations of it for each of the classes. These observations form the training set for the development of the classifier. Once the classifier is obtained, it can be used to classify a new object whose class is unknown [1].

Energy Consumption Forecasting: DSLMs combined with the Bayes Classifier can predict energy consumption patterns, helping industries and households optimize energy usage. This approach assists in load balancing, demand response, and efficient utilization of renewable energy sources, contributing to reduced carbon emissions.

We can exemplify a wide range of objects to be studied in DA [2], such as individuals to be associated with sick and non-sick classes; plants to be associated with different species; digital images of tumors to be classified as benign or malignant; mass spectrometry signals to be classified as coming from different sources. DA is one of the techniques within the area of supervised pattern recognition, and for several other application examples, see, for example, A. Jumaa et al. (2023) [3]. Using the idea of Bayes classifier and the Dynamic Linear Model (DLM) to construct a self-calibration tool through a collection of observations whose classes are known, we present a new approach to problems in DA where the feature vectors are time series.

The Sony AIBO Robot is a diminutive quadruped robot in the form of a dog that is fitted with a variety of sensors, including a tri-axial accelerometer. The set of observations created by Turki, A. I (2023) [4] in which accelerometer measurements were recorded while the robot walked in circles on two types of surfaces: cement and carpet [5]. The data obtained for a horizontal axis is available in [6]. Ahmed, F. M et al. (2023) [6] under the name Sony AIBO Robot Surface. Each time series represents a complete loop. 621 turns were recorded, 349 on cement and 272 on carpet. Cement is harder than carpet, which makes for more surface variability. Consider each surface as a class.

Discriminant Analysis (DA), a classification technique within the field of Supervised Pattern Recognition (SPR), addresses problems where we must allocate objects whose classes are previously known [3]. Objects can be of any nature, people, plants, digital images, etc., which are described by observations derived from a set of variables. In particular, we are interested in cases where observations are obtained over time, forming a time series. Classes are previously defined categories where objects must be allocated. The observations obtained regarding an object are modelled as a random vector $X^T = (X_1, X_2, \cdot\cdot\cdot, X_t)$, where its components are also random *variables*. Such vector X is called a "feature vector" [7]. The Naive Bayes Classifier assumes that in each class, the variables that make up the feature vector X are independent. Although this assumption is not generally true, it simplifies the estimation of conditional densities drastically. Despite these rather optimistic assumptions, the Naive Bayes classifier often outperforms much more sophisticated alternatives and is quite appropriate in DA problems when the number of variables in the X feature vector is very large, in particular when the sample size is much smaller. then the number of observations [3].

Al-Obeidi (2021) [8] discuss the consistency conditions of Naive Bayes and demonstrate that the classifier can be optimal even when the assumption of independence is violated, considering the loss function $0-1$. Bickel and Levina (2004) [9] discuss the consistency of Naive Bayes when the number of variables increases faster than the number of observations. The authors consider only two classes modelled with multivariate normal distributions. In addition to theoretical deductions from consistency conditions, the authors conclude that sometimes Naive Bayes performs better than other models that estimate the dependency structure of observations.

Transportation and Logistics: Sustainable transportation planning relies on accurate predictions of traffic patterns and travel demand. Time series classification can enhance route optimization, reduce congestion, and promote eco-friendly transportation options.

The objective is to identify which of the two surfaces the Robot is walking on again based on the observation of the time series. The problem described is typical of the ones we mentioned; that is, the observed data for the feature vector provide a time series, which is challenging in DA. Based on the above information, this work is to build a classifier capable of distinguishing the characteristics of a time series, considering a training sample, assuming the structure of a dynamic linear model to the data.

Time series data is prevalent in sustainability-related applications, where understanding historical trends and predicting future behaviors is essential. Sustainable practices require efficient utilization of resources, reduction of waste, and minimizing negative environmental impacts. Time series classification plays a crucial role by enabling early identification of trends and anomalies, leading to proactive interventions and informed decision-making. DLMs and the Bayes Classifier offer a powerful combination for time series classification, enabling accurate predictions and informed decision-making in areas such as energy consumption, waste management, and resource allocation. In sustainability applications, DSLMs can help model the interactions between resource consumption, environmental variables, and other relevant factors.

2. Methodology-organization of simulations

Before testing the performance of BCDLM on real-time series, it is important to gain knowledge about its performance in the case where we know the true structure of the series. This can be done through simulation studies, in which we can generate time series from a given DLM. For $i = 1, \ldots, K$, suppose we want to generate $l^{(i)}$ time t series of length according to the dynamic linear model $\{\mathbb{F}, \mathbb{G}, \mathbb{V}, \mathbb{W}\}_t^{(i)}$. We use the following algorithm to generate the time series in all simulations:

1. Simulate $\theta_t^{(i)} \sim N[m_o, C_o]$.
2. For j varying from 1 to t, simulate $\theta_J \sim N\left[G_J^{(i)}\, \theta_{J-1}^{(i)}, \mathbb{W}_J^{(i)}\right]$.
3. For each j varying from 1 to t manage, independently, $y_1, \ldots\ldots\ldots, y_{l(i)} \sim N\left[F_J^{T(i)}\, \theta_J^{(i)}, \mathbb{V}_J^{(i)}\right]$

Note that in each class, a single sequence of parameters is generated where the time series within each class are simulated from these. Therefore, it is possible that two or more classes come from the same DLM, and their differences will be given by the latent vector of parameters. This strategy and simulation match the real series that we have observed in practice.

In this section, it was presents two simulation studies (hereinafter denoted by S3 and S4) carried out to compare the performance of the classifiers discussed in this dissertation. In scenario S3, we generate 180-time series of length 30 with two classes (90 series for each class) generated from a polynomial DLM of order 1. In scenario S4, we generate 20 samples of time series of length 30 with two classes (10 series for each) also generated from order 1 polynomial DLM. For the evaluation of the two studies, cross-validation was used with repeated random subsamples, with 500 repetitions, always using 70% of the time series as a training set, with the minimum number of time series of a class in the training set being 40 and 5 for S3 and S4 respectively (these values were empirically determined, ensuring that it would be possible to apply QDA in S3 and RDA in S4).

2.1 Comparing strategies to estimate variance

It was present two simulation studies (hereinafter denoted by S1 and S2) carried out with the purpose of evaluating the impact of the estimation (or elicitation) strategies of the variances present in the BCDLM. In scenario S1 we generate 100 time series of length 20 with two classes (50 series for each) generated from a polynomial DLM of order 1. In scenarioS2, we generate 100 time series of length 20 with two classes (50 for each) generated from a trigonometric DLM with period 6 and 1 harmonic. For the evaluation of the two studies, cross-validation was used with repeated random subsamples, with 500 repetitions, always using 70% of the time series as a training set.

The QDA, RDA and Naive Bayes classifiers estimate the variance of each class. In this sense, it is important that the dynamic linear models have the advantage of adequately estimating the variances $\mathcal{V}$t and $\mathbb{W}_t$ (aiming to achieve good accuracy rates in comparative classification studies).

Below, we list the three strategies that were evaluated in this dissertation:

- ($\mathbb{V}$,Δ): In this strategy, $\mathcal{V}$t is estimated at each time by its maximum likelihood estimator $\mathcal{V}$t, while $\mathbb{W}_t$ is elicited through discount factors.
- (φ,Δ): In this strategy, $\mathcal{V}$t is considered fixed over time, resulting in the classifier that uses the t-Student model.
- ($\mathbb{V}$, $\mathbb{W}$): In this strategy, $\mathcal{V}$t is estimated at each time by its maximum likelihood estimator $\mathbb{V}$t. For superimposed models, $\mathbb{W}_t$=diag{ $\mathbb{W}_{\mathrm{1t}}$, . . . , $\mathbb{W}_{\mathrm{ht}}$ } and each $\mathbb{W}_{\mathrm{jt}}$=$\mathbb{W}_{\mathrm{1I}}$. The hyperparameters 1, . . . , $\mathbb{W}_h$ are estimated by maximizing the predictive distribution.

In order to determine which strategies are adequate, we performed two simulation studies. The following algorithm was used in simulation studies S_1 and S_2 in order to determine the error rates associated with the compared strategies:

Algorithm A

1. Take a simple random sample without replacement of 70 time series to form the training sample.
2. Use the training sample to build the classifiers with each variance strategy.
3. Sort the remaining time series using the classifiers obtained in Step 2.
4. Save the classification error rate for each classifier defined by the different strategies.

In the first computer simulation study (S1), the following polynomial models of order 1 were considered:

• Class 1:{1,1,10,.1}com m_0 =0
• Class 2:{1,1,10,.1}com m_0=1

Figure 1a shows the graph of examples of two classes simulated from the order 1 model. Table 1 shows the Average and standard deviation of error rates expressed as a percentage. We can see that the strategy that employs $\mathbb{V}_t$, both with the elicitation and the estimation of W, performed better than the strategy that employs $\mathcal{V}$=1/φ.

Table 1. Average and standard deviation of error rates (in %) with different variance estimation strategies for scenario S1

	$\mathbb{V}_t$, Δ	φ, Δ	$\mathbb{V}_t$, $\mathbb{W}$
Average	0.195195	0.45255	0.19761
Standard deviation	0.06783	0.141225	0.06825

Table 2. Performance in terms of error rates (in %) of classification with different strategies for estimating variance for scenario S1

		$\mathbb{V}_t$. Δ	φ. Δ	$\mathbb{V}_t$. $\mathbb{W}$
$\mathbb{V}_t$. Δ	Equal	-	4.2	59.22
	Minor	-	95.34	24.57
	Total	-	99.54	83.79
φ. Δ	Equal	4.2	-	4.41
	Minor	5.46	-	6.3
	Total	9.66	-	10.71
$\mathbb{V}_t$. $\mathbb{W}$	Equal	59.22	4.41	-
	Minor	21.21	94.29	-
	Total	80.43	98.7	-

In Table 2, it presents the percentage of the number of times (Total) that a method has an error rate less than or equal to another. From this table, we observe that the strategy with (φ, Δ), in terms of the error rate, was superior around only 10% of the time compared to the other strategies.

In the second simulation study, studyS_2, observations of time series in two classes were simulated from the same trigonometric model where the only differential between the two classes is based on θ. This situation, in general, is what we observe in real data, where the data are very overlapping and, when classified, the DLM classifies them through the fluctuation of the observed series around $F_t\theta$.

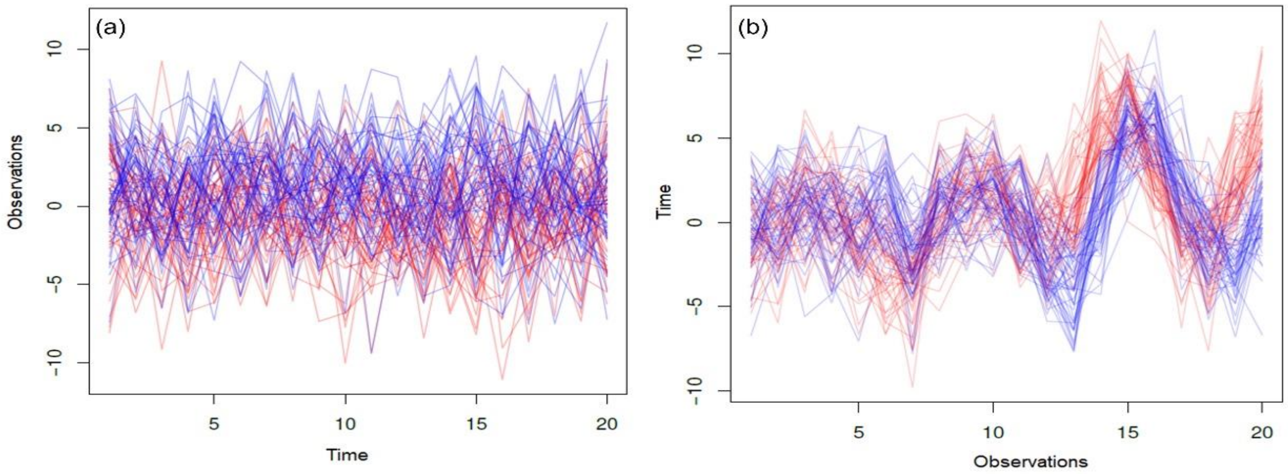


Figure 1. (a) Simulated series with two classes from the order 1 polynomial DLM; (b) Simulated series with two classes from the period 6 trigonometric DLM with one harmonic

$$\left\{\begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}\cos\left(\frac{\pi}{3}\right) & \sin\left(\frac{\pi}{3}\right)\\ -\sin\left(\frac{\pi}{3}\right) & \cos\left(\frac{\pi}{3}\right)\end{pmatrix}, 5, \begin{pmatrix}\frac{1}{2} & 0\\ 0 & \frac{1}{2}\end{pmatrix}\right\}$$

With this trigonometric model employed, the classes will be differentiated by the latent structures from the simulated θ_j sequences. In this way, series corresponding to classes with marked overlap were obtained, which constituted an attempt to reproduce the behavior of series observed in real problems.

Figure 1b presents examples of the series simulated in the second simulation study. Table 3 shows the Average and standard deviation of the error rates obtained for each strategy. Note that the average rate was lower for $\mathcal{V}_t$ estimated by $\mathbb{V}_t$ and $\mathbb{W}$ optimized. Attract through Table 4, we noticed that the methods $\mathbb{V}_t$performed well, being that in this study, the method that optimizes $\mathbb{W}$ was superior to the discount factors. Anticipating a comment on the use of the BCDLM with series derived from a set of real data, we also noticed that the discount factor performed slightly worse than the other methods. Therefore, we decided to employ the strategy ($\mathbb{V}_t$, $\mathbb{W}$), which proved to be superior in this initial analysis, in the simulation studies comparing classifiers presented in the following section, and in the applications with real data presented in Chapter 6.

Table 3. Average and standard deviation of error rates (in %) comparing different variance strategies for scenario S2

	$\mathbb{V}_t$,Δ	φ, Δ	$\mathbb{V}_t$, $\mathbb{W}$
Average	0.00217	0.29246	7.04E-05
Standard deviation	0.009011	0.198056	0.001566

2.2 Comparisons of BCDLM with other classifiers

In this section, we compare the performance of the proposed classifier, BCDLM [10-11], with variance estimation strategy ($\mathbb{V}_t$,$\mathbb{W}$), with the other classifiers discussed in this dissertation. This simulation study considers the model {1,1,10,1}, Polynomial Model of Order 1. In this model, the variance of the observations is greater than that of the parameters, something common in practice. Therefore, we will have a series of the same class that are distant from each other with little application Sustainability apparent structural behavior.

We created two scenarios, S3 and S4, as previously described. In the first S3, the number of series in the training set is greater than the length of the time series. In this case, all classification methods discussed in this dissertation can be used. In the second, S4, the number of sets in the training set is smaller than the length of the sets and classification methods such as LDA and QDA cannot be used.

Table 4. Performance in terms of error rates (in %) comparing different variance strategies for scenario S2

		$\mathbb{V}_t, \Delta$	φ, Δ	$\mathbb{V}_t, \mathbb{W}$
$\mathbb{V}_t$, Δ	Equal	-	0.0861	0.9912
	Minor	-	0.9639	0
	Total	-	1050	0.9912
φ, Δ	Equal	0.0861	-	0.0798
	Minor	0	-	0
	Total	0.0861	-	0.0798
$\mathbb{V}_t$, $\mathbb{W}$	Equal	0.9912	0.0798	-
	Minor	0.0588	0.9702	-
	Total	1	1	-

In this second simulation study, the following algorithm was considered:

Algorithm B

1. Take a simple random sample without replacement from the T (number of runs in the training set) time series.
    (a) Check how many series of each class were selected. If the total series for one of the classes is less than n C (number of elements in the test set) , go back to Step 1. If not, go to Step 2.
2. Use the training sample to build the classifiers.
3. Sort the remaining time series using the classifiers obtained in Step 2.
4. Save the classification error rate for each classifier.

The procedure established in Algorithm B was repeated 500 times in simulation studies S3 and S4.

- In the first scenario, study S3, 180 time series of length 30 were generated, 90 for each class. Of these series, in each repetition of Algorithm B, n T =112 (which corresponds to 70% of the data) and n C =40 (to avoid problems with LDA and QDA) were used.
- In the second scenario, study S4, 20 time series of length 30 were generated, 10 of each class. Of these series, in each repetition of Algorithm B, n T =14 (which corresponds to 70% of the data) and n C =5 were used. In these cases, the LDA and QDA classifiers were not used.

In Figure 2, we present examples of series simulated in scenarios S3 and S4. The Average and standard deviation of the classification error rates for each method in scenario S3 are recorded in Table 5. From the values presented in Table 5, if we consider confidence intervals with normal approximation for the average error rate and a significance level of 5%, we observe that the performance of the BCDLM was equivalent to that of the RDA, NBND and NBK classifiers. Still, in this scenario, the QDA and NNC methods had poor performance, being the only ones significantly inferior to the others.

In Table 6, the proportions of times that each method presented an error rate lower than or equal to that of another method are described. In comparative terms, we can see that the RDA classifier performed better against all alternatives, followed by BCDLM.

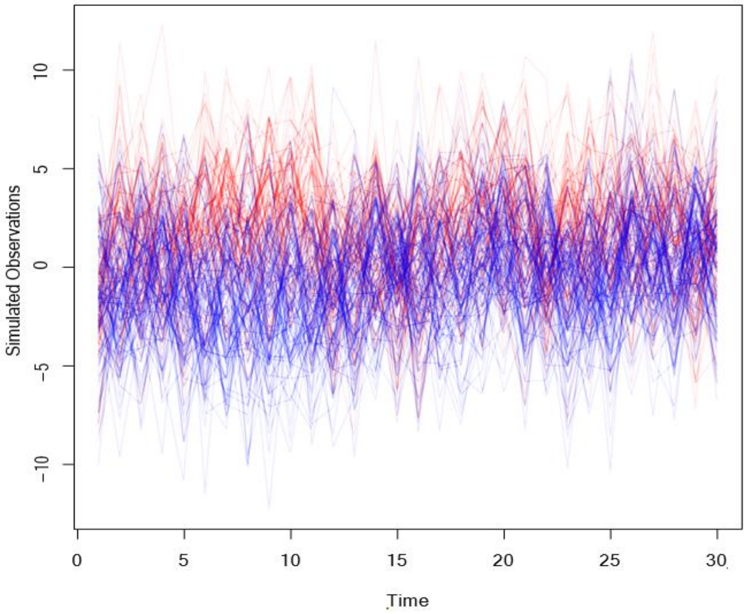


Figure 2. Simulated series with two classes from a polynomial DLM of order 1 for scenarios S3 and S4

The Average and standard deviation of the classification error rates for each method in scenario S4 are recorded in Table 7. From the values presented in Table 7, we have that the BCDLM, NBND and NBK classifiers presented equivalent results in performance. However, the best results were those presented by the RDA classifiers, with the best performance, followed by NNC.

In Table 8, the proportions of times that each method presented an error rate smaller than or equal to that of another method, even in scenario S4, are described. From Table 8, in agreement with the results of Table 7, we observe that the BCDLM classifier presented the performance equivalent to the NBND and inferior to the NNC the RDA. Also noteworthy, in this scenario, is the performance of the RDA superior to all other classifiers.

Table 5. Average and standard deviation of error rates (in %) for scenario S3

Statistic	BCDLM	NNC	LDA	QDA	RDA	NBND	NBK
Average	0.0665	0.1765	0.08	0.2435	0.0592	0.0673	0.0941
Standard deviation	0.0301	0.0453	0.0346	0.0648	0.0294	0.0296	0.0348

Table 6. Performance in terms of error rates (in %) between classifiers for scenario S3

		BCDLM	NNC	LDA	QDA	RDA	NBND	NBK
BCDLM	Equal	-	1	22.2	0.2	26.4	41.4	17.6
	Minor	-	98.2	52.6	99.6	26.6	32	71.6
	Total	-	99.2	74.8	99.8	53	73.4	89.2
NNC	Equal	1	-	2	7.8	0.8	1.2	3.2
	Minor	0.8	-	2.2	76.6	0.6	0.6	4.2
	Total	1.8	-	4.2	84.4	1.4	1.8	7.4
LDA	Equal	22.2	2	-	0.2	20	22.8	17.2
	Minor	25.2	95.8	-	99.4	15.2	27	56.2
	Total	47.4	97.8	-	99.6	35.2	49.8	73.4
QDA	Equal	0.2	7.8	0.2	-	0	0.4	0.4
	Minor	0.2	15.6	0.4	-	0.2	0.2	1.2
	Total	0.4	23.4	0.6	-	0.2	0.6	1.6

		BCDLM	NNC	LDA	QDA	RDA	NBND	NBK
RDA	Equal	26.4	0.8	20	0	-	29	12.4
	Minor	47	98.6	64.8	99.8	-	48.8	77.2
	Total	73.4	99.4	84.8	99.8	-	77.8	89.6
NBND	Equal	41.4	1.2	22.8	0.4	29	-	17.8
	Minor	26.6	98.2	50.2	99.4	22.2	-	70.4
	Total	68	99.4	73	99.8	51.2	-	88.2
NBK	Equal	17.6	3.2	17.2	0.4	12.4	17.8	-
	Minor	10.8	92.6	26.6	98.4	10.4	11.8	-
	Total	28.4	95.8	43.8	98.8	22.8	29.6	-

Table 7. Average and standard deviation of error rates (in %) for scenario S4

	BCDLM	NNC	RDA	NBND	NBK
Average	0.0527	0.0033	0.0013	0.052	0.0653
Standard deviation	0.105	0.0234	0.0149	0.1115	0.1233

Table 8. Performance in terms of error rates (in %) among classifiers for scenario S4

		BCDLM	NNC	RDA	NBND	NBK
BCDLM	Equal	-	73.2	75.2	88.6	64
	Minor	-	2	0.2	5.6	20.2
	Total	-	75.2	75.4	94.2	84.2
NNC	Equal	73.2	-	97.2	75.8	70.8
	Minor	24.8	-	0.8	22.2	27.6
	Total	98	-	98	98	98.4
RDA	Equal	75.2	97.2	-	77.6	72.2
	Minor	24.6	2	-	22.2	27.6
	Total	99.8	99.2	-	99.8	99.8
NBND	Equal	88.6	75.8	77.6	-	66.4
	Minor	5.8	2	0.2	-	19.8
	Total	94.4	77.8	77.8	-	86.2
NBK	Equal	64	70.8	72.2	66.4	-
	Minor	15.8	1.6	0.2	13.8	-
	Total	79.8	72.4	72.4	80.2	-

3. Results and discussion

3.1 Real data applications

In this section, we analyze some time series available in the UCR Time Series Classification [6]. By the time this dissertation was written, there were 85 time series files for classification. For each of them, we performed a visual inspection trying to identify series that could be adjusted with simple dynamic linear models (such as polynomials and trigonometric ones). As the computational analysis demands a lot of time, we selected some sets of data and highlighted three of them in this dissertation for proper comparisons of the classification results using the BCDLM with those of the usual classifiers already mentioned. Such datasets are composed of real-time series (data from SONY AIBO Robot and spectrometry of coffee types) and pseudo-time series (data from Swedish leaves adapted for time series).

3.2 Soil classification by SONY AIBO robot

Returning to the DA problem dataset of the SONY AIBO robot, which is a small dog-shaped quadruped robot equipped with multiple sensors, including a triaxial accelerometer. As already mentioned, we have accelerometer measurements on the horizontal axis that were recorded while the robot walked in circles on two types of surfaces: cement and carpet. Each time series represents a complete loop. A total of 621 turns were recorded, 349 on cement and 272 on carpet, all series 71 in length. Cement is harder than carpet, which makes for more surface variability. Considering each surface as a class, the aim is to classify the series with respect to the two types of surfaces. For each class in the training set, a first-order polynomial dynamic linear model was fitted using the variance strategy ($\mathbb{V}$,$\mathbb{W}$). The series observed with the one-step-ahead forecast are shown in Figure 3.

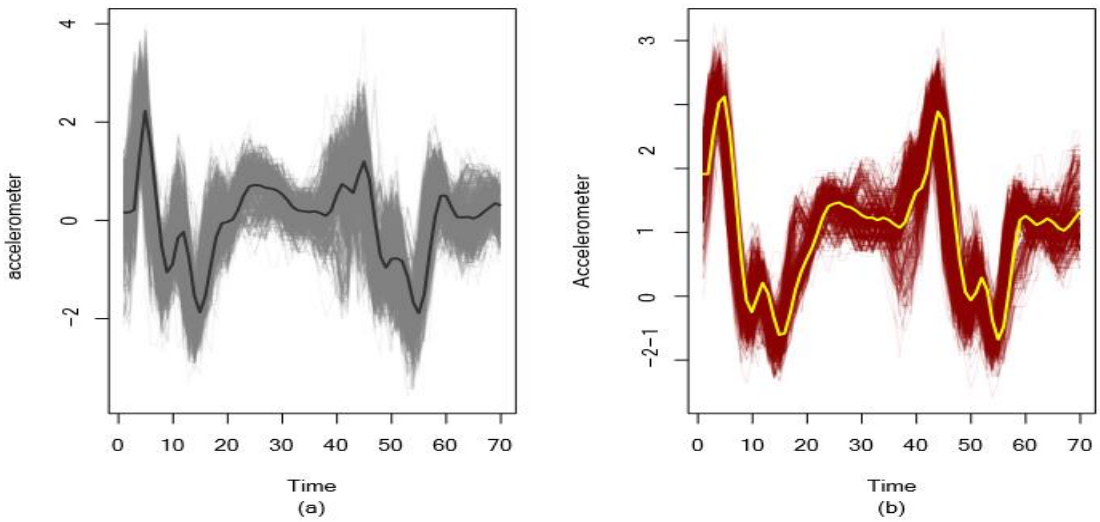


Figure 3. Sony AIBO Robot accelerometer series on the two surfaces (a) Cement and (b) Carpet, with the prediction one step ahead adjusted by the first order polynomial DLM

Once we have chosen an appropriate dynamic linear model, we proceed to evaluate the performance of the classifiers. We carried out a study taking a random sample of time series of size 140 to constitute the training set, and we used the rest as a test set. This procedure was repeated 500 times. Table 9 shows the Averages and standard deviations of the error rates of the classifiers; the results indicate that our classifier was not the best, although if we consider confidence intervals with a normal approximation for the Average error rate and a level of significance of 5%, there is no significant difference between the methods when compared 2 by 2. Table 10 shows the performance of the BCDLM compared to the usual classifiers in terms of the percentage of times it had the same or lesser error rate.

Table 9. Average and standard deviation of error rates (in %) of the classifiers for the SONY AIBO Robot series

	BCDLM	NNC	LDA	QDA	RDA	NBND	NBK
Average	0.03163	0.02577	0.05875	0.33517	0.01488	0.03189	0.02339
Standard deviation	0.012476	0.007846	0.012595	0.080105	0.012529	0.012554	0.009526

Table 10. Performance in terms of error rates (in %) among the classifiers for the SONY AIBO Robot series

		BCDLM	NNC	LDA	QDA	RDA	NBND	NBK
BCDLM	Equal	-	8	1.4	0	2.2	79.6	6.6
	Minor	-	30.8	93.6	100	15.6	15	7
	Total	-	38.8	95	100	17.8	94.6	13.6

		BCDLM	NNC	LDA	QDA	RDA	NBND	NBK
NNC	Equal	8	-	0.6	0	4	8.6	8.6
	Minor	61.2	-	99.2	100	19.6	61.4	34.8
	Total	69.2	-	99.8	100	23.6	70	43.4
LDA	Equal	1.4	0.6	-	0	0	1	0.2
	Minor	5	0.2	-	99.8	0.8	5.4	1.6
	Total	6.4	0.8	-	99.8	0.8	6.4	1.8
QDA	Equal	0	0	0	-	0	0	0
	Minor	0	0	0.2	-	0	0	0
	Total	0	0	0.2	-	0	0	0
RDA	Equal	2.2	4	0	0	-	2	3.2
	Minor	82.2	76.4	99.2	100	-	82.6	76.4
	Total	84.4	80.4	99.2	100	-	84.6	79.6
NBND	Equal	79.6	8.6	1	0	2	-	5.6
	Minor	5.4	30	93.6	100	15.4	-	6.8
	Total	85	38.6	94.6	100	17.4	-	12.4
NBK	Equal	6.6	8.6	0.2	0	3.2	5.6	-
	Minor	86.4	56.6	98.2	100	20.4	87.6	-
	Total	93	65.2	98.4	100	23.6	93.2	-

From this Table 10, we observe that the BCDLM performed worse than NNC, RDA and NBK but with better performance than LDA, QDA and NBND.

Table 10 illustrate the results of the case by case only in the Total item, indicating in black how efficient the first classifier was in relation to the second classifier represented in grey.

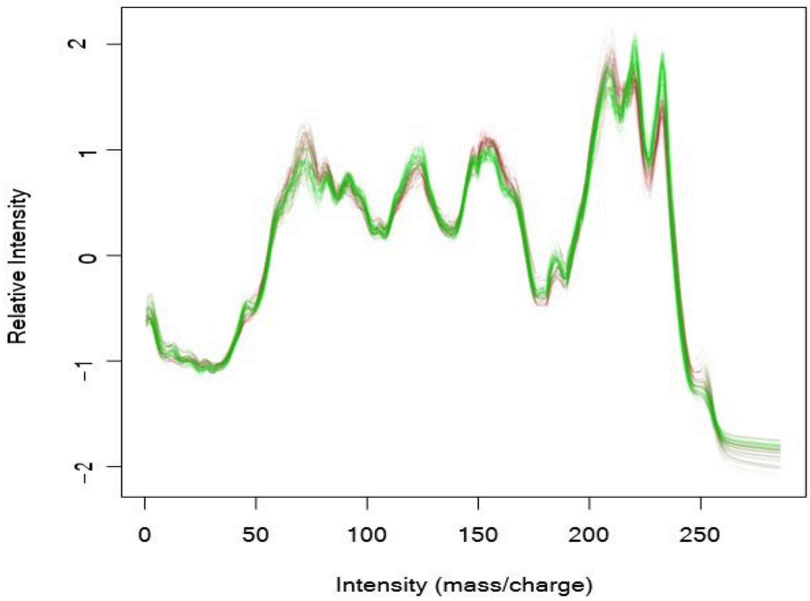


Figure 4. Mass spectrum of coffee samples, *Canephora* (brown) and *Arabica* (green)

The two main species of coffee cultivated in the world are *Arabica* and *Canephora*. They are different in flavour, growing medium and commercial value, with *Arabica* being more expensive than *Canephora*, although the latter is less susceptible to disease. Fifty-six dehydrated and frozen coffee samples were analyzed by mass spectrometry, 29 from the *Canephora* species and 27 from *Arabica*, all series 236 in length.

Mass spectrometry is a technique in which molecules in a sample are converted into ions in gaseous form, which are separated according to the ratio of their mass to their charge. The end result is the mass spectrum - a graph showing the abundance of each intensity (mass/charge). Taking it to be the observed abundance (also called relative intensity) at intensity t, we obtain a mass spectrum as a time series.

For a training sample consisting of 70% of the original series, we again fit a polynomial model of order 1. Then we draw a new training sample at random and rank the remaining ones. We repeat this 500 times. As the size of the series is greater than the number of available series, classified LDA and QDA were not considered. Additionally, we discarded training samples that had less than 10 sets in any of the classes.

The results of the classification of types of coffee are summarized in Tables 9 and 10. In Table 11, it can be seen that the BCDLM method did not present classification errors, while the RDA performed worse than the other classifiers. The NNC classifier obtained the second-best performance, superior to NBK and NBND, and these performances are close to each other. The values in Table 12 show that the NNC classifier in 64.4% of the time presented an error rate equal to that of the BCDLM, that is, without a classification error. On the other hand, the RDA presented a classification error in all repetitions.

The results obtained with the BCDLM in this problem of classifying types of coffee using mass spectrometry data indicate the relevance of this proposal for a classifier. This statement is justifiable since, as mentioned, the NNC classifier is considered the "gold standard" in the literature on time series classification, yet here we have a case where the proposed classifier is superior to the NNC.

Table 11. Average and standard deviation of error rates (in %) of the classifiers for a series of types of coffee

Statistic	BCDLM	NNC	RDA	NBND	NBK
Average	0	0.0214	0.5153	0.0552	0.0782
Standard deviation	0	0.0338	0.0717	0.0546	0.0604

Table 12. Performance in terms of error rates (in %) among the classifiers for a series of types of coffee

		BCDLM	NNC	RDA	NBND	NBK
BCDLM	Equal	-	64.4	0	34	20
	Minor	-	35.6	100	66	80
	Total	-	100	100	100	100
NNC	Equal	64.4	-	0	36.2	26.6
	Minor	0	-	100	54.6	71.2
	Total	64.4	-	100	90.8	97.8
RDA	Equal	0	0	-	0	0
	Minor	0	0	-	0	0
	Total	0	0	-	0	0
NBND	Equal	34	36.2	0	-	41.2
	Minor	0	9.2	100	-	44.2
	Total	34	45.4	100	-	85.4
NBK	Equal	20	26.6	0	41.2	-
	Minor	0	2.2	100	14.6	-
	Total	20	28.8	100	55.8	-

3.3 Classification of Swedish Leaves

We analyzed the data set that we called "Leaves Suecas" (from the original Swedish Leaf) [6]. This dataset is composed of 1125 images of Swedish leaves divided into 15 classes. Each image was converted into a 'pseudo time series' of length 128, where yt is the distance from the tth point from the edge of the leaf to its centroid. In

Figure 5, steps (a), (b) and (c) illustrate the construction of pseudo time series through measurements of the Euclidean distances from the leaf centroid to its edges. Figure 6 and 7 show the time series obtained for the 15 classes.

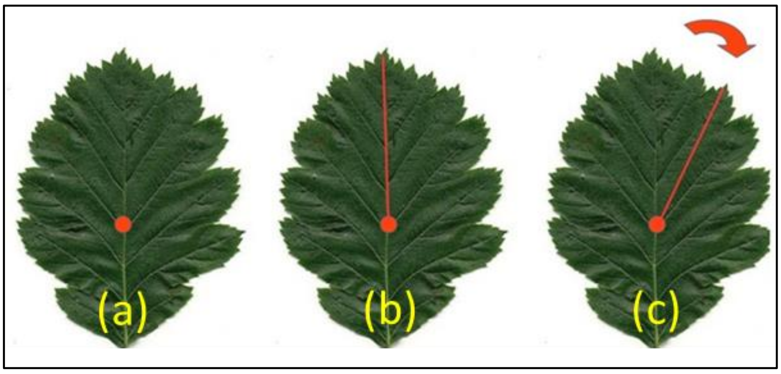


Figure 5. Steps to obtain pseudo time series for Swedish leaves

As in the previous applications, 70% of the time series samples were selected to compose the training set, if each class had at least 35-time series, we built the classifiers and classified the remaining series. This procedure was performed 300 times. After some analysis for a training sample, we identified a seasonal period equal to 128 and chose the harmonics for the trigonometric dynamic linear models:

- Class 1: harmonics 1, 2 and 3
- Class 2: harmonics 2, 3, 4, 5 and 6
- Class 3: harmonics 1, 2 and 3
- Class 4: harmonics 2 and 3
- Class 5: harmonic 2 and 3
- Class 6: harmonic 2

Table 13. Average and standard deviation of error rates (in %) of classifiers for Swedish leaf types

Statistic	BCDLM	NNC	RDA	NBND	NBK
Average	0.1698	0.1956	0.1744	0.17	0.1529
Standard deviation	0.0183	0.0207	0.029	0.0184	0.0184

- Class 7: harmonic 2
- Class 8: harmonics 2 and 3
- Class 9: harmonics 1, 2 and 3
- Class 10: harmonic 2
- Class 11: harmonic 2
- Class 12: harmonics 1, 2, 3, 4, 5, 6 and 7
- Class 13: harmonics 2 and 3
- Class 14: harmonic 2
- Class 15: 2^{nd} and 3^{rd} harmonics.

Table 13 shows the Average and standard deviation of error rates for each classifier, which are graphically represented in Figure 7(A). From this table, it can be seen that the Average error rates are very close (with no significant difference!), although the BCDLM method showed an average (0.1698%) lower only than that of the NBK (0.1529%).

In Table 14, the superiority of the BCDLM is observed, mainly in relation to the NNC and the RDA, since it presented a lower error rate than that of these methods in 82.94% and 50.5% of the times in the repetitions, respectively. Compared with NBK, which had the lowest average error rate, BCDLM still had a lower error rate of 9.36% of the time. The results of this Table, in the total item, are graphically illustrated in Figure 7B.

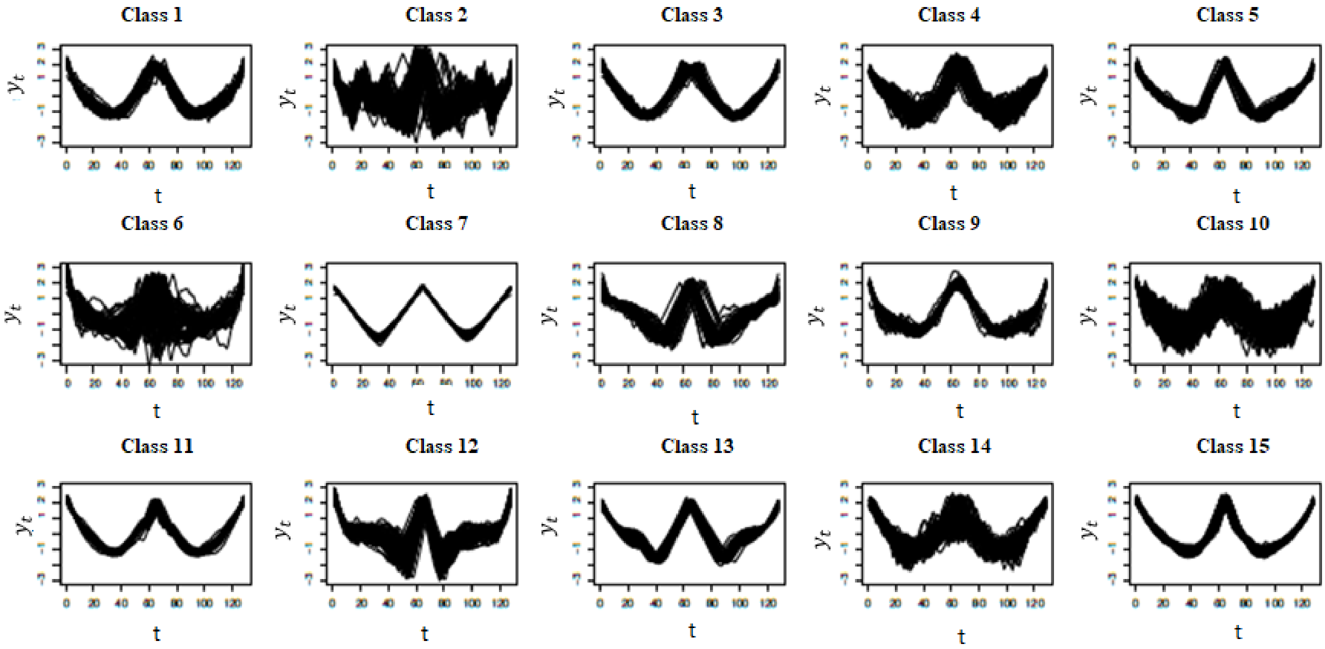
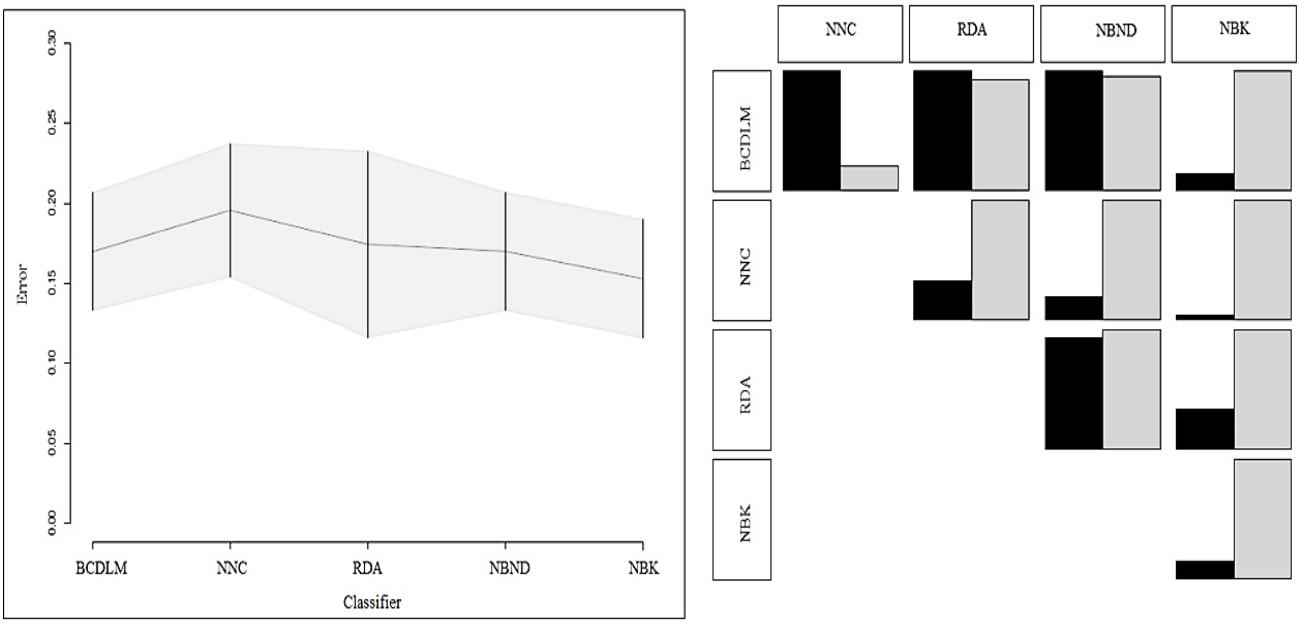


Figure 6. Pseudo time series obtained for the 15 types of Swedish leaves

Table 14. Performance in terms of error rates (in %) among classifiers for series of Swedish leaf types

		BCDLM	NNC	RDA	NBND	NBK
BCDLM	Equal	-	2.457	3.1605	54.075	3.1605
	Minor	-	87.087	53.025	27.384	9.828
	Total	-	89.544	56.1855	81.459	12.988
NNC	Equal	2.457	-	2.457	2.8035	0.693
	Minor	15.4455	-	23.52	14.742	2.8035
	Total	17.9025	-	25.977	17.545	3.4965
RDA	Equal	3.1605	2.457	-	5.2605	2.8035
	Minor	48.804	79.012	-	48.100	23.877
	Total	51.964	81.469	-	53.361	26.680
NBND	Equal	54.075	2.8035	5.2605	-	5.2605
	Minor	23.52	87.433	51.618	-	9.1245
	Total	77.595	90.237	56.8785	-	14.385
NBK	Equal	3.1605	0.693	2.8035	5.2605	-
	Minor	92.001	101.483	78.309	90.594	-
	Total	95.161	102.176	81.112	95.854	-



(A) (B)

Figure 7. (A) Comparison of Confidence Intervals of classifier error rates for series of Swedish leaf types; (B) Comparison of error rates (in %) between classifiers for series of Swedish leaf types

4. Conclusions

In this work, we present a new approach for discriminant analysis (DA) of time series, proposing a version for Bayes classifier employing Dynamic Linear Models. We carry out simulation studies with simple models that are useful in practice. Such studies, although not exhaustive, suggest that the BCDLM is configured as an efficient proposal for a classifier, provided that the appropriate Dynamic Linear Model is used. In these simulation studies carried out, we established comparisons between some variance estimation strategies or covariance matrix: $\mathbb{V}_t$ estimated in each shelf t, and requested through discount factor ($\mathbb{V}$,Δ); $\mathcal{V}_t$ considered fixed and requested through factor of discount(φ,Δ); $\mathbb{V}_t$ estimated on each shelf, and $\mathbb{W}_t$ considering superimposed models ($\mathbb{V}$,$\mathbb{W}$). The simulation studies carried out with these strategies, analyzing the error rates of the classifier, indicated the strategy ($\mathbb{V}$, $\mathbb{W}$) as the one that presented the best results.

The integration of Dynamic System Linear Models and the Bayes Classifier for time series classification holds immense potential for fostering sustainability. By accurately predicting future behaviors, optimizing resource utilization, and enabling proactive interventions, this approach contributes to more informed and responsible decision-making across various sectors. As data collection and analysis techniques advance, the synergy between DLMs and the Bayes Classifier will continue to play a pivotal role in shaping a more sustainable future.

As a disadvantage for BCDLM, we can mention the computational cost of application in the adjustment phase (classifier training) using cross-validation. In the cross-validation process, the classifier is estimated as many times as the number of observations in the training set; this Averages that the matrix $\mathbb{W}$ must be estimated in all these repetitions, which represents a high computational cost. However, after adjusting the model parameters in BCDLM, this classifier can be used in practice for several problems whose observations come from time series.

In general, considering the real-time series analyzed in this work, we can state that the BCDLM proved to be competitive compared to the results of the classifiers NNC, RDA, NBND and NBK and superior to the methods LDA and QDA. It is important to note that both NNC and NBK, which are recognized in the literature as classified.

Declaration of competing interest

The authors declare that they have no known financial or non-financial competing interests in any material discussed in this paper.

Funding information

No funding was received from any financial organization to conduct this research.

Definitions

LDA= Linear Discriminant Analysis
QDA=Quadratic Discriminant Analysis,
RDA=Regularized Discriminant Analysis
NBND=Naive Bayes with Normal Distribution,
NBHFF=Naive Bayes with Head Function Estimators
NNC=Nearest Neighbor Classifier
NBK= Naive Bayes Kernel
BCDLM= Bayes classifier employing Dynamic Linear Models

References

[1] A. El-Habil and M. El-Jazzar, "A comparative study between linear discriminant analysis and multinomial logistic regression," *An-Najah Univ J Res-B (Humanities)*, vol. 28, no. 6, pp. 1525–1548, 2014.

[2] J. K. W. Ognian Asparoukhov, "A comparison of discriminant procedures for binary variables," *Computational Statistics & Data Analysis*, vol. 38, no. 2, pp. 139–160, 2001.

[3] H. A. Jumaa and M. M. Abed, "Co-retractable modules and semi-simplicity property:," *sjpas*, vol. 5, no. 2, pp. 151–159, 2023.

[4] A. A. Hamad *et al.*, "Secure complex systems: A dynamic model in the synchronization," *Comput. Intell. Neurosci.*, vol. 2021, p. 9719413, 2021.

[5] F. M. Ahmed and D. B. S. Mohammed, "Feasibility of breast cancer detection through a convolutional neural network in mammographs," *Tamjeed Journal of Healthcare Engineering and Science Technology*, vol. 1, no. 2, pp. 36–43, 2023.

[6] O. F. Khalaf and N. A. H. Al-Samarrai, "Using information systems and modern technologies to know the reasons for the disappearance of archaeological sites and the extinction of their features. Samarra Journal of Pure and Applied Science," *sjpas*, vol. 5, no. 1, pp. 118–125, 2023.

[7] F. T. Jasim and Karthick, "Artificial intelligence innovation and human resource recruitment," *Tamjeed Journal of Healthcare Engineering and Science Technology*, vol. 1, no. 2, pp. 20–29, 2023.

[8] A. S. Al-Obeidi, S. Fawzi Al-Azzawi, A. Abdullah Hamad, M. L. Thivagar, Z. Meraf, and S. Ahmad, "A novel of new 7D hyperchaotic system with self-excited attractors and its hybrid synchronization," *Comput. Intell. Neurosci.*, vol. 2021, p. 3081345, 2021.

[9] E. P. J. Bickel, "Some theory for Fisher's linear discriminant function,'naive Bayes', and some alternatives when there are many more variables than observations," *Bernoulli*, pp. 989–1010, 2004.

[10] G. Saggio *et al.*, "Comparison of different classifiers on a reduced set of features for mental tasks-based brain-computer interface," in *Proceedings of the Third International Conference on Bio-inspired Systems and Signal Processing*, vol. 0, Valencia, Spain, 2010, pp. 174–179.

[11] P. D. S. Jha, M. F. Khalf, and Karthick, "Convolutional Neural Networks for breast cancer detection using regions of interest from infrared images," *Tamjeed Journal of Healthcare Engineering and Science Technology*, vol. 1, no. 2, pp. 44–53, 2023.

This page intentionally left blank.