Preprints.org

Article

Not peer-reviewed version

Geometric Path Features (GPF) via the Generalized Cross Product for Time Series Embeddings

Yunis Kahalan * and Samir Belhaouari

Posted Date: 23 March 2026

doi: 10.20944/preprints202603.1778.v1

Keywords: time series classification; time series regression; multivariate time series; sequential data; time series representation learning; path signatures

Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a Creative Commons CC BY 4.0 license, which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 March 2026 doi:10.20944/preprints202603.1778.v1

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

Geometric Path Features (GPF) via the Generalized Cross Product for Time Series Embeddings

Yunis Kahalan * and Samir Belhaouari

Division of Information and Computing Technology, Hamad Bin Khalifa University, Doha 34110, Qatar
* Correspondence: yuka34154@hbku.edu.qa

Abstract

Multivariate time series such as EEG or seismic data can be viewed as parameterized curves in a multidimensional space. Path signatures achieve state of the art performance on many machine learning tasks by embedding curves into geometric feature spaces. But at higher degrees, their magnitudes decay factorially and become difficult to interpret. Inspired by Lévy areas, we propose a novel generalization of path signatures where features are defined between pairs or groups of curves. These features are (i) areas of ruled surfaces formed by projecting curves into pairs of same dimension subspaces (ii) volumes enclosed by surfaces from projections into multiple subspaces of same dimension, computable through the generalized cross product. Unlike path signatures of high degree, the features remain interpretable geometrically even in high dimensional settings while also capturing richer interactions across subsets of channels of the time series. Our generalized features can complement path signatures of low degree or be used independently. We demonstrate this across 143 classification and 45 regression datasets where they consistently improve over baseline path signatures where in some cases, accuracy gains of up to +45%, and MSE reduction up to 14.2x are achieved, while providing a richer and sparse framework for time series embeddings.

Keywords: time series classification; time series regression; multivariate time series; sequential data; time series representation learning; path signatures

1. Introduction

In the last two decades, there has been an increasing interest in time series analysis research, stemming from multiple application domains where time series data are generated at a fast-growing rate and scale. Some examples are stock market data in financial applications, smartphone and GPS sensor data, and medical data like electrocardiograms or electroencephalograms. These high-dimensional datasets open up new theoretical and practical challenges, as techniques need to be adapted to their sequential nature. What sets time series data apart from other datasets is the natural order of the values and the correlation among these variables that must be considered by the ML algorithms [1]. To deal with this, distance based methods have been used such as One Nearest Neighbour Classifier with Dynamic Time Warping and Proximity Forest where features are measures of similarity through pairwise comparison of the whole time series [1]. But methods have been popularized which are feature based, for example taking nearly 800 features from the TSFresh toolbox [2] with it being popular enough to be reduced to 22 features as Catch22 [3] or QUANT, [4], or interval based methods like Random Interval Forest which summarise statistics over random intervals [5]. It is noted that dilation based methods like WEASEL V2.0 [6], BOSS [7], Shapelets [8], RDST [9], and related methods such as gRSF [10] are strong classifiers, with the highly popular convolution based classifiers falling under this category such as Hydra [11], ROCKET [12], MINIROCKET [13], and MultiROCKET [14]. Due to the recent trend of deep learning, deep learning models have also been constructed for use in time series analysis such as Inception [15], MUSE [16], Tapnet [17], and MLCN [18].

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 March 2026 doi:10.20944/preprints202603.1778.v1

Recently, a feature based method, path signature [19–21] has found many applications in time series analysis, especially in analyzing non-smooth or irregular signals with logistic regression [22], for handwriting recognition with neural networks [23], image recognition [24], or to generate multivariate time series with a Generative Adversarial Network [25]. Unlike the previous methods, signatures captures higher order dependencies and handles irregularities. The advantage of signatures compared to other feature based methods is that it has desirable attributes which are theoretically guaranteed. For example, a function whose domain is defined by path signatures can be well approximated using linear functions [26].

In this work, we build on the Lévy area [27], a difference of path signature terms of multivariate time series to propose its extension to multidimensional time series, and introduce new signal features called Geometric Path Features (GPF) based on these multivariate extensions. Our contributions are:

1. We propose a set of new features that are information dense and geometrically interpretable, based on path signatures that summarize geometric information (Lévy area).
2. We extend these features using the generalized cross product to capture higher dimensional interaction and we show that these features are compatible with extensions of signatures such as dyadic techniques and also introduce preprocessing techniques using Chord and Curvature.

2. Preliminaries

2.1. Definitions and Main Properties of Path Signatures

In this subsection, the notion of signature is introduced and some of its important properties are reviewed. The reader is referred to [20] or [21] for a more involved mathematical treatment with proofs. Throughout the paper, the basic objects are paths, that is, functions from $[0,1] \to \mathbb{R}^d$, where $d \in \mathbb{N}^*$. The main assumption is that these paths are of bounded variation, e.g. they have finite length.

Definition 1. *Let X_t be a d-dimensional path indexed by time T. The path X_t is of finite p-variation for certain $p \geq 1$ if the p-variation of X_t is defined by*

$$\|X_t\|_{p-var} = \left(\sup_{\mathcal{D}} \sum_{t_i \in D} \|X_{t_i} - X_{t_{i-1}}\|^p \right)^{1/p} < \infty$$

where the supremum is taken over all possible finite partitions $\mathcal{D} = \{(t_1, ..., t_l) | l \geq 1, t_0 = 0 < t_1 < ... < t_{l-1} < t_l = 1\}$, and $\|.\|$ denotes the Euclidean norm on $\mathbb{R}^d$. The path X_t is said to be of bounded variation if its total variation is finite.

The consequence of bounded variation is covered in Appendix A. In this section, a practical description of the path signature is presented.

Definition 2. *The path signature coefficient for the time series or path of bounded variation $X_t : T \to \mathbb{R}^d$ of degree $k \geq 1$ and the multi-index $(i_1, ..., i_k) \in \{1, ..., d\}^k$ with length k is defined,*

$$\begin{aligned} \Phi(X_t)^{i_1,...,i_k}_{[0,t]} &= \int_{0<s<t} \Phi(X_t)^{i_1,...,i_{k-1}}_{[0,s]} dX^{i_k}_s \\ &= \int_0^t ... \Big(\int_0^{t_3} \Big(\int_0^{t_2} dX^{i_1}_{t_1} \Big) dX^{i_2}_{t_2} \Big) ... dX^{i_k}_{t_k} \end{aligned}$$

Then $\Phi(X_t)^{i_1,...,i_k}_{[0,t]}$ is called the k-fold iterated integral of X_t over the indices, and the integration limits correspond to the integration over the simplex. For any $k \geq 1$, the signature truncated at degree k is the infinite series that contains all possible terms and is denoted by

$$\Phi(X_t)_{0,1} = (1, \Phi(X_t)^1_{[0,1]}, \dots, \Phi(X_t)^d_{[0,1]}, \Phi(X_t)^{1,1}_{[0,1]}, \dots)$$

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 March 2026 doi:10.20944/preprints202603.1778.v1

where the zeroth term by convention is 1. For the truncated path signature of degree k,

$$\Phi^k(X_t)_{[0,1]} = (1, \Phi(X_t)^1_{[0,1]}, \ldots, \Phi(X_t)^d_{[0,1]},$$
$$\Phi(X_t)^{1,1}_{[0,1]}, \ldots, \Phi(X_t)^{d,d}_{[0,1]}, \ldots, \Phi(X_t)^{\overbrace{d,\ldots,d}^{\text{k terms}}}_{[0,1]})$$

Both the discrete and continuous signature are calculated through example in Appendix B and C for clarity. It is important to mention that, for a path in $\mathbb{R}^d$, there are d^i terms of degree i. The signature truncated at degree k is therefore a vector of dimension $\sum_{i=0}^{k} d^i = \frac{d^{k+1}-1}{d-1}$ if $d \neq 1$, growing exponentially with k. For instance, consider a time series with $d = 6$ features and a path signature of degree $k = 7$, then the resulting signature space has 335922 dimensions. The path signature is a high-dimensional embedding of the time-series capturing its intricate temporal dependencies and patterns.

3. Related work

Concepts:

As stated above, path signature was first introduced in [19] for smooth paths. Later, it was developed for rough paths by Lyons and others [28]. Signature can capture higher order dependencies in time series data, providing a more nuanced understanding of the underlying structure and dynamics. It is invariant under reparameterization of the time axis, it contains information about the joint evolution of tuples of coordinates [23] and possesses several properties that make it a good statistical summary of paths. Finally, it can handle irregularly sampled data, making them suitable for time series with uneven time intervals between observations.

Applications:

Path signature gained popularity in the ML community recently and was successfully used in multiple applications. Moore et al. (2019) used path signature to classify patients with Alzheimer's disease using logistic regression [22]. Yang et al. (2016) combined it with neural networks in handwriting recognition [23], Ibrahim and Lyons, (2022) used it for image recognition [24]. Lastly, Ni et al. (2021) proposed a Generative Adversarial Network capable of generating time series data of variable lengths and dimensions [25].

Dimensionality:

Handling and interpreting high-dimensional signatures can be challenging and can lead to increased computational demands and potential issues with overfitting. In addition to this, increasing the degree or dimensionality of signatures to high degrees won't necessarily lead to a significant improvement in performance due to factorial decay [29] which suggests that predictive power of path signatures lie in lower degrees.

4. Methodology

4.1. Lévy-based Area Between a Curve and a Line

We begin with the standard path signature on a univariate time series and then state its relation with the Lévy area.

Lemma 1. *For a one-dimensional time series X_t on $t \in [a, b]$, the features of the signature transformation of X_t of degree k, $\Phi_k(X_t)$ are given by,*

$$\left(1, X^1_b - X^1_a, \frac{(X^1_b - X^1_a)^2}{2!}, \ldots, \frac{(X^1_b - X^1_a)^k}{k!}\right)$$

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 March 2026 doi:10.20944/preprints202603.1778.v1

For a multivariate linear time series $X_t = (X_t^1, X_t^2, X_t^3, \ldots, X_t^d)$ *on* $t \in [a, b]$*, then the coefficient of degree* k *of the signature is given by,*

$$\Phi^{(i_1,i_2,\ldots,i_k)}(X_t)_{[a,b]} = \frac{1}{k!}\prod_{j=1}^{k}(X_b^{i_j} - X_a^{i_j})$$

We prove this Lemma in Appendix D. If X_t is of bounded p-variation with $1 < p < 2$, then the signature is still well defined using the Young integral. If $p \geq 2$, the iterated integrals are no longer uniquely defined. This is where rough path theory comes into play to show their existence for various stochastic processes of the Brownian motion [27]. It is noted that $\dot{X}_t$ denotes the derivative of X_t with respect to t,

$$\dot{X}_t = \frac{dX_t}{dt}.$$

Theorem 1. *Let* $X_t = \{X_t^1, X_t^2, \ldots, X_t^d\} \in \mathbb{R}^d$ *a time series defined in the interval* $[a, b]$ *with dimension* d*. Lévy signatures are the signed (S) and unsigned areas (U) of any two dimensions* $(i_1, i_2) \in \{1, 2, \ldots, d\}^2$ *of* X_t*, which is defined as the enclosed area between the 2-dimensional curve* $\{x = x_{i_1}, y = x_{i_2}\}$ *and the straight line connecting its endpoints of equation* $y = x$ *(Figure 1) where unsigned area is the absolute value.*

$$A_{\text{signed}} = \frac{1}{2}\int_0^{t_1}\left(\dot{X}_t^2 X_t^1 - X_t^2 \dot{X}_t^1\right)dt,$$

$$A_{\text{unsigned}} = \frac{1}{2}\int_0^{t_1}\left|\dot{X}_t^2 X_t^1 - X_t^2 \dot{X}_t^1\right|dt.$$

The proof is given in detail in Appendix F. In short, the signature is an infinite vector corresponding to the iterated integrals of X_t against itself on a simplex. It has a natural interpretation in terms of areas under the path, as shown in Figure 1. The first degree terms are just the increments of the path: for any $1 \leq i \leq d$, $\Phi^{(i)}(X_t) = X_b^i - X_a^i$. The second degree terms correspond to areas under the curve delimited by each pair of coordinates. Moreover, the Lévy area, which is the signed area between a curve and the line connecting its two endpoints, can be recovered from these second degree terms.

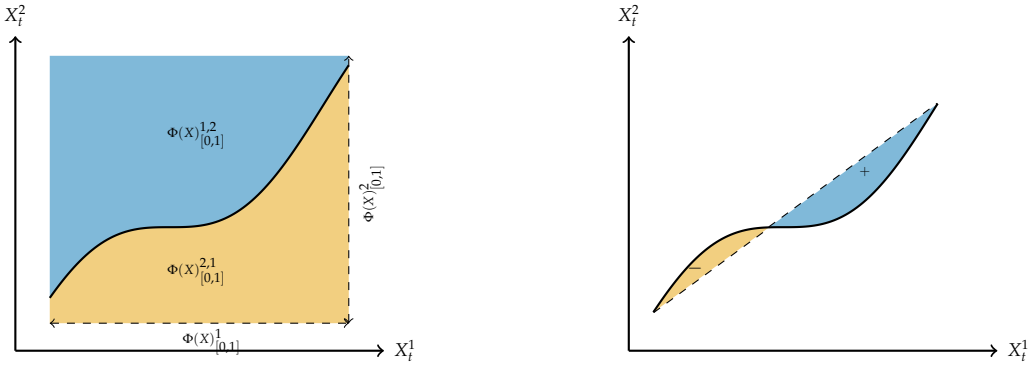

Figure 1. Left: Degree 1 and 2 terms correspond to increments and areas. **Right**: The Lévy area corresponds to the orange and blue areas.

In the following, we will extend the same idea to derive Lévy based features using the Lévy area between two curves and more, then use them as a feature based on path signature noted by GPF_S, GPF_U and GPF_{SU}.

4.2. Proposed Lévy Based Area Methods

In the previous section, we have shown the geometric nature of the Lévy area, which is the enclosed area between a 2-dimensional parametric curve and the straight line connecting its endpoints (Chord), based on that, a new feature complementing path signature features and adding information about the relative variation between different dimensions of the curve. This section proposes a

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 March 2026 doi:10.20944/preprints202603.1778.v1

generalization of the Lévy area, that computes the enclosed area between pairs of curves and volume between groups of curves.

4.2.1. GPF: Lévy Area in Two Dimensions

The case of an arbitrary pair of 2-dimensional parametric curves is considered in this section.

Theorem 2. *Let $X_t = \{X_t^1, X_t^2, \ldots, X_t^d\} \in \mathbb{R}^d$ a time series defined in the interval $[a,b]$ with dimension d. Let $\Phi(X_t)_{[a,b]}$ be a signature of X_t. Then the signed and unsigned GPF, which is the enclosed area between any two 2-dimensional parametric curves from X_t, $C_1 = \{x = X_t^i, y = X_t^j\}$ and $C_2 = \{x = X_t^k, y = X_t^l\}$ where $(i,j,k,l) \in \{1,2,\ldots,d\}^4$ (Figure 2 Left), is derived from the second degree signature terms as follows:*

$$\begin{aligned} A = \int_a^b dA &= \int_a^b (X_t^j \dot{X}_t^i - X_t^i \dot{X}_t^j + X_t^k \dot{X}_t^j - X_t^l \dot{X}_t^i) dt \\ &= \Phi^{(j,i)}(X_t) + \Phi^{(k,j)}(X_t) - \Phi^{(i,j)}(X_t) - \Phi^{(l,i)}(X_t) \end{aligned}$$

Following Theorem 2, the signed and unsigned *GPF* features are given by:

$$\mathrm{GPF}_S = \int_a^b (X_t^j \dot{X}_t^i - X_t^i \dot{X}_t^j + X_t^k \dot{X}_t^j - X_t^l \dot{X}_t^i) dt$$

$$\mathrm{GPF}_U = \int_a^b \left| X_t^j \dot{X}_t^i - X_t^i \dot{X}_t^j + X_t^k \dot{X}_t^j - X_t^l \dot{X}_t^i \right| dt$$

This is a new feature called $\Phi^{(i,j);(k,l)}$. The proof is given in the Appendix G.

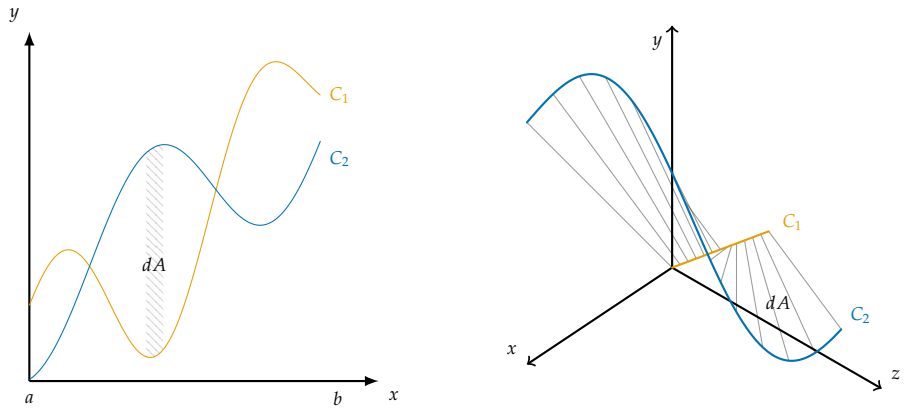

Figure 2. Left: Area element dA between two 2D curves C_1 and C_2 which is a geometric interpretation of the area between both channels (*GPF*). **Right:** Area between two 3D curves C_1 and C_2 where line segments represent instances of the area element dA.

4.2.2. GPF Vector: Lévy Area Between Pairs of 3-Dimensional curves

We now start to generalize into pairs of parametric curves of higher dimensions and include the proof of the general case. Proofs in this section

Theorem 3. *Let $X_t = \{X_t^1, X_t^2, \ldots, X_t^d\} \in \mathbb{R}^d$ a time series defined in the interval $[a,b]$ with dimension d. Let $\Phi(X_t)_{[a,b]}$ be a signature of X_t. Then the GPF Vector, which is the enclosed area between any two 3-dimensional parametric curves from the multidimensional time series X_t, $C^1 = C^{i_1,i_2,i_3} = \{x = X_t^{i_1}, y = X_t^{i_2}, z = X_t^{i_3}\}$ and $C^2 = C^{i_4,i_5,i_6} = \{x = X_t^{i_4}, y = X_t^{i_5}, z = X_t^{i_6}\}$ where for $(i_1,i_2,i_3,i_4,i_5,i_6) \in \{1,2,\ldots,d\}^6$ (Figure 2 Right), is given by*

$$\Phi^{(i_1,i_2,i_3);(i_4,i_5,i_6)} = \int_a^b dA = \int_a^b \sqrt{\Lambda^2 + \Psi^2 + \Omega^2}\, dt$$

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 March 2026 doi:10.20944/preprints202603.1778.v1

$$\begin{aligned} with \quad \Lambda =& \dot{X}_t^{i_2}(X_t^{i_6} - X_t^{i_3}) - \dot{X}_t^{i_3}(X_t^{i_5} - X_t^{i_2}) \\ \Psi =& \dot{X}_t^{i_1}(X_t^{i_6} - X_t^{i_3}) - \dot{X}_t^{i_3}(X_t^{i_4} - X_t^{i_1}) \\ \Omega =& \dot{X}_t^{i_1}(X_t^{i_5} - X_t^{i_2}) - \dot{X}_t^{i_2}(X_t^{i_4} - X_t^{i_1}) \end{aligned}$$

We provide a sketch of proof below for the general case, but its full form is in Appendix I with proof of the specific case here in Appendix H.

Proof. Consider we have the d-dimensional path $X_t : [a,b] \to \mathbb{R}^d = (X_t^1, X_t^2, ..., X_t^{2p-1}, X_t^{2p})$. Assume we split the path uniquely into two p-dimensional paths for $(i_1, ..., i_{2p}) \in \{1, 2, \ldots, d\}^{2p}$: $C_1(t) = \{X_t^{i_1}, ..., X_t^{i_p}\}$ and $C_2(t) = \{X_t^{i_{p+1}}, ..., X_t^{i_{2p}}\}$. Then the same signature feature can be used as it is invariant to dimension of the paths by calculating with a norm of the cross product as shown in proof of Proposition 5 [30],

$$\begin{aligned} A = \int_a^b dA = \int_a^b \sqrt{M} = \Phi^{(1,2,\ldots,p);(p+1,p+2,\ldots,2p)}, \\ M = \left|\left|\overrightarrow{C_1(t)C_1(t+dt)}\right|\right|^2 \cdot \left|\left|\overrightarrow{C_1(t)C_2(t)}\right|\right|^2 \\ - \left\langle \overrightarrow{C_1(t)C_1(t+dt)}, \overrightarrow{C_1(t)C_2(t)} \right\rangle^2 \end{aligned}$$

□

It is important to observe that based on how we encode the pairs of k-dimensional vectors, we can have P_{2k}^d features without repetition or with repetition, d^{2k}. We now demonstrate how to expand these features by considering tuples of curves or more through the generalized cross product.

4.2.3. GPF: Lévy Volume Between p Curves That Are p-Dimensional

This is our last feature which will allow us to take the volume between multiple parametric curves in higher dimensions.

Theorem 4. *Let $X_t = \{X_t^1, X_t^2, \ldots, X_t^d\} \in \mathbb{R}^d$ a time series defined in the interval $[a,b]$ with dimension d. Let $\Phi(X_t)_{[a,b]}$ be a signature of X_t. Then the signed and unsigned GPF, which is the enclosed volume between any p p-dimensional parametric curves from the multidimensional time series X_t, $C^1 = C^{i_1,i_2,\ldots,i_p} = (X_t^{i_1}, X_t^{i_2}, ..., X_t^{i_p})^T$,..., $C^p = C^{i_{p(p-1)+1},\ldots,i_{p^2}} = (X_t^{i_{p(p-1)+1}}, ..., X_t^{i_{p^2}})^T$ for $(i_1, ..., i_{p^2}) \in \{1, 2, \ldots, d\}^{p^2}$, is given by*

$$\begin{aligned} \Phi^{(1,2,..,p);...;(p^2-p+1,p^2-p+2,...,p^2)} = \int_{t_1}^{t_2} dA = \int_{t_1}^{t_2} M dt, \\ M = \left|\left| cross\left(\overrightarrow{C_1(t)C_2(t)}, \overrightarrow{C_1(t)C_3(t)}, ..., \overrightarrow{C_1(t)C_p(t)}\right)\right|\right| \end{aligned}$$

We replicate the proof in Appendix J, but for better intuition and understanding it is suggested to read Appendix J in detail.

Proof. The notion of parallelogram can be used again through cross product but needs to be in its general form in high dimension as in Definition 1 of [30]. Therefore, the area element M can be written as,

$$\begin{aligned} M &= \left|\left| cross\left(\overrightarrow{C_1(t)C_2(t)}, \overrightarrow{C_1(t)C_3(t)}, ..., \overrightarrow{C_1(t)C_p(t)}\right)\right|\right| \\ &= \left|\left| cross\left(\begin{pmatrix} X^{i_{p+1}} - X_t^{i_1} \\ \vdots \\ X^{i_{2p}} - X_t^{i_p} \end{pmatrix}, ..., \begin{pmatrix} X_t^{i_{p^2-p+1}} - X_t^{i_1} \\ \vdots \\ X_t^{i_{p^2}} - X_t^{i_p} \end{pmatrix}\right)\right|\right| \end{aligned}$$

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 March 2026 doi:10.20944/preprints202603.1778.v1

Where the cross product is computed with a determinant as per Definition 1 leading to the above cross product evaluated as below,

$$\left\| \begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \dots & \mathbf{e}_p \\ X_t^{i_{p+1}} - X_t^{i_1} & X_t^{i_{p+2}} - X_t^{i_2} & \dots & X_t^{i_{2p}} - X_t^{i_p} \\ \vdots & \vdots & \dots & \vdots \\ X_t^{i_{p^2-p+1}} - X_t^{i_1} & X_t^{i_{p^2-p+2}} - X_t^{i_2} & \dots & X_t^{i_{p^2}} - X_t^{i_p} \end{vmatrix} \right\|$$

Now, the signed and unsigned signature feature is defined respectively,

$$GPF_S = \int_{t_1}^{t_2} dA = \int_{t_1}^{t_2} M dt$$
$$= \Phi^{(1,2,..,p);(p+1,p+2,...,2p);...;(p^2-p+1,p^2-p+2,...,p^2)}$$

$$GPF_U = \int_{t_1}^{t_2} |dA|$$

□

It is noted that to have the signed feature, we avoid calculating the element dA as a vector norm, we do this by considering the case where C_1 is a curve of zeros, and assume a specific configuration of the curves such that for C_i, all entries are empty except at $i-1$ which is X_t^i. This results in a determinant calculation which is just the product of the channels at time t where we can take its sum or sum of its absolute values to get signed and unsigned features respectively.

4.2.4. Curvature Transform

When dealing with univariate time series, we require another channel so that we can compute GPF features, to achieve this, the Curvature Transform is used,

Definition 3. *For a one-dimensional time series X_t on $t \in [a,b]$, The curvature transform results in a one-dimensional time series $r(t)$ which assigns a local curvature radius at each point. This is given by $r(t) = \frac{(1+\dot{X}_t^2)^{3/2}}{|\ddot{X}_t|}$.*

For the discrete case, we define the values at endpoints a, and b as, $r(a) = r(a+1)$, and $r(b) = r(b-1)$ respectively. The curvature radius at each point indicates smoothness, where a high value indicates smooth properties (acceleration or $\ddot{X}_t$ is small), and a low value indicates rapid change in slope ($\ddot{X}_t$ is high).

4.2.5. GPF vs Path Signature

First, GPF is robust to noise and missing values as it is designed to capture the geometric structure of paths rather than specific point values and the integration processes incorporate the available information without requiring consistent intervals. Second, path signature methods are limited by their reliance on paths of bounded variation, the proposed Lévy signature generalization offers a more robust and adaptable approach for modeling time series data, particularly in scenarios where nonlinear patterns and sudden shifts are prevalent.

In addition, path signatures suffer from lack of geometric interpretability of terms at higher degrees with the additional drawback that as the degree increases, the terms start experiencing factorial decay thus become factors smaller compared to terms of lower degree contributing little to performance [31]. GPF holds the advantage that it is interpretable as each term is either an area or volume representing cross channel interaction, and all features are significant (don't factorially decay). The exponential growth $O(Ld^k)$ of k degree path signatures in the number of generated features with the increase in truncation degree is noted for a d-dimensional $(d > 1)$ time series with L observations in each dimension whilst for GPF it is $O(L^2 d)$. Along with this, the space complexity is $\frac{d^{k+1}-1}{d-1}$ and

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 March 2026 doi:10.20944/preprints202603.1778.v1

$\sum_{i=1}^{d} \binom{d}{i}$ for k degree path signatures and GPF respectively. In the end of Appendix J, we show how to generate even more features so that features explode faster than path signatures due to being a generalization of path signature methods.

5. Illustrations

5.1. Experimental Setup

The proposed solution is evaluated on two main machine learning tasks. In each task, the proposed solution was compared to the path signature method either as an independent method or as features appended to path signature features. In every experiment we set the path signature degree to 8 for the univariate case and degree to 4 for the multivariate case and use a Random Forest Classifier/Regressor with 5-fold validation. All experiments were done using only CPU Google Colab and for the Comparison With Other Feature Sets we use only Kaggle GPU P1 compute. For every task, we have statistical validation of our features either independently or appended to path signature features in Appendix K verifying that our features can work alone or to complement path signatures features. We also have dataset details in Appendix L for completeness.

Classification:

We evaluate GPF in the case of univariate time series classification using the UCR archives which consists of 128 datasets. And we evaluate GPF on a subset of the UEA archive for multivariate time series classification using 15 datasets. In the univariate case we evaluate using both Curvature (Lévy area between a univariate channel and its Curvature) and Chord + Curvature (concatenating Lévy areas between a univariate channel and its Chord/Curvature) features. In the multivariate case, we take Curvature and Chord + Curvature features by taking GPF along dyadic lengths of the time series and concatenate them, for this we consider the whole series, halves of the series, and quarters of the series for our dyadic features (dyadic level 2). Our metric for this task is accuracy.

Regression:

In the univariate regression case we evaluate GPF with both Curvature and Chord + Curvature features on 26 datasets from the Time Series Extrinsic Regression (TSER) repository. And similarly, in the multivariate regression we evaluate with Chord and Chord + Curvature features on 19 datasets of the Soton subset of the (TSER) repository. Our metric for this task is MSE.

5.2. Methods

We can now build generalized GPF features as shown below where we concatenate features from lower degree of k just like path signatures. It is noted that each degree k results in $\binom{d}{k}$ features for a d channel time series. For all our experiments with multivariate time series, we compute up to $k = 4$ and append $k = d$.

$k = 1$

We consider each channel individually, computing signed and unsigned GPF requires four dimensions that form two 2-dimensional time series. The dimensions of the first time series are a dimension from the time series and time index. The dimensions of the second time series are a derivative time series of the channel, e.g. (Its Chord, Curvature, or Absolute Value) and time index.

$k = 2$

This is similar to when $k = 1$ except we dont need a derivative of the channel and rather consider another channel in the time series to compute signed and unsigned GPF which means we have one channel from the time series with a time index and another channel from the time series with a time index.

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 March 2026 doi:10.20944/preprints202603.1778.v1

$k = 3+$

We calculate both the signed and unsigned Lévy volume of the k channels.

5.3. Results

For each dataset row, scores are normalised as:

- $\tilde{s} = a/a^*$ for classification. where a^* is the best accuracy in the row
- $\tilde{s} = m^*/m$ for regression, where m^* is the best MSE in the row.

We then average over the D datasets: $\bar{s} = \frac{1}{D}\sum_{d=1}^{D} \tilde{s}_d$. The outermost polygon indicates better performance. Full results for both metrics are in Appendix Q-T with feature counts in Appendix M-P. We denote the baseline path signature as Sig, with standalone GPF variants GPF_S, GPF_U, and GPF_{SU}. When GPF is appended with signature features we denote as GPF + Sig.

5.3.1. Performance Over All Tasks

For univariate classification, in Figure 3, we observe that the path signature is the innermost polygon indicating it underperforms relative to our features and that both GPF_S and GPF_U perform either similarly to path signatures or better than it. While for multivariate classification as shown in Figure 3, most variants of GPF outperform path signatures except GPF_U indicating that signed areas are important and GPF_U is best used to complement feature sets. With regards to regression in general, from Figure 3, the path signature performs than GPF as it is the innermost polygon for both cases. We also observe that there is a significant distance between the GPF variants and the path signature polygon, this is further the case when GPF_{SU} is used alone or in tandem with path signatures. It is worth noting that in the case of the univariate tasks, path signatures have 8 features while GPF_S, GPF_U, and GPF_{SU} have 4, 4, and 8 features respectively.

To demonstrate specific examples where the improvements are substantial, Table 1 is presented for Chord + Curvature $GPF_{SU} + Sig$. As seen here, in some cases like ShapeletSim which is a binary classification dataset, path signatures essentially perform random guessing but when we add GPF_{SU}, we are now certain enough to have perfect accuracy. This indicates that GPF captures geometric cross channel interactions very well which complements path signatures. We can see a similar trend in the regression task where we can reduce MSE of the path signatures by adding our features for up to 14.2× reduction.

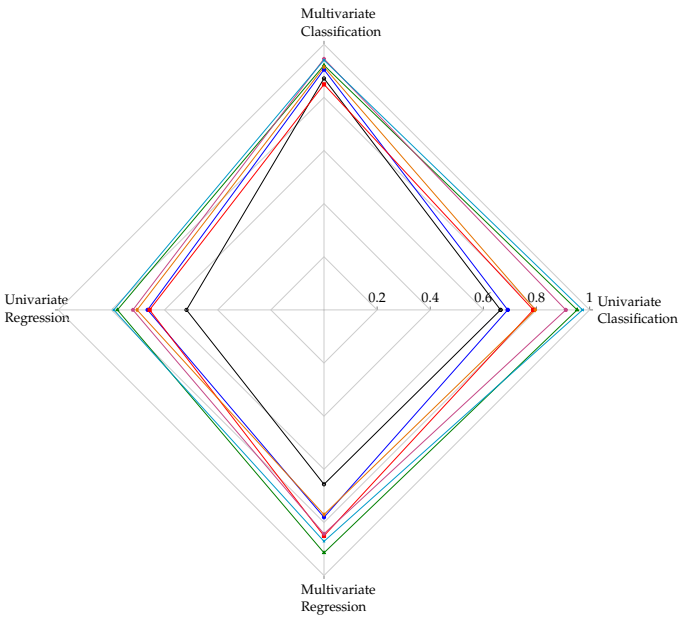

Figure 3. Normalised performance radar chart across feature sets. Method colour key: Sig; GPF_S; GPF_U; GPF_{SU}; GPF_S+Sig; GPF_U+Sig; GPF_{SU}+Sig.

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 March 2026 doi:10.20944/preprints202603.1778.v1

Table 1. Representative datasets where GPF$_{SU}$+Sig substantially outperforms signatures alone. Classification reports accuracy (↑); regression reports MSE (↓) with improvement expressed as a multiplicative reduction (×).

	Dataset	Sig	GPF$_{SU}$+Sig	Improv.
Classif.	ShapeletSim	0.55	**1.00**	+0.45
	InsectEPGRegularTrain	0.55	**0.98**	+0.43
	Mallat	0.18	**0.60**	+0.42
	MelbournePedestrian	0.34	**0.73**	+0.39
Regress.	DhakaHourlyAirQuality	4003.48	**282.78**	14.2×
	HouseholdPowerConsumption1	174093.58	**39484.04**	4.4×
	WindTurbinePower	744821.80	**254201.11**	2.9×
	IEEEPPG	247.17	**149.15**	1.7×

5.3.2. Comparison with Other Feature Sets

Here we make a pipeline to properly compare with well established feature sets, we have the Original Series, their Chords, Curvatures, and Absolute Values(Absolute value at each time point for each series) as well as the z-score normalized versions of all these channels. We use these channels on both *GPF* where $k = 2$, and degree two path signatures. Full details such as results and feature counts are in Appendix V. The other feature sets run on the raw time series as intended, but for fair comparison we feed the feature sets into random forest classifier with the same hyperparameters. We compare across feature sets using Friedman rank [32], and Pareto optimality [33]. The former measures the average position of a method across all datasets, and the latter finds which feature sets are not dominated by others with regards to both accuracy and feature count at the same time. It is also noted that some feature sets are univariate only where we have applied the feature extraction of those feature sets on a per channel basis.

Table 2 is to contextualise our *GPF* features against ten other established feature sets across the same fifteen UEA multivariate datasets which we have used in the previous experiment to compare path signatures with our features. RDST ranks first with a mean rank of 4.90 and 75.2% accuracy, but at the cost of 27,900 features on average. While GPF_{SU} ranks second, within a narrow margin of RDST while using roughly 3× fewer features, matching RDST on the Pareto frontier at 3× fewer features and with full geometric interpretability. MultiROCKET ranks at third but requires 7,392 random convolutions which have no geometric interpretability, essentially a black box feature set. Notably, all GPF variants append to path signatures rank in the top five, leading to outperforming all other feature sets, confirming that GPF features capture complementary information beyond what path signatures provide. The GPF_U underperforms notably without the signed area/volume information, suggesting that signed features are the primary driver of discriminative power within the *GPF* family. BOSS, TSFresh, and Interval occupy the bottom three positions, which is consistent with their limitation being specific to univariate time series. It is important to note, we can easily scale our results easily by increasing k for *GPF*, and the degree of path signatures to improve results.

Table 2. Friedman rank [32], mean accuracy, and mean feature count across UEA datasets (rank 1 = best). ⋆ denotes Pareto-optimal [33] methods (accuracy vs. feature count).

#	Method	Mean Rank ↓	Mean Acc.	Mean Feats
1	RDST⋆	4.90	75.2%	27,900
2	GPF$_{SU}$+Sig⋆	5.70	74.8%	8,609
3	MultiROCKET⋆	6.37	74.7%	7,392
4	GPF$_{SU}$	6.53	73.6%	4,268
5	GPF$_U$+Sig⋆	6.63	73.9%	6,502
6	MiniROCKET⋆	7.83	73.1%	924
7	ROCKET⋆	8.23	73.7%	2,000
8	GPF$_S$+Sig	8.77	73.6%	6,502
9	HYDRA	8.97	72.4%	4,984

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 March 2026 doi:10.20944/preprints202603.1778.v1

Table 2. *Cont.*

#	Method	Mean Rank ↓	Mean Acc.	Mean Feats
10	Catch22⋆	9.33	72.7%	143
11	Sig	9.50	72.9%	4,368
12	GPF_S	10.03	72.5%	2,134
13	Shapelet	10.47	72.4%	774
14	GPF_U	11.93	68.8%	2,134
15	TSFresh	12.10	59.3%	4,870
16	Interval⋆	12.83	68.1%	113
17	BOSS	12.87	68.4%	76,790

6. Conclusion

We propose a new set of features, grounded in path signatures, that shows promise for ML applications. By leveraging mathematical representations of sequential patterns, it excels in capturing temporal dependencies in data. Its adaptability and efficiency make it valuable for tasks like sequence classification and time-series analysis. Our algorithm offers a fresh approach to interpretable features through geometric grounding and addressing challenges associated with time series data.

Data Availability Statement: Data sharing is not applicable to this article as no new data were created or analyzed in this study.

Acknowledgments: The authors thank the Qatar National Library (QNL) for their support in providing resources and funding the APC, which was instrumental in completing this research.

Conflicts of Interest: On behalf of all authors, the corresponding author states that there is no Conflict of interest.

Appendix A. Impact of Definition 1

The set of bounded variation paths is exactly the set of functions whose first derivatives exist almost everywhere. Being of bounded variation is therefore not a particularly restrictive assumption. It contains, for example, all Lipschitz functions. In particular, if X_t is continuously differentiable, and $\dot{X}_t$ denotes its first derivative with respect to t e.g. $\dot{X}_t = \frac{dX_t}{dt}$, then the length of X_t can be defined,

$$\|X\|_v = \int_0^1 \|\dot{X}_t\| dt$$

The assumption of bounded variation allows us to define Riemann-Stieljes integrals along paths. More details is given in Lyons et al. (2007). Therefore we will assume that the integral of a continuous path $Y : [0,1] \to \mathbb{R}^d$ against a path of bounded variation $X : [0,1] \to \mathbb{R}^d$ is well-defined on any interval $[s,t] \subset [0,1]$, and denoted by,

$$\int_s^t Y_h dX_h = \begin{pmatrix} \int_s^t Y_h^1 dX_h^1 \\ \vdots \\ \int_s^t Y_h^d dX_h^d \end{pmatrix} \in \mathbb{R}^d$$

where $X_t = (X_t^1, \ldots, X_t^d)$, and $Y_t = (Y_t^1, \ldots, Y_t^d)$. When X_t is continuously differentiable, this integral is equal to the standard Riemann integral, that is,

$$\int_s^t Y_h dX_h = \int_s^t Y_h \dot{X}_h dh$$

It is noted for paths of arbitrary domain the above applies as the domain can be rescaled such as in a time series.

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 March 2026 doi:10.20944/preprints202603.1778.v1

Appendix B. Example of Signature for Continuous Time Series

Let $X_t = (X_t^1, X_t^2) = (t, t^2) \in \mathbb{R}^2$, a time series defined on $[0,4]$, and its derivative defined by $dX_t = (dX_t^1, dX_t^2) = (dt, 2tdt)$. Then the signature of order 2, $\Phi_2(X_t)_{[0,4]}$, is given by,

$$\left(1, \Phi^{(1)}, \Phi^{(2)}, \Phi^{(1,1)}, \Phi^{(1,2)}, \Phi^{(2,1)}, \Phi^{(2,2)}\right)$$

where the signature terms of order 1 are,

$$\Phi^{(1)} = \Phi^{(1)}(X_t)_{[0,4]} = \int_0^4 dX_t^1 = X_4^1 - X_0^1 = 4$$

$$\Phi^{(2)} = \Phi^{(2)}(X_t)_{[0,4]} = \int_0^4 dX_t^2 = X_4^2 - X_0^2 = 16$$

The first term of order 2 is,

$$\Phi^{(1,1)} = \Phi^{(1,1)}(X_t)_{[0,4]} = \int_0^4 \int_0^t dX_u^1 dX_t^1 = \frac{\left(X_4^1 - X_0^1\right)^2}{2!} = 8$$

Similarly,

$$\Phi^{(1,2)} = \Phi^{(1,2)}(X_t)_{[0,4]} = \int_0^4 \int_0^t dX_u^1 dX_t^2 = \int_0^4 2t_2^2 dt_2 = 128/3$$

$$\Phi^{(2,1)} = \Phi^{(2,1)}(X_t)_{[0,4]} = \int_0^4 \int_0^t dX_u^2 dX_t^1 = \int_0^4 t_2^2 dt_2 = 64/3$$

Finally,

$$\Phi^{(2,2)} = \Phi^{(2,2)}(X_t)_{[0,4]} = \int_0^4 \int_0^t dX_u^2 dX_t^2 = \frac{\left(X_4^2 - X_0^2\right)^2}{2!} = 128$$

Appendix C. Example of Signature for Discrete Time Series

In practice, most of the data are discrete, like time series. The computation of the signature of discrete data is achieved through the following three steps:

1. Transform the discrete data into a piecewise linear continuous curve through linear interpolation.
2. For each liner piece, compute its signature
3. Compute the signature of the entire curve by applying Chen's identity [34].

Let X_t be a piecewise linear curve and $t_0 < t_1 < \ldots < t_n$ the timestamps on which X_t is observed such that on each partition $[t_{(j-1)}, t_j]$, X_t is linear. Then $\Phi(X_t)_{[t_{j-1}, t_j]}$ is computed. Then, by applying Chen's identity we have:

$$\Phi(X_t)_{[t_0,t_n]} = \Phi(X_t)_{[t_0,t_1]} \otimes \Phi(X_t)_{[t_1,t_2]} \ldots \otimes \Phi(X_t)_{[t_{n-1},t_n]}$$

where $A \otimes B$ is the algebraic tensor product between A and B.

Consider $X_t = (X_t^1, X_t^2)$, a two-dimensional time series as in the following

$$X_t = \begin{cases} X_t^1 = (0,1,2) \\ X_t^2 = (0,2,5) \end{cases}$$

Then X_t can be partitioned into two piecewise linear partitions P_1 and P_2

$$P_1 = \begin{cases} (0,1) \\ (0,2) \end{cases} \quad P_2 = \begin{cases} (1,2) \\ (2,5) \end{cases}$$

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 March 2026 doi:10.20944/preprints202603.1778.v1

The signature of order 2 for each partition can be computed as follows. For P_1,

$$\Phi^{(1)}(P_1) = 1,\ \Phi^{(2)}(P_1) = \frac{2}{1!} = 2,$$

$$\Phi^{(1,1)}(P_1) = \frac{1}{2!} = 0.5,\ \Phi^{(2,2)}(P_1) = \frac{4}{2!} = 2,$$

$$\Phi^{(1,2)}(P_1) = \Phi^{(2,1)}(P_1) = \frac{2}{2!} = 1$$

For P_2,

$$\Phi^{(1)}(P_2) = 1,\ \Phi^{(2)}(P_2) = \frac{3}{1!} = 3,$$

$$\Phi^{(1,1)}(P_2) = \frac{1}{2!} = 0.5,\ \Phi^{(2,2)}(P_2) = \frac{9}{2!} = 4.5,$$

$$\Phi^{(1,2)}(P_2) = \Phi^{(2,1)}(P_1) = \frac{3}{2!} = 1.5$$

Then,

$$\Phi_2(X_t) = \Phi_2(P_1) \otimes \Phi_2(P_2) = (1, 2, 5, 2, 5.5, 4.5, 12.5)$$

Which arises from the following by evaluating the tensor products from Chen's identity for the order zero term by notation,

$$\Phi^{\emptyset}(X_t) = \Phi^{\emptyset}(P_1) \cdot \Phi^{\emptyset}(P_2)$$

For the order 1 terms,

$$\Phi^{(i)}(X_t) = \Phi^{(i)}(P_1) + \Phi^{(i)}(P_2)$$

And for the order 2 terms,

$$\Phi^{(i,j)}(X_t) = \Phi^{(i,j)}(P_1) + \Phi^{(i)}(P_1)\Phi^{(j)}(P_2) + \Phi^{(i,j)}(P_2)$$

Appendix D. Signature for 1-Dimensional Time Series

Proof. This is indeed straightforward. The signature term of order 1,

$$\Phi^{(1)}(X_t) = \int_a^b dX_t^1 = X_b^1 - X_a^t$$

The signature term of order 2 is,

$$\Phi^{(1,1)}(X_t) = \int_a^b \int_a^{t_2} dX_{t_1}^1 dX_{t_2}^1 = \frac{(X_b^1 - X_a^1)^2}{2!}$$

$$\Phi^{(1,1,1)}(X_t) = \int_a^b \int_a^{t_3} \int_a^{t_2} dX_{t_1}^1 dX_{t_2}^1 dX_{t_3}^1$$
$$= \frac{(X_b^1 - X_a^1)^3}{3!}$$

And the signature term of order k is,

$$\Phi^{(1,..,1)}(X_t) = \frac{(X_b^1 - X_a^1)^k}{k!}$$

□

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 March 2026 doi:10.20944/preprints202603.1778.v1

Appendix E. Some Useful Facts About The Generalized Cross Product

The facts mentioned in this section come from [30] and will be useful for calculating areas and volumes between curves in higher dimensions. Sketchs of proofs are provided where possible. To start, we define the generalized cross product which will be useful for volumes,

Definition A1. *For a vector space* $\mathbf{V}^N$ *with basis* $\{\mathbf{e}_1, \ldots, \mathbf{e}_N\}$*, and vectors* $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{N-1} \in \mathbf{V}^N$*, the generalized cross product is defined by,*

$$cross(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{N-1}) = \begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \cdots & \mathbf{e}_N \\ \mathbf{v}_{1,1} & \mathbf{v}_{1,2} & \cdots & \mathbf{v}_{1,N} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{v}_{N-1,1} & \mathbf{v}_{N-1,2} & \cdots & \mathbf{v}_{N-1,N} \end{vmatrix}$$

The determinant approach allows us to compute a vector orthogonal to the input vectors. We now show the proposition which allows us to compute the area between high dimensional curves,

Proposition A1. *For* $\mathbf{x}, \mathbf{y} \in \mathbf{V}^N$*, the sine of the angle* θ *between them can be expressed using their magnitudes and dot products.*

Proof. We use the Pythagorean identity below to derive the magnitude of the cross product,

$$|\mathbf{x} \times \mathbf{y}| = \|\mathbf{x}\|\|\mathbf{y}\| \sin(\theta) = \sqrt{\|\mathbf{x}\|^2\|\mathbf{y}\|^2(1 - \cos^2(\theta))}$$

We substitute $\cos(\theta) = \frac{\mathbf{x} \cdot \mathbf{y}}{\|\mathbf{x}\|\|\mathbf{y}\|}$ and we are done. □

It is thus observed that $|\mathbf{x} \times \mathbf{y}| = \sqrt{\|\mathbf{x}\|^2\|\mathbf{y}\|^2 - (\mathbf{x} \cdot \mathbf{y})^2}$ as required, which is precisely the area of the parallelogram between both vectors. We now demonstrate the relationship between the generalized cross product and the volume of an N-dimensional parallelotope which will allow us to compute the volume between high dimensional curves,

Theorem A1. *The magnitude of the generalized cross product is equal to the volume of the parallelotope formed by the vectors* $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{N-1} \in \mathbf{V}^N$*.*

To make notation easier to read, we use the convention of column vectors for the generalized cross product, but as the determinant is symmetric it is still equivalent as to using them as row vectors.

Proof. The volume of a parallelotope defined by vertex vectors can be derived from the Jacobian matrix which means,

$$\det([\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_N]) = \text{volume}(\mathbf{v}_1, \ldots, \mathbf{v}_N).$$

Now consider a rotation matrix $\mathbf{A}$. It is noted that its determinant is 1 thus we can absorb it inside the determinant on the right hand side. For this case, $\mathbf{v}_N = \mathbf{e}_1 + \mathbf{e}_2 + \ldots + \mathbf{e}_N$,

$$\det([\mathbf{v}_1, \ldots, \mathbf{v}_{N-1}, \mathbf{v}_N]) = \det(\mathbf{A}[\mathbf{v}_1, \ldots, \mathbf{v}_{N-1}, \mathbf{v}_N])$$

We use $\mathbf{A}$ to zero out $\mathbf{e}_N$ for $\mathbf{v}_1, \ldots, \mathbf{v}_{N-1}$,

$$= \det\left(\left[\begin{pmatrix} \mathbf{v}'_{11} \\ \vdots \\ 0 \end{pmatrix}, \ldots, \begin{pmatrix} \mathbf{v}'_{N-1,1} \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ \vdots \\ 1 \end{pmatrix}\right]\right) = \begin{vmatrix} \mathbf{v}'_{1,1} & \cdots & \mathbf{v}'_{N-1,1} \\ \vdots & \vdots & \vdots \\ \mathbf{v}'_{1,N-1} & \cdots & \mathbf{v}'_{N-1,N-1} \end{vmatrix} = \text{volume}(\mathbf{v}_1, \ldots, \mathbf{v}_{N-1})$$

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 March 2026 doi:10.20944/preprints202603.1778.v1

Thus we can generalize to the case for an arbitrary orthogonal vector $\mathbf{v}_N = \text{cross}(\mathbf{v}_1, \ldots, \mathbf{v}_{N-1})$,

$$\text{volume}(\mathbf{v}_1, \ldots, \mathbf{v}_{N-1}) = \det\left(\left[\mathbf{v}_1, \ldots, \mathbf{v}_{N-1}, \frac{\text{cross}(\mathbf{v}_1, \ldots, \mathbf{v}_{N-1})}{\|\text{cross}(\mathbf{v}_1, \ldots, \mathbf{v}_{N-1})\|}\right]\right)$$

It is observed scaling by the magnitude is necessary as we see expansion on cofactors. □

We now have the tools to build our Lévy based features in the subsequent sections.

Appendix F. Proof for The Lévy Area Feature

Proof. Let X_t^1 and X_t^2 be two 1-dimensional paths which together form the 2-dimensional path $X_t = (X_t^1, X_t^2)$ where for the sake of example, it is constrained to the domain $[0, t]$, the orthogonal signed distance from the point,

$$M = \begin{pmatrix} X_t^1 \\ X_t^2 \end{pmatrix}$$

To the line $Y = X$ is given by:

$$d = \frac{X_t^1 - X_t^2}{\sqrt{2}}$$

The distance arises from the observation that the line $Y = X$ rearranged can be interpreted as the dot product $< (X, Y), (1, -1) >= 0$. Equivalently, $< (X, Y), (1/\sqrt{2}, -1/\sqrt{2}) >= 0$. Thus points across this line have distance 0 and d otherwise as $(1/\sqrt{2}, -1/\sqrt{2})$ is perpendicular to the line. Therefore

$$dA = dl \cdot d = d \cdot (< d\vec{V}, \vec{u} >)$$

Where $\vec{u} = (1\sqrt{2}, 1/\sqrt{2})$ is the unit direction of the line $Y = X$ and $d\vec{V} = M_{t+dt} - M_t$. So,

$$d\vec{V} = \begin{pmatrix} X_{t+dt}^2 - X_t^2 \\ X_{t+dt}^1 - X_t^1 \end{pmatrix} = \begin{pmatrix} \dot{X}_t^2 \\ \dot{X}_t^1 \end{pmatrix} dt$$

And taking the dot product of the unit vector along the line with the above,

$$dA = \frac{X_t^1 - X_t^2}{\sqrt{2}} \cdot \frac{(\dot{X}_t^2 + \dot{X}_t^1)dt}{\sqrt{2}}$$

The signed area can be reformulated as

$$dA = \frac{1}{2}\Big((X_t^1 - X_t^2)(\dot{X}_t^2 + \dot{X}_t^1)\Big)dt$$

Therefore, the total area enclosed between the line and the curve is calculated as,

$$A = \int_0^{t_1} \left(\frac{1}{2}\left(\dot{X}_t^1 X_t^1 - X_t^2 \dot{X}_t^2\right) + \frac{1}{2}(X_t^1 \dot{X}_t^2 - X_t^2 \dot{X}_t^1)\right)dt$$

Or equivalently,

$$A = \int_0^{t_1} \frac{1}{2}\left(X_t^1 \dot{X}_t^2 - X_t^2 \dot{X}_t^1\right)dt + \int_0^{t_1} \frac{1}{2} X_t^1 \dot{X}_t^1 dt - \int_0^{t_1} \frac{1}{2} X_t^2 \dot{X}_t^2 dt$$

Using the signature definition and $\dot{X}_t^i dt = dX_t^i$, we have the below without the scaling factor $\frac{1}{2}$,

$$\int_0^{t_1} \dot{X}_t^1 X_t^1 dt = \frac{(X_{t_1}^1 - X_0^1)^2}{2} = \Phi^{(1,1)}(X_t)$$

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 March 2026 doi:10.20944/preprints202603.1778.v1

And,

$$\int_0^{t_1} \dot{X}_t^2 X_t^2 dt = \frac{(X_{t_1}^2 - X_0^2)^2}{2} = \Phi^{(2,2)}(X_t)$$

And lastly the Lévy area which is the focus of the proof,

$$A = \frac{1}{2}\int_0^{t_1} \left(\dot{X}_t^2 X_t^1 - X_t^2 \dot{X}_t^1\right) dt = \frac{1}{2}\int_0^{t_1} X_t^1 \dot{X}_t^2 dt - \frac{1}{2}\int_0^{t_1} X_t^2 \dot{X}_t^1 dt$$

$$= \frac{\Phi^{(1,2)}(X_t) - \Phi^{(2,1)}(X_t)}{2}$$

□

From the above, the signed and unsigned Lévy based signature features can be defined as follows,

$$A_{signed} = \frac{1}{2}\int_0^{t_1} \left(\dot{X}_t^2 X_t^1 - X_t^2 \dot{X}_t^1\right) dt,$$

$$A_{unsigned} = \frac{1}{2}\int_0^{t_1} \left|\dot{X}_t^2 X_t^1 - X_t^2 \dot{X}_t^1\right| dt.$$

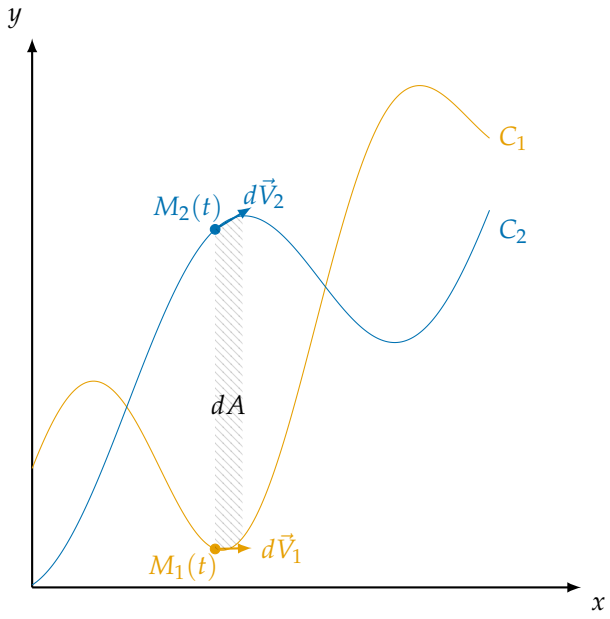

Figure A1. Area Element Between Two 2-Dimensional Curves.

Appendix G. Proof for The Lévy Area Between Two 2-Dimensional Curves

Proof. Let's assume we have two paths (Figure 1) which arise from a 4-dimensional path $X_t : [a, b] \to \mathbb{R}^4$:

$$M_1(t) = \begin{pmatrix} X_t^1 \\ X_t^2 \end{pmatrix} \quad M_2(t) = \begin{pmatrix} X_t^3 \\ X_t^4 \end{pmatrix}$$

Then, it is observed a vector between the two paths can be formed and along with a cross product the differential area can be computed,

$$dA = \|d\vec{V}_1 \times \overrightarrow{M_1(t)M_2(t)}\| = \|\overrightarrow{M_1(t)M_1(t+dt)} \times \overrightarrow{M_1(t)M_2(t)}\|$$

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 March 2026 doi:10.20944/preprints202603.1778.v1

Which is equivalent to the below since the cross product can be defined in terms of the determinant,

$$dA = \begin{vmatrix} \dot{X}_t^1 dt & \dot{X}_t^2 dt \\ X_t^1 - X_t^3 & X_t^2 - X_t^4 \end{vmatrix}$$

So,

$$dA = (X_t^2 - X_t^4)\dot{X}_t^1 dt - (X_t^1 - X_t^3)\dot{X}_t^2 dt$$

Therefore,

$$A = \int_a^b \left(X_t^2 \dot{X}_t^1 + X_t^2 X_t^3 - X_t^4 \dot{X}_t^1 - \dot{X}_t^2 X_t^1 \right) dt$$

And,

$$A = \int_a^b \left(X_t^2 \dot{X}_t^1 - X_t^1 \dot{X}_t^2 \right) dt + \int_a^b \left(X_t^3 \dot{X}_t^2 - X_t^4 \dot{X}_t^1 \right) dt$$

Thus, from the definition of the signature,

$$A = \Phi^{(1,2);(3,4)}(X_t) = \Phi^{(2,1)}(X_t) + \Phi^{(3,2)}(X_t) - \Phi^{(1,2)}(X_t) - \Phi^{(4,1)}(X_t)$$

□

$\Phi^{(1,2);(3,4)}$ is *GPF* feature which is a generalization of signature feature (Lévy area) from a single 2 dimensional curve to two 2 dimensional curves.

From this, the signed and unsigned *GPF* features can be defined as follows:

$$A_{\text{signed}} = \text{GPF}_S = \int_a^b \left(X_t^2 \dot{X}_t^1 - \dot{X}_t^2 X_t^1 \right) dt + \int_a^b \left(\dot{X}_t^2 X_t^3 - X_t^4 \dot{X}_t^1 \right) dt$$

$$= \Phi^{(2,1)}(X_t) + \Phi^{(3,2)}(X_t) - \Phi^{(1,2)}(X_t) - \Phi^{(4,1)}(X_t)$$

And,

$$A_{\text{unsigned}} = \text{GPF}_U = \int_a^b \left| X_t^2 \dot{X}_t^1 - \dot{X}_t^2 X_t^1 + \dot{X}_t^2 X_t^3 - X_t^4 \dot{X}_t^1 \right| dt$$

Where the unsigned Lévy area has no closed form in terms of signature terms.

Appendix H. Proof for The Lévy Area Between Two 3-Dimensional Curves

Proof. Assume we have two 3-dimensional paths from the 6-dimensional path $X_t : [a, b] \to \mathbb{R}^6$:

$$M_1(t) = \begin{pmatrix} X_t^1 \\ X_t^2 \\ X_t^3 \end{pmatrix} \quad M_2(t) = \begin{pmatrix} X_t^4 \\ X_t^5 \\ X_t^6 \end{pmatrix}$$

Then just as in the previous section, the signed area is defined as the following:

$$dA = ||\overrightarrow{M_1(t)M_1(t+dt)} \times \overrightarrow{M_1(t)M_2(t)}|| = \left|\left| \begin{pmatrix} \dot{X}_t^1 \\ \dot{X}_t^2 \\ \dot{X}_t^3 \end{pmatrix} \times \begin{pmatrix} X_t^4 - X_t^1 \\ X_t^5 - X_t^2 \\ X_t^6 - X_t^3 \end{pmatrix} \right|\right| dt$$

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 March 2026 doi:10.20944/preprints202603.1778.v1

The differential element dt can be taken out of the determinant when computing the cross product due to the scaling property of determinants,

$$dA = \left\| \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ \dot{X}_t^1 & \dot{X}_t^2 & \dot{X}_t^3 \\ X_t^4 - X_t^1 & X_t^5 - X_t^2 & X_t^6 - X_t^3 \end{vmatrix} \right\| dt = \Bigg\| \vec{i} \underbrace{\left[\dot{X}_t^2 (X_t^6 - X_t^3) - \dot{X}_t^3 (X_t^5 - X_t^2) \right]}_{\Lambda}$$

$$- \vec{j} \underbrace{\left[\dot{X}_t^1 (X_t^6 - X_t^3) - \dot{X}_t^3 (X_t^4 - X_t^1) \right]}_{\Psi} + \vec{k} \underbrace{\left[\dot{X}_t^1 (X_t^5 - X_t^2) - \dot{X}_t^2 (X_t^4 - X_t^1) \right]}_{\Omega} \Bigg\| dt$$

□

Therefore, the signature feature is defined succintly below,

$$A = \int_{t_1}^{t_2} dA = \int_{t_1}^{t_2} \sqrt{\Lambda^2 + \Psi^2 + \Omega^2}\, dt = \Phi^{(1,2,3);(4,5,6)}$$

Appendix I. Proof for The Lévy Area Between Two 4-Dimensional Or Two Higher Dimensional Curves

Proof. Consider we have two 4-dimensional paths from the 8-dimensional path $X_t : [a, b] \to \mathbb{R}^8 = (X_t^1, X_t^2, ..., X_t^7, X_t^8)$. Assume we split the paths uniquely into two 4-dimensional paths:

$$M_1(t) = \begin{pmatrix} X_t^1 \\ X_t^2 \\ X_t^3 \\ X_t^4 \end{pmatrix} \quad M_2(t) = \begin{pmatrix} X_t^5 \\ X_t^6 \\ X_t^7 \\ X_t^8 \end{pmatrix}$$

Then similar to the previous section, the signed area is defined using the notion of the parallelogram spanned between two vectors,

$$dA = \left\| \overrightarrow{M_1(t)M_1(t+dt)} \right\| \cdot \left\| \overrightarrow{M_1(t)M_2(t)} \right\| \cdot sin\left(\overrightarrow{M_1(t)M_1(t+dt)}, \overrightarrow{M_1(t)M_2(t)} \right)$$

Proposition A1 of Section 5 is then used to compute the signed area of the parallelogram,

$$= \sqrt{\left\| \overrightarrow{M_1(t)M_1(t+dt)} \right\|^2 \cdot \left\| \overrightarrow{M_1(t)M_2(t)} \right\|^2 - \left\langle \overrightarrow{M_1(t)M_1(t+dt)}, \overrightarrow{M_1(t)M_2(t)} \right\rangle^2}$$

Thus, the signature feature is defined as,

$$A = \int_{t_1}^{t_2} dA = \Phi^{(1,2,3,4);(5,6,7,8)}$$

□

From the above, higher dimensional paths can be dealt with similarly,

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 March 2026 doi:10.20944/preprints202603.1778.v1

Proof. Consider we have two p-dimensional points from the $2p$-dimensional path $X_t : [a,b] \to \mathbb{R}^{2p} = (X_t^1, X_t^2, ..., X_t^{2p-1}, X_t^{2p})$. Assume we split the paths uniquely into two p-dimensional points:

$$M_1(t) = \begin{pmatrix} X_t^1 \\ X_t^2 \\ \vdots \\ X_t^p \end{pmatrix} \quad M_2(t) = \begin{pmatrix} X_t^{p+1} \\ X_t^{p+2} \\ \vdots \\ X_t^{2p} \end{pmatrix}$$

Then the same signature feature can be used as it is invariant to dimension of the points,

$$A = \int_{t_1}^{t_2} dA = \int_{t_1}^{t_2} \sqrt{\left\|\overrightarrow{M_1(t)M_1(t+dt)}\right\|^2 \cdot \left\|\overrightarrow{M_1(t)M_2(t)}\right\|^2 - \left\langle \overrightarrow{M_1(t)M_1(t+dt)}, \overrightarrow{M_1(t)M_2(t)} \right\rangle^2}$$

$$= \Phi^{(1,2,...,p);(p+1,p+2,...,2p)}$$

□

Appendix J. Proof for The Lévy Volume Between Three 3-Dimensinoal Curves Or p Curves That Are p-Dimensional

The Lévy volume for the case of three 3D curves will serve as the prototype for the higher dimensional analogues and thus will be proved first.

Proof. Consider we have three 3-dimensional paths from the 9-dimensional path $X_t : [a,b] \to \mathbb{R}^9 = (X_t^1, X_t^2, ..., X_t^8, X_t^9)$. Assume we split the paths uniquely into three 3-dimensional paths:

$$M_1(t) = \begin{pmatrix} X_t^1 \\ X_t^2 \\ X_t^3 \end{pmatrix} \quad M_2(t) = \begin{pmatrix} X_t^4 \\ X_t^5 \\ X_t^6 \end{pmatrix} \quad M_3(t) = \begin{pmatrix} X_t^7 \\ X_t^8 \\ X_t^9 \end{pmatrix}$$

In this case, it is appropriate that instead of measuring the area between pairs of curves, the volume between multiple curves have to be considered. To measure volume, the integration in this case uses the area formed between the curves at time t similarly as in the previous section. It is observed that the area is a shape of a triangle which can be computed using the cross product as follows,

$$dA = \left\| \frac{\overrightarrow{M_1(t)M_2(t)} \times \overrightarrow{M_1(t)M_3(t)}}{2} \right\| dt = \left\| \frac{1}{2} \left(\begin{pmatrix} X^4 - X_t^1 \\ X^5 - X_t^2 \\ X^6 - X_t^3 \end{pmatrix} \times \begin{pmatrix} X_t^7 - X_t^1 \\ X_t^8 - X_t^2 \\ X_t^9 - X_t^3 \end{pmatrix} \right) \right\| dt$$

To visualize what is being done for intuition, the below Figure A2 is presented,

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 March 2026 doi:10.20944/preprints202603.1778.v1

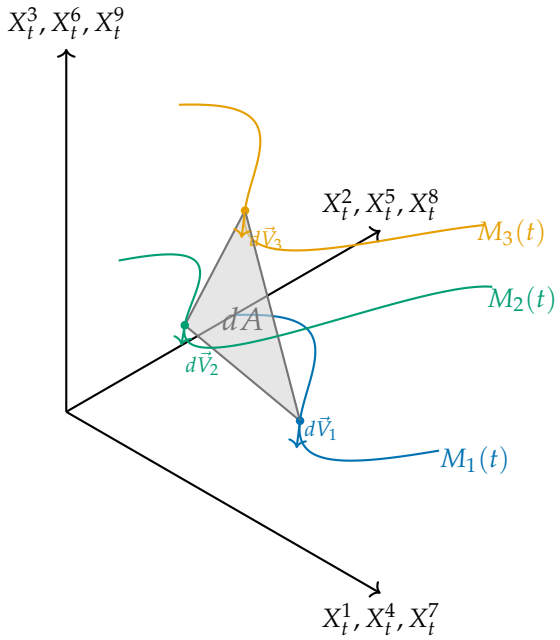

Figure A2. Volume Element Between Three 3-Dimensional Curves.

To make notation more concise and so that it is easier generalize the notion of volume with higher dimensional curves, the scaling factor $\frac{1}{2}$ is taken out.

$$dA = \left|\left|\overrightarrow{M_1(t)M_2(t)} \times \overrightarrow{M_1(t)M_3(t)}\right|\right| dt = \left|\left|\left(\begin{pmatrix} X^4 - X_t^1 \\ X^5 - X_t^2 \\ X^6 - X_t^3 \end{pmatrix} \times \begin{pmatrix} X_t^7 - X_t^1 \\ X_t^8 - X_t^2 \\ X_t^9 - X_t^3 \end{pmatrix}\right)\right|\right| dt$$

Computing the cross product requires the determinant definition to be used,

$$dA = \left|\left|\begin{vmatrix} \mathbf{e}_i & \mathbf{e}_j & \mathbf{e}_k \\ X^4 - X_t^1 & X^5 - X_t^2 & X^6 - X_t^3 \\ X_t^7 - X_t^1 & X_t^8 - X_t^2 & X_t^9 - X_t^3 \end{vmatrix}\right|\right| dt$$

Now, the signed signature feature is defined as usual,

$$A_{signed} = \int_{t_1}^{t_2} dA = \Phi^{(1,2,3);(4,5,6);(7,8,9)}$$

With the unsigned signature feature as,

$$A_{unsigned} = \int_{t_1}^{t_2} |dA|$$

□

Now this approach can be repeated for p curves that are p-dimensional,

Proof. Consider there are p p-dimensional curves from the p^2-dimensional path $X_t : [a, b] \to \mathbb{R}^8 = (X_t^1, X_t^2, ..., X_t^{p^2-1}, X_t^{p^2})$. Assume the paths are split uniquely into p p-dimensional paths,

$$M_1(t) = \begin{pmatrix} X_t^1 \\ X_t^2 \\ \vdots \\ X_t^p \end{pmatrix} \quad M_2(t) = \begin{pmatrix} X_t^{p+1} \\ X_t^{p+2} \\ \vdots \\ X_t^{2p} \end{pmatrix}, ..., \quad M_p(t) = \begin{pmatrix} X_t^{p^2-p+1} \\ X_t^{p^2-p+2} \\ \vdots \\ X_t^{p^2} \end{pmatrix}$$

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 March 2026 doi:10.20944/preprints202603.1778.v1

The notion of cross product can be used again but in its general form as in Definition A1 of Section 5. Therefore, the volume element can be written as,

$$dA = \left|\left|cross\left(\overrightarrow{M_1(t)M_2(t)}, \overrightarrow{M_1(t)M_3(t)}, ..., \overrightarrow{M_1(t)M_p(t)}\right)\right|\right|dt = \left|\left|cross\left(\begin{pmatrix} X^{p+1} - X_t^1 \\ X^{p+2} - X_t^2 \\ \vdots \\ X^{2p} - X_t^p \end{pmatrix}, ..., \begin{pmatrix} X_t^{p^2-p+1} - X_t^1 \\ X_t^{p^2-p+2} - X_t^2 \\ \vdots \\ X_t^{p^2} - X_t^p \end{pmatrix}\right)\right|\right|dt$$

Where the cross product is computed with a determinant as per Definition 1,

$$dA = \left|\left|\left|\begin{matrix} \mathbf{e}_1 & \mathbf{e}_2 & ... & \mathbf{e}_p \\ X^{p+1} - X_t^1 & X^{p+2} - X_t^2 & ... & X^{2p} - X^p \\ \vdots & \vdots & \dots & \vdots \\ X_t^{p^2-p+1} - X_t^1 & X_t^{p^2-p+2} - X_t^2 & ... & X_t^{p^2} - X_t^p \end{matrix}\right|\right|\right|dt$$

Now, the signed signature feature is defined,

$$A_{signed} = \int_{t_1}^{t_2} dA = \Phi^{(1,2,..,p);(p+1,p+2,...,2p);...;(p^2-p+1,p^2-p+2,...,p^2)}$$

With the unsigned signature feature as,

$$A_{unsigned} = \int_{t_1}^{t_2} |dA|$$

□

It is important to observe that for an n-dimensional path or time series, the above features are possible to calculate as long as we find a square number p such that $p \leq n$ where we use p of the original time series to compute the above features. This means there are P_p^n possible signature features without repetition. If we allow repetition we have n^p possible signature features. But this is not the limit of the signature features defined here as the point is that the area element is reliant on how the points are encoded. For example, below is a possible encoding for the points,

$$dA = \left|\left|cross\left(\overrightarrow{M_1(t)(2M_2(t))}, \overrightarrow{M_1(t)(2M_3(t))}, ..., \overrightarrow{M_1(t)(2M_p(t))}\right)\right|\right| = \left|\left|\left|\begin{matrix} \mathbf{e}_1 & \mathbf{e}_2 & ... & \mathbf{e}_p \\ X^{p+1} - 2X_t^1 & X^{p+2} - 2X_t^2 & ... & X^{2p} - 2X^p \\ \vdots & \vdots & \dots & \vdots \\ X_t^{p^2-p+1} - 2X_t^1 & X_t^{p^2-p+2} - 2X_t^2 & ... & X_t^{p^2} - 2X_t^p \end{matrix}\right|\right|\right|$$

Therefore by extension, linear combinations are possible or other derivative time series can be used. The implication is that we can say an infinite number of signature features are possible.

Appendix K. Per-Dataset Comparison with Statistical Tests

In this section, the comparison between the baseline signature method is compared with the signed and unsigned *GPF* features as standalone features or to complement baseline signatures. The statistical validation is done through the Wilcoxon signed rank tests [32]. The comparison and statistics are calculated on a task basis where we have four tasks and compare both methods. The tasks compared are in Univariate Classification (Tables 1 And 2), Multivariate Classification (Tables 3 And 4), Univariate Regression (Tables 5 And 6), and Multivariate Regression (Tables 7 And 8). For both

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 March 2026 doi:10.20944/preprints202603.1778.v1

Classification and Regression, when **p+** < 0.05, we say our method is Significantly Better, and when **p-** < 0.05 it is significantly Worse, otherwise there is No Significant Difference.

Table A1. Wilcoxon Signed-Rank Test for Univariate Classification (Curvature) (128 Datasets, Metric: Accuracy, Higher is Better).

Method	W+	p+	W-	p-	Conclusion
GPF_S	1177.00	1.000000	6951.00	0.000000	Significantly Worse
GPF_U	2082.00	0.999999	6046.00	0.000001	Significantly Worse
GPF_{SU}	4597.00	0.099624	3531.00	0.900376	No Significant Difference
GPF_S + Sig	7217.00	0.000000	784.00	1.000000	Significantly Better
GPF_U + Sig	7770.00	0.000000	231.00	1.000000	Significantly Better
GPF_{SU} + Sig	7887.00	0.000000	114.00	1.000000	Significantly Better

Table A2. Wilcoxon Signed-Rank Test for Univariate Classification (Chord + Curvature) (128 Datasets, Metric: Accuracy, Higher is Better).

Method	W+	p+	W-	p-	Conclusion
GPF_S	5279.00	0.001744	2849.00	0.998256	Significantly Better
GPF_U	6821.00	0.000000	1435.00	1.000000	Significantly Better
GPF_{SU}	8162.00	0.000000	94.00	1.000000	Significantly Better
GPF_S + Sig	7552.00	0.000000	198.00	1.000000	Significantly Better
GPF_U + Sig	8240.00	0.000000	16.00	1.000000	Significantly Better
GPF_{SU} + Sig	8255.00	0.000000	1.00	1.000000	Significantly Better

Table A3. Wilcoxon Signed-Rank Test for Multivariate Classification (Dyadic Curvature) (15 Datasets, Metric: Accuracy, Higher is Better).

Method	W+	p+	W-	p-	Conclusion
GPF_S	67.00	0.181343	38.00	0.818657	No Significant Difference
GPF_U	30.00	0.958374	90.00	0.047302	Significantly Worse
GPF_{SU}	58.00	0.068049	20.00	0.931951	No Significant Difference
GPF_S + Sig	104.00	0.000613	1.00	0.999387	Significantly Better
GPF_U + Sig	88.00	0.001488	3.00	0.998512	Significantly Better
GPF_{SU} + Sig	78.00	0.001109	0.00	0.998891	Significantly Better

Table A4. Wilcoxon Signed-Rank Test for Multivariate Classification (Dyadic Chord + Curvature) (15 Datasets, Metric: Accuracy, Higher is Better).

Method	W+	p+	W-	p-	Conclusion
GPF_S	86.00	0.075714	34.00	0.932312	No Significant Difference
GPF_U	35.00	0.924286	85.00	0.084412	No Significant Difference
GPF_{SU}	73.00	0.027312	18.00	0.972688	Significantly Better
GPF_S + Sig	83.00	0.004387	8.00	0.995613	Significantly Better
GPF_U + Sig	104.00	0.000613	1.00	0.999387	Significantly Better
GPF_{SU} + Sig	83.00	0.004387	8.00	0.995613	Significantly Better

Table A5. Wilcoxon Signed-Rank Test for Univariate Regression (Curvature) (26 Datasets, Metric: MSE, Lower is Better).

Method	W+	p+	W-	p-	Conclusion
GPF_S	262.00	0.013598	89.00	0.987325	Significantly Better
GPF_U	207.00	0.218723	144.00	0.788649	No Significant Difference
GPF_{SU}	310.00	0.000141	41.00	0.999876	Significantly Better
GPF_S + Sig	345.00	0.000000	6.00	1.000000	Significantly Better
GPF_U + Sig	320.00	0.000035	31.00	0.999970	Significantly Better
GPF_{SU} + Sig	323.00	0.000022	28.00	0.999981	Significantly Better

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 March 2026 doi:10.20944/preprints202603.1778.v1

Table A6. Wilcoxon Signed-Rank Test for Univariate Regression (Chord + Curvature) (26 Datasets, Metric: MSE, Lower is Better).

Method	W+	p+	W-	p-	Conclusion
GPF_S	314.00	0.000083	37.00	0.999928	Significantly Better
GPF_U	290.00	0.001330	61.00	0.998795	Significantly Better
GPF_{SU}	295.00	0.000802	56.00	0.999278	Significantly Better
GPF_S + Sig	324.00	0.000019	27.00	0.999984	Significantly Better
GPF_U + Sig	332.00	0.000005	19.00	0.999996	Significantly Better
GPF_{SU} + Sig	297.00	0.000649	54.00	0.999417	Significantly Better

Table A7. Wilcoxon Signed-Rank Test for Multivariate Regression (Curvature) (19 Datasets, Metric: MSE, Lower is Better).

Method	W+	p+	W-	p-	Conclusion
GPF_S	145.00	0.022280	45.00	0.979935	significantly better
GPF_U	157.00	0.005413	33.00	0.995274	significantly better
GPF_{SU}	161.00	0.003090	29.00	0.997335	significantly better
GPF_S + Sig	168.00	0.001011	22.00	0.999153	significantly better
GPF_U + Sig	188.00	0.000006	2.00	0.999996	significantly better
GPF_{SU} + Sig	188.00	0.000006	2.00	0.999996	significantly better

Table A8. Wilcoxon Signed-Rank Test for Multivariate Regression (Chord + Curvature) (19 Datasets, Metric: MSE, Lower is Better).

Method	W+	p+	W-	p-	Conclusion
GPF_S	181.00	0.000063	9.00	0.999952	Significantly Better
GPF_U	174.00	0.000322	16.00	0.999739	Significantly Better
GPF_{SU}	187.00	0.000010	3.00	0.999994	Significantly Better
GPF_S + Sig	168.00	0.001011	22.00	0.999153	Significantly Better
GPF_U + Sig	188.00	0.000006	2.00	0.999996	Significantly Better
GPF_{SU} + Sig	187.00	0.000010	3.00	0.999994	Significantly Better

Appendix L. Dataset Details

Here we include the details of dataset such as the number of samples, the number of channels, as well as the length of the channels. For all cases, we have equal length channels. When we have classification, we also give the number of classes for the specific dataset.

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 March 2026 doi:10.20944/preprints202603.1778.v1

Table A9. Univariate Classification Datasets.

Dataset	Samples	Channels	Length	Classes
ACSF1	200	1	1460	10
Adiac	781	1	176	37
AllGestureWiimoteX	1000	1	500	10
AllGestureWiimoteY	1000	1	500	10
AllGestureWiimoteZ	1000	1	500	10
ArrowHead	211	1	251	3
BME	180	1	128	3
Beef	60	1	470	5
BeetleFly	40	1	512	2
BirdChicken	40	1	512	2
CBF	930	1	128	3
Car	120	1	577	4
Chinatown	363	1	24	2
ChlorineConcentration	4307	1	166	3
CinCECGTorso	1420	1	1639	4
Coffee	56	1	286	2
Computers	500	1	720	2
CricketX	780	1	300	12
CricketY	780	1	300	12
CricketZ	780	1	300	12
Crop	24000	1	46	24
DiatomSizeReduction	322	1	345	4
DistalPhalanxOutlineAgeGroup	539	1	80	3
DistalPhalanxOutlineCorrect	876	1	80	2
DistalPhalanxTW	539	1	80	6
DodgerLoopDay	144	1	288	7
DodgerLoopGame	144	1	288	2
DodgerLoopWeekend	144	1	288	2
ECG200	200	1	96	2
ECG5000	5000	1	140	5
ECGFiveDays	884	1	136	2
EOGHorizontalSignal	724	1	1250	12
EOGVerticalSignal	724	1	1250	12
Earthquakes	461	1	512	2
ElectricDevices	16637	1	96	7

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 March 2026 doi:10.20944/preprints202603.1778.v1

Table A9. *Cont.*

Dataset	Samples	Channels	Length	Classes
EthanolLevel	1004	1	1751	4
FaceAll	2250	1	131	14
FaceFour	112	1	350	4
FacesUCR	2250	1	131	14
FiftyWords	905	1	270	50
Fish	350	1	463	7
FordA	4921	1	500	2
FordB	4446	1	500	2
FreezerRegularTrain	3000	1	301	2
FreezerSmallTrain	2878	1	301	2
Fungi	204	1	201	18
GestureMidAirD1	338	1	360	26
GestureMidAirD2	338	1	360	26
GestureMidAirD3	338	1	360	26
GesturePebbleZ1	304	1	455	6
GesturePebbleZ2	304	1	455	6
GunPoint	200	1	150	2
GunPointAgeSpan	451	1	150	2
GunPointMaleVersusFemale	451	1	150	2
GunPointOldVersusYoung	451	1	150	2
Ham	214	1	431	2
HandOutlines	1370	1	2709	2
Haptics	463	1	1092	5
Herring	128	1	512	2
HouseTwenty	159	1	2000	2
InlineSkate	650	1	1882	7
InsectEPGRegularTrain	311	1	601	3
InsectEPGSmallTrain	266	1	601	3
InsectWingbeatSound	2200	1	256	11
ItalyPowerDemand	1096	1	24	2
LargeKitchenAppliances	750	1	720	3
Lightning2	121	1	637	2
Lightning7	143	1	319	7
Mallat	2400	1	1024	8
Meat	120	1	448	3

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 March 2026 doi:10.20944/preprints202603.1778.v1

Table A9. *Cont.*

Dataset	Samples	Channels	Length	Classes
MedicalImages	1141	1	99	10
MelbournePedestrian	3457	1	24	10
MiddlePhalanxOutlineAgeGroup	554	1	80	3
MiddlePhalanxOutlineCorrect	891	1	80	2
MiddlePhalanxTW	553	1	80	6
MixedShapesRegularTrain	2925	1	1024	5
MixedShapesSmallTrain	2525	1	1024	5
MoteStrain	1272	1	84	2
NonInvasiveFetalECGThorax1	3765	1	750	42
NonInvasiveFetalECGThorax2	3765	1	750	42
OSULeaf	442	1	427	6
OliveOil	60	1	570	4
PLAID	1074	1	500	11
PhalangesOutlinesCorrect	2658	1	80	2
Phoneme	2110	1	1024	39
PickupGestureWiimoteZ	100	1	361	10
PigAirwayPressure	312	1	2000	52
PigArtPressure	312	1	2000	52
PigCVP	312	1	2000	52
Plane	210	1	144	7
PowerCons	360	1	144	2
ProximalPhalanxOutlineAgeGroup	605	1	80	3
ProximalPhalanxOutlineCorrect	891	1	80	2
ProximalPhalanxTW	605	1	80	6
RefrigerationDevices	750	1	720	3
Rock	70	1	2844	4
ScreenType	750	1	720	3
SemgHandGenderCh2	900	1	1500	2
SemgHandMovementCh2	900	1	1500	6
SemgHandSubjectCh2	900	1	1500	5
ShakeGestureWiimoteZ	100	1	385	10
ShapeletSim	200	1	500	2
ShapesAll	1200	1	512	60
SmallKitchenAppliances	750	1	720	3
SmoothSubspace	300	1	15	3
SonyAIBORobotSurface1	621	1	70	2

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 March 2026 doi:10.20944/preprints202603.1778.v1

Table A9. *Cont.*

Dataset	Samples	Channels	Length	Classes
SonyAIBORobotSurface2	980	1	65	2
StarLightCurves	9236	1	1024	3
Strawberry	983	1	235	2
SwedishLeaf	1125	1	128	15
Symbols	1020	1	398	6
SyntheticControl	600	1	60	6
ToeSegmentation1	268	1	277	2
ToeSegmentation2	166	1	343	2
Trace	200	1	275	4
TwoLeadECG	1162	1	82	2
TwoPatterns	5000	1	128	4
UMD	180	1	150	3
UWaveGestureLibraryAll	4478	1	945	8
UWaveGestureLibraryX	4478	1	315	8
UWaveGestureLibraryY	4478	1	315	8
UWaveGestureLibraryZ	4478	1	315	8
Wafer	7164	1	152	2
Wine	111	1	234	2
WordSynonyms	905	1	270	25
Worms	258	1	900	5
WormsTwoClass	258	1	900	2
Yoga	3300	1	426	2

Table A10. Multivariate Classification Datasets.

Dataset	Samples	Channels	Length	Classes
ArticularyWordRecognition	575	9	144	25
AtrialFibrillation	30	2	640	3
BasicMotions	80	6	100	4
ERing	300	4	65	6
HandMovementDirection	234	10	400	4
Handwriting	1000	3	152	26
LSST	4925	6	36	14
Libras	360	2	45	15
NATOPS	360	24	51	6
PenDigits	10992	2	8	10
RacketSports	303	6	30	4
SelfRegulationSCP1	561	6	896	2
SelfRegulationSCP2	380	7	1152	2
StandWalkJump	27	4	2500	3
UWaveGestureLibrary	440	3	315	8

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 March 2026 doi:10.20944/preprints202603.1778.v1

Table A11. Univariate Regression Datasets.

Dataset	Samples	Channels	Length
AcousticContaminationMadrid	238	1	365
AluminiumConcentration	629	1	2542
BoronConcentration	626	1	2542
CalciumConcentration	635	1	2307
CopperConcentration	629	1	2542
Covid19Andalusia	204	1	91
Covid3Month	201	1	84
DhakaHourlyAirQuality	2068	1	24
FloodModeling1	673	1	266
FloodModeling2	667	1	266
FloodModeling3	613	1	266
GasSensorArrayAcetone	464	1	7500
GasSensorArrayEthanol	464	1	7500
IronConcentration	611	1	1716
LPGasMonitoringHomeActivity	2882	1	100
MagnesiumConcentration	2229	1	3578
ManganeseConcentration	611	1	1716
MethaneMonitoringHomeActivity	2882	1	100
NaturalGasPricesSentiment	93	1	20
PhosphorusConcentration	2248	1	3578
PotassiumConcentration	2230	1	3578
SodiumConcentration	607	1	1716
SulphurConcentration	635	1	2307
WaveDataTension	1893	1	57
WindTurbinePower	852	1	144
ZincConcentration	636	1	2307

Table A12. Multivariate Regression Datasets.

Dataset	Samples	Channels	Length
BarCrawl6min	201	3	360
BeijingIntAirportPM25Quality	1571	6	24
BinanceCoinSentiment	263	2	24
BitcoinSentiment	332	2	24
CardanoSentiment	107	2	24
ChilledWaterPredictor	459	4	168
DailyOilGasPrices	191	2	30
ElectricityPredictor	810	4	168
EthereumSentiment	356	2	24
HotwaterPredictor	351	4	168
HouseholdPowerConsumption1	1431	5	1440
HouseholdPowerConsumption2	1431	5	1440
IEEEPPG	3096	5	1000
MetroInterstateTrafficVolume	1214	4	24
OccupancyDetectionLight	340	3	60
PrecipitationAndalusia	672	4	365
SierraNevadaMountainsSnow	500	3	30
SolarRadiationAndalusia	672	2	365
SteamPredictor	300	4	168

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 March 2026 doi:10.20944/preprints202603.1778.v1

Appendix M. Feature Counts for Univariate Classification

Table A13. Feature Counts for Univariate Classification with *GPF* Curvature.

Dataset	Sig	GPF_S	GPF_U	GPF_{SU}	GPF_S + Sig	GPF_U + Sig	GPF_{SU} + Sig
All Datasets	8	1	1	2	9	9	10

Table A14. Feature Counts for Univariate Classification with *GPF* Chord + Curvature.

Dataset	Sig	GPF_S	GPF_U	GPF_{SU}	GPF_S + Sig	GPF_U + Sig	GPF_{SU} + Sig
All Datasets	8	2	2	4	10	10	12

Appendix N. Feature Counts for Multivariate Classification

Table A15. Feature Counts for Multivariate Classification with Dyadic *GPF* Curvature.

Dataset	Sig	GPF_S	GPF_U	GPF_{SU}	GPF_S + Sig	GPF_U + Sig	GPF_{SU} + Sig
ArticularyWordRecognition	7380	1785	1785	3570	9165	9165	10950
AtrialFibrillation	30	21	21	42	51	51	72
BasicMotions	1554	392	392	784	1946	1946	2338
ERing	340	112	112	224	452	452	564
HandMovementDirection	11110	2695	2695	5390	13805	13805	16500
Handwriting	120	56	56	112	176	176	232
Libras	30	21	21	42	51	51	72
LSST	1554	392	392	784	1946	1946	2338
NATOPS	346200	90650	90650	181300	436850	436850	527500
PenDigits	30	21	21	42	51	51	72
RacketSports	1554	392	392	784	1946	1946	2338
SelfRegulationSCP1	1554	392	392	784	1946	1946	2338
SelfRegulationSCP2	2800	686	686	1372	3486	3486	4172
StandWalkJump	340	112	112	224	452	452	564
UWaveGestureLibrary	120	56	56	112	176	176	232

Table A16. Feature Counts for Multivariate Classification with Dyadic *GPF* Chord + Curvature.

Dataset	Sig	GPF_S	GPF_U	GPF_{SU}	GPF_S + Sig	GPF_U + Sig	GPF_{SU} + Sig
ArticularyWordRecognition	7380	1848	1848	3696	9228	9228	11076
AtrialFibrillation	30	35	35	70	65	65	100
BasicMotions	1554	434	434	868	1988	1988	2422
ERing	340	140	140	280	480	480	620
HandMovementDirection	11110	2765	2765	5530	13875	13875	16640
Handwriting	120	77	77	154	197	197	274
Libras	30	35	35	70	65	65	100
LSST	1554	434	434	868	1988	1988	2422
NATOPS	346200	90818	90818	181636	437018	437018	527836
PenDigits	30	35	35	70	65	65	100
RacketSports	1554	434	434	868	1988	1988	2422
SelfRegulationSCP1	1554	434	434	868	1988	1988	2422
SelfRegulationSCP2	2800	735	735	1470	3535	3535	4270
StandWalkJump	340	140	140	280	480	480	620
UWaveGestureLibrary	120	77	77	154	197	197	274

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 March 2026 doi:10.20944/preprints202603.1778.v1

Appendix O. Feature Counts for Univariate Regression

Table A17. Feature Counts for Univariate Regression with *GPF* Curvature.

Dataset	Sig	GPF_S	GPF_U	GPF_{SU}	GPF_S + Sig	GPF_U + Sig	GPF_{SU} + Sig
All Datasets	8	1	1	2	9	9	10

Table A18. Feature Counts for Univariate Regression with *GPF* Chord + Curvature.

Dataset	Sig	GPF_S	GPF_U	GPF_{SU}	GPF_S + Sig	GPF_U + Sig	GPF_{SU} + Sig
All Datasets	8	2	2	4	10	10	12

Appendix P. Feature Counts for Multivariate Regression

Table A19. Feature Counts for Multivariate Regression with *GPF* Chord + Curvature.

Dataset	Sig	GPF_S	GPF_U	GPF_{SU}	GPF_S + Sig	GPF_U + Sig	GPF_{SU} + Sig
BarCrawl6min	120	11	11	22	131	131	142
BeijingIntAirportPM25Quality	1554	62	62	124	1616	1616	1678
BinanceCoinSentiment	30	30	5	5	5	5	10
BitcoinSentiment	30	5	5	10	35	35	40
CardanoSentiment	30	5	5	10	35	35	40
ChilledWaterPredictor	340	20	20	40	360	360	380
DailyOilGasPrices	30	30	5	5	5	5	10
ElectricityPredictor	340	50	20	20	20	20	40
EthereumSentiment	30	5	5	10	35	35	40
HotwaterPredictor	340	20	20	40	360	360	380
HouseholdPowerConsumption1	780	35	35	70	815	815	830
HouseholdPowerConsumption2	780	35	35	70	815	815	830
IEEEPPG	780	35	35	70	815	815	830
MetroInterstateTrafficVolume	340	20	20	40	360	360	380
OccupancyDetectionLight	120	11	11	22	131	131	142
PrecipitationAndalusia	340	20	20	40	360	360	380
SierraNevadaMountainsSnow	120	11	11	22	131	131	142
SolarRadiationAndalusia	30	5	5	10	35	35	40
SteamPredictor	340	20	20	40	360	360	380

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 March 2026 doi:10.20944/preprints202603.1778.v1

Table A20. Feature Counts for Multivariate Regression with *GPF* Chord + Curvature.

Dataset	Sig	GPF_S	GPF_U	GPF_{SU}	GPF_S + Sig	GPF_U + Sig	GPF_{SU} + Sig
BarCrawl6min	120	14	14	28	134	134	148
BeijingIntAirportPM25Quality	1554	68	68	136	1622	1622	1690
BinanceCoinSentiment	30	7	7	14	37	37	44
BitcoinSentiment	30	7	7	14	37	37	44
CardanoSentiment	30	7	7	14	37	37	44
ChilledWaterPredictor	340	24	24	48	364	364	388
DailyOilGasPrices	30	7	7	14	37	37	44
ElectricityPredictor	340	24	24	48	364	364	388
EthereumSentiment	30	7	7	14	37	37	44
HotwaterPredictor	340	24	24	48	364	364	388
HouseholdPowerConsumption1	780	40	40	80	820	820	860
HouseholdPowerConsumption2	780	40	40	80	820	820	860
IEEEPPG	780	40	40	80	780	780	820
MetroInterstateTrafficVolume	340	24	24	48	364	364	388
OccupancyDetectionLight	120	14	14	28	134	134	148
PrecipitationAndalusia	340	24	24	48	364	364	388
SierraNevadaMountainsSnow	120	14	14	28	134	134	148
SolarRadiationAndalusia	30	7	7	14	37	37	44
SteamPredictor	340	24	24	48	364	364	388

Appendix Q. Results for Univariate Classification

Table A21. Metrics Results (Accuracy) for Univariate Classification with *GPF* Curvature.

Dataset	Sig	GPF_S	GPF_U	GPF_{SU}	GPF_S + Sig	GPF_U + Sig	GPF_{SU} + Sig
ACSF1	0.47±0.09	0.28±0.07	0.24±0.08	0.51±0.04	0.57±0.07	0.54±0.09	0.56±0.06
Adiac	0.16±0.02	0.04±0.02	0.09±0.02	0.10±0.01	0.15±0.02	0.22±0.03	0.22±0.03
AllGestureWiimoteX	0.16±0.02	0.15±0.02	0.15±0.01	0.21±0.02	0.17±0.02	0.23±0.01	0.25±0.02
AllGestureWiimoteY	0.19±0.02	0.15±0.02	0.11±0.02	0.15±0.03	0.25±0.02	0.27±0.03	0.32±0.03
AllGestureWiimoteZ	0.17±0.02	0.12±0.01	0.13±0.02	0.17±0.02	0.21±0.02	0.22±0.02	0.25±0.02
ArrowHead	0.46±0.04	0.40±0.04	0.49±0.06	0.49±0.10	0.46±0.04	0.52±0.05	0.51±0.04
Beef	0.33±0.11	0.12±0.04	0.28±0.08	0.22±0.08	0.38±0.12	0.38±0.14	0.43±0.14
BeetleFly	0.45±0.17	0.53±0.15	0.45±0.17	0.45±0.10	0.50±0.11	0.55±0.10	0.53±0.12
BirdChicken	0.70±0.10	0.57±0.06	0.62±0.19	0.68±0.10	0.65±0.09	0.68±0.06	0.70±0.06
BME	0.38±0.06	0.28±0.09	0.38±0.08	0.43±0.04	0.40±0.11	0.39±0.07	0.43±0.10
Car	0.36±0.08	0.28±0.05	0.22±0.09	0.43±0.13	0.43±0.08	0.49±0.05	0.45±0.08
CBF	0.37±0.04	0.33±0.02	0.34±0.01	0.34±0.03	0.37±0.01	0.36±0.02	0.38±0.02
Chinatown	0.75±0.03	0.66±0.05	0.62±0.08	0.69±0.04	0.80±0.03	0.79±0.03	0.83±0.03
ChlorineConcentration	0.41±0.01	0.39±0.02	0.40±0.01	0.48±0.02	0.44±0.01	0.47±0.01	0.51±0.02
CinCECGTorso	0.45±0.01	0.31±0.02	0.44±0.02	0.55±0.03	0.52±0.02	0.66±0.01	0.71±0.02
Coffee	0.37±0.18	0.45±0.16	0.50±0.16	0.45±0.19	0.34±0.15	0.42±0.18	0.43±0.14
Computers	0.64±0.05	0.48±0.04	0.50±0.05	0.57±0.02	0.64±0.06	0.64±0.04	0.64±0.05
CricketX	0.12±0.02	0.12±0.02	0.09±0.01	0.15±0.02	0.16±0.01	0.15±0.03	0.17±0.02
CricketY	0.09±0.02	0.08±0.01	0.10±0.03	0.12±0.01	0.10±0.02	0.11±0.02	0.13±0.03
CricketZ	0.10±0.02	0.13±0.01	0.10±0.03	0.17±0.04	0.15±0.02	0.14±0.03	0.18±0.02
Crop	0.21±0.00	0.32±0.01	0.33±0.00	0.38±0.01	0.35±0.01	0.37±0.01	0.42±0.01
DiatomSizeReduction	0.52±0.05	0.35±0.05	0.47±0.03	0.59±0.04	0.60±0.04	0.74±0.03	0.77±0.03
DistalPhalanxOutlineAgeGroup	0.74±0.03	0.48±0.02	0.58±0.02	0.65±0.02	0.77±0.03	0.76±0.02	0.77±0.02
DistalPhalanxOutlineCorrect	0.68±0.04	0.53±0.03	0.53±0.02	0.55±0.03	0.73±0.04	0.74±0.04	0.75±0.03
DistalPhalanxTW	0.67±0.03	0.31±0.05	0.47±0.03	0.52±0.07	0.71±0.04	0.70±0.04	0.72±0.05
DodgerLoopDay	0.20±0.09	0.10±0.04	0.19±0.06	0.17±0.08	0.16±0.06	0.17±0.06	0.14±0.05
DodgerLoopGame	0.49±0.05	0.46±0.10	0.56±0.03	0.55±0.04	0.51±0.06	0.49±0.02	0.51±0.05
DodgerLoopWeekend	0.74±0.10	0.56±0.10	0.56±0.06	0.64±0.10	0.77±0.12	0.73±0.11	0.75±0.11
Earthquakes	0.78±0.03	0.67±0.04	0.68±0.03	0.75±0.02	0.78±0.03	0.81±0.02	0.79±0.01
ECG200	0.56±0.07	0.56±0.07	0.60±0.06	0.62±0.06	0.57±0.07	0.62±0.09	0.67±0.07
ECG5000	0.50±0.01	0.48±0.01	0.51±0.01	0.59±0.02	0.55±0.02	0.61±0.01	0.65±0.02
ECGFiveDays	0.56±0.02	0.50±0.04	0.53±0.03	0.59±0.04	0.61±0.00	0.60±0.03	0.65±0.01
ElectricDevices	0.51±0.01	0.31±0.01	0.38±0.00	0.56±0.01	0.57±0.01	0.59±0.01	0.65±0.01
EOGHorizontalSignal	0.17±0.02	0.12±0.03	0.14±0.02	0.21±0.01	0.24±0.03	0.27±0.04	0.31±0.03
EOGVerticalSignal	0.11±0.02	0.10±0.03	0.08±0.01	0.10±0.02	0.14±0.03	0.16±0.04	0.18±0.04
EthanolLevel	0.24±0.02	0.23±0.02	0.25±0.02	0.28±0.04	0.24±0.02	0.29±0.01	0.29±0.03
FaceAll	0.22±0.01	0.09±0.00	0.11±0.01	0.14±0.01	0.25±0.00	0.26±0.02	0.27±0.02
FaceFour	0.33±0.08	0.39±0.04	0.26±0.06	0.45±0.12	0.36±0.10	0.30±0.05	0.43±0.13
FacesUCR	0.22±0.01	0.11±0.01	0.13±0.02	0.16±0.01	0.23±0.01	0.26±0.01	0.26±0.01
FiftyWords	0.10±0.03	0.07±0.02	0.06±0.01	0.11±0.03	0.12±0.02	0.13±0.02	0.15±0.02
Fish	0.25±0.04	0.18±0.03	0.13±0.03	0.19±0.05	0.34±0.04	0.28±0.05	0.34±0.03
FordA	0.49±0.02	0.51±0.02	0.53±0.02	0.56±0.01	0.50±0.02	0.55±0.01	0.55±0.01
FordB	0.49±0.01	0.48±0.01	0.51±0.01	0.52±0.02	0.49±0.01	0.50±0.01	0.50±0.01
FreezerRegularTrain	0.65±0.02	0.59±0.01	0.66±0.03	0.74±0.02	0.72±0.02	0.74±0.02	0.78±0.02

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 March 2026 doi:10.20944/preprints202603.1778.v1

Table A21. *Cont.*

Dataset	Sig	GPF_S	GPF_U	GPF_{SU}	GPF_S + Sig	GPF_U + Sig	GPF_{SU} + Sig
FreezerSmallTrain	0.66±0.01	0.61±0.01	0.66±0.02	0.75±0.01	0.72±0.01	0.73±0.01	0.78±0.01
Fungi	0.13±0.05	0.07±0.02	0.12±0.02	0.15±0.04	0.13±0.04	0.16±0.04	0.17±0.03
GestureMidAirD1	0.16±0.04	0.10±0.03	0.08±0.04	0.12±0.03	0.21±0.06	0.20±0.04	0.26±0.07
GestureMidAirD2	0.10±0.05	0.05±0.01	0.05±0.02	0.06±0.02	0.13±0.04	0.14±0.05	0.14±0.04
GestureMidAirD3	0.09±0.02	0.07±0.03	0.06±0.01	0.07±0.05	0.11±0.03	0.10±0.03	0.12±0.04
GesturePebbleZ1	0.20±0.05	0.13±0.01	0.20±0.04	0.18±0.05	0.18±0.06	0.22±0.05	0.23±0.06
GesturePebbleZ2	0.28±0.03	0.16±0.04	0.20±0.04	0.17±0.06	0.21±0.03	0.24±0.05	0.23±0.04
GunPoint	0.53±0.04	0.46±0.03	0.48±0.03	0.50±0.03	0.48±0.06	0.56±0.06	0.55±0.08
GunPointAgeSpan	0.55±0.03	0.75±0.04	0.47±0.03	0.84±0.02	0.81±0.04	0.57±0.05	0.83±0.02
GunPointMaleVersusFemale	0.46±0.06	0.71±0.02	0.61±0.04	0.80±0.03	0.71±0.04	0.63±0.05	0.75±0.04
GunPointOldVersusYoung	0.83±0.05	0.89±0.03	0.88±0.02	0.94±0.02	0.96±0.01	0.95±0.02	0.96±0.02
Ham	0.43±0.08	0.54±0.07	0.46±0.04	0.52±0.05	0.50±0.10	0.46±0.04	0.46±0.07
HandOutlines	0.59±0.02	0.55±0.02	0.52±0.03	0.58±0.02	0.60±0.03	0.62±0.03	0.63±0.03
Haptics	0.24±0.07	0.22±0.05	0.21±0.04	0.31±0.04	0.25±0.06	0.24±0.05	0.25±0.05
Herring	0.56±0.13	0.46±0.08	0.45±0.09	0.50±0.07	0.56±0.11	0.53±0.06	0.54±0.07
HouseTwenty	0.60±0.06	0.58±0.06	0.59±0.09	0.69±0.10	0.66±0.04	0.65±0.06	0.65±0.05
InlineSkate	0.24±0.03	0.19±0.02	0.30±0.05	0.33±0.06	0.24±0.03	0.30±0.04	0.30±0.03
InsectEPGRegularTrain	0.55±0.06	0.59±0.04	0.95±0.03	0.96±0.02	0.70±0.05	0.96±0.01	0.97±0.01
InsectEPGSmallTrain	0.59±0.04	0.62±0.09	0.95±0.03	0.96±0.03	0.70±0.05	0.96±0.02	0.97±0.02
InsectWingbeatSound	0.11±0.01	0.10±0.01	0.10±0.01	0.11±0.01	0.11±0.01	0.13±0.01	0.14±0.01
ItalyPowerDemand	0.56±0.02	0.49±0.03	0.53±0.03	0.55±0.03	0.57±0.03	0.58±0.02	0.58±0.02
LargeKitchenAppliances	0.54±0.03	0.37±0.04	0.36±0.02	0.43±0.04	0.53±0.04	0.58±0.03	0.60±0.01
Lightning2	0.60±0.04	0.60±0.07	0.66±0.06	0.58±0.12	0.62±0.05	0.70±0.06	0.68±0.06
Lightning7	0.34±0.04	0.26±0.08	0.20±0.05	0.34±0.05	0.29±0.03	0.36±0.07	0.36±0.04
Mallat	0.18±0.02	0.15±0.01	0.26±0.01	0.38±0.02	0.26±0.01	0.34±0.02	0.45±0.03
Meat	0.50±0.10	0.40±0.08	0.29±0.04	0.26±0.08	0.47±0.09	0.56±0.10	0.53±0.11
MedicalImages	0.39±0.01	0.33±0.02	0.35±0.03	0.45±0.03	0.47±0.03	0.48±0.03	0.53±0.04
MelbournePedestrian	0.34±0.02	0.25±0.01	0.32±0.01	0.46±0.02	0.46±0.02	0.52±0.03	0.59±0.02
MiddlePhalanxOutlineCorrect	0.60±0.01	0.53±0.04	0.56±0.04	0.56±0.03	0.64±0.01	0.63±0.02	0.65±0.03
MiddlePhalanxOutlineAgeGroup	0.64±0.05	0.42±0.03	0.39±0.04	0.45±0.04	0.68±0.03	0.67±0.04	0.70±0.03
MiddlePhalanxTW	0.50±0.04	0.27±0.03	0.23±0.02	0.31±0.03	0.54±0.02	0.52±0.01	0.55±0.02
MixedShapesRegularTrain	0.63±0.03	0.24±0.01	0.30±0.01	0.39±0.02	0.64±0.02	0.70±0.03	0.72±0.03
MixedShapesSmallTrain	0.63±0.02	0.23±0.02	0.32±0.02	0.39±0.01	0.65±0.01	0.71±0.01	0.73±0.01
MoteStrain	0.72±0.03	0.61±0.03	0.50±0.04	0.66±0.01	0.81±0.01	0.76±0.01	0.83±0.01
NonInvasiveFetalECGThorax1	0.06±0.01	0.03±0.01	0.07±0.01	0.07±0.01	0.05±0.00	0.11±0.01	0.11±0.01
NonInvasiveFetalECGThorax2	0.09±0.01	0.03±0.00	0.06±0.01	0.08±0.01	0.10±0.01	0.16±0.01	0.18±0.01
OliveOil	0.37±0.08	0.43±0.08	0.20±0.10	0.30±0.07	0.48±0.08	0.33±0.07	0.47±0.14
OSULeaf	0.25±0.07	0.23±0.05	0.37±0.03	0.43±0.04	0.29±0.03	0.47±0.06	0.47±0.05
PhalangesOutlinesCorrect	0.57±0.01	0.55±0.01	0.55±0.01	0.59±0.01	0.63±0.02	0.63±0.02	0.65±0.01
Phoneme	0.06±0.01	0.05±0.01	0.09±0.01	0.11±0.02	0.07±0.01	0.10±0.01	0.12±0.01
PickupGestureWiimoteZ	0.24±0.06	0.06±0.06	0.36±0.13	0.30±0.07	0.23±0.07	0.27±0.07	0.29±0.09
PigAirwayPressure	0.06±0.02	0.07±0.03	0.07±0.03	0.11±0.04	0.08±0.03	0.09±0.01	0.12±0.02
PigArtPressure	0.02±0.01	0.03±0.02	0.03±0.02	0.06±0.03	0.04±0.02	0.04±0.01	0.08±0.03
PigCVP	0.02±0.01	0.03±0.01	0.05±0.02	0.08±0.01	0.02±0.01	0.05±0.03	0.07±0.05
PLAID	0.32±0.02	0.24±0.02	0.26±0.04	0.34±0.03	0.40±0.02	0.44±0.03	0.48±0.04
Plane	0.30±0.06	0.22±0.06	0.13±0.03	0.25±0.07	0.34±0.04	0.35±0.09	0.41±0.10
PowerCons	0.79±0.06	0.83±0.02	0.85±0.03	0.91±0.04	0.84±0.05	0.90±0.03	0.91±0.04
ProximalPhalanxOutlineCorrect	0.69±0.01	0.58±0.02	0.59±0.04	0.64±0.03	0.73±0.03	0.72±0.01	0.74±0.02
ProximalPhalanxOutlineAgeGroup	0.77±0.03	0.38±0.03	0.42±0.03	0.40±0.04	0.80±0.03	0.81±0.04	0.81±0.03
ProximalPhalanxTW	0.73±0.03	0.31±0.01	0.28±0.02	0.36±0.06	0.76±0.04	0.76±0.03	0.76±0.02
RefrigerationDevices	0.40±0.04	0.33±0.03	0.39±0.03	0.45±0.04	0.40±0.06	0.43±0.05	0.44±0.08
Rock	0.47±0.07	0.41±0.19	0.37±0.14	0.53±0.10	0.53±0.03	0.60±0.03	0.69±0.10
ScreenType	0.44±0.06	0.37±0.03	0.36±0.03	0.47±0.04	0.45±0.07	0.47±0.03	0.49±0.05
SemgHandGenderCh2	0.58±0.04	0.58±0.03	0.61±0.04	0.69±0.02	0.63±0.02	0.63±0.04	0.69±0.04
SemgHandMovementCh2	0.21±0.02	0.23±0.03	0.29±0.04	0.34±0.02	0.30±0.02	0.34±0.02	0.37±0.03

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 March 2026 doi:10.20944/preprints202603.1778.v1

Table A21. *Cont.*

Dataset	Sig	GPF_S	GPF_U	GPF_{SU}	GPF_S + Sig	GPF_U + Sig	GPF_{SU} + Sig
SemgHandSubjectCh2	0.28±0.04	0.26±0.03	0.31±0.04	0.36±0.03	0.32±0.02	0.38±0.01	0.39±0.03
ShakeGestureWiimoteZ	0.31±0.04	0.13±0.07	0.19±0.09	0.27±0.10	0.30±0.06	0.39±0.02	0.35±0.05
ShapeletSim	0.55±0.06	0.43±0.03	1.00±0.00	1.00±0.00	0.55±0.04	1.00±0.00	1.00±0.00
ShapesAll	0.14±0.03	0.03±0.00	0.05±0.01	0.07±0.02	0.17±0.02	0.24±0.04	0.24±0.03
SmallKitchenAppliances	0.58±0.04	0.41±0.02	0.42±0.02	0.57±0.04	0.62±0.03	0.61±0.03	0.65±0.03
SmoothSubspace	0.44±0.05	0.31±0.05	0.33±0.08	0.37±0.07	0.41±0.05	0.46±0.04	0.43±0.02
SonyAIBORobotSurface1	0.65±0.04	0.58±0.03	0.53±0.05	0.56±0.03	0.71±0.07	0.67±0.04	0.70±0.06
SonyAIBORobotSurface2	0.72±0.03	0.58±0.03	0.57±0.03	0.67±0.04	0.75±0.02	0.76±0.02	0.78±0.02
StarLightCurves	0.78±0.01	0.49±0.01	0.76±0.00	0.82±0.01	0.80±0.01	0.89±0.00	0.90±0.01
Strawberry	0.60±0.02	0.57±0.02	0.55±0.02	0.60±0.02	0.60±0.04	0.64±0.02	0.63±0.03
SwedishLeaf	0.23±0.02	0.08±0.01	0.12±0.01	0.14±0.02	0.25±0.02	0.33±0.03	0.32±0.03
Symbols	0.61±0.02	0.43±0.02	0.27±0.02	0.57±0.03	0.74±0.02	0.65±0.02	0.75±0.02
SyntheticControl	0.51±0.03	0.20±0.02	0.19±0.03	0.25±0.03	0.52±0.05	0.53±0.04	0.53±0.04
ToeSegmentation1	0.57±0.04	0.59±0.09	0.48±0.03	0.55±0.03	0.60±0.05	0.58±0.06	0.60±0.05
ToeSegmentation2	0.62±0.06	0.68±0.06	0.75±0.09	0.81±0.06	0.67±0.03	0.78±0.06	0.79±0.06
Trace	0.49±0.10	0.36±0.07	0.23±0.04	0.39±0.09	0.56±0.05	0.48±0.07	0.56±0.06
TwoLeadECG	0.58±0.02	0.49±0.02	0.53±0.03	0.54±0.02	0.60±0.03	0.60±0.02	0.61±0.01
TwoPatterns	0.27±0.01	0.28±0.01	0.26±0.01	0.31±0.02	0.31±0.02	0.29±0.01	0.33±0.02
UMD	0.43±0.06	0.34±0.10	0.48±0.06	0.44±0.04	0.43±0.07	0.46±0.07	0.47±0.06
UWaveGestureLibraryAll	0.20±0.01	0.20±0.02	0.15±0.01	0.26±0.01	0.29±0.02	0.24±0.01	0.35±0.01
UWaveGestureLibraryX	0.22±0.02	0.18±0.01	0.15±0.01	0.24±0.01	0.32±0.02	0.28±0.01	0.37±0.02
UWaveGestureLibraryY	0.29±0.01	0.22±0.01	0.13±0.01	0.27±0.01	0.35±0.01	0.32±0.02	0.37±0.01
UWaveGestureLibraryZ	0.23±0.01	0.20±0.01	0.14±0.01	0.26±0.01	0.30±0.01	0.29±0.01	0.37±0.01
Wafer	0.96±0.01	0.82±0.01	0.91±0.00	0.95±0.01	0.98±0.00	0.99±0.00	0.99±0.00
Wine	0.50±0.10	0.55±0.05	0.63±0.08	0.54±0.05	0.50±0.07	0.58±0.13	0.54±0.10
WordSynonyms	0.18±0.04	0.10±0.02	0.10±0.02	0.16±0.03	0.19±0.02	0.20±0.02	0.23±0.04
Worms	0.44±0.05	0.49±0.03	0.56±0.04	0.66±0.06	0.46±0.05	0.55±0.05	0.56±0.04
WormsTwoClass	0.59±0.06	0.64±0.02	0.65±0.05	0.74±0.08	0.59±0.07	0.66±0.04	0.68±0.03
Yoga	0.60±0.01	0.53±0.02	0.51±0.01	0.56±0.02	0.64±0.02	0.64±0.02	0.68±0.02

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 March 2026 doi:10.20944/preprints202603.1778.v1

Table A22. Metrics Results (Accuracy) for Univariate Classification with *GPF* Chord + Curvature.

Dataset	Sig	GPF_S	GPF_U	GPF_{SU}	GPF_S + Sig	GPF_U + Sig	GPF_{SU} + Sig
ACSF1	0.47±0.09	0.57±0.08	0.60±0.04	0.75±0.06	0.56±0.08	0.68±0.05	0.68±0.06
Adiac	0.16±0.02	0.14±0.02	0.26±0.04	0.41±0.03	0.20±0.02	0.39±0.02	0.37±0.03
AllGestureWiimoteX	0.16±0.02	0.17±0.02	0.26±0.02	0.31±0.03	0.17±0.03	0.30±0.02	0.33±0.03
AllGestureWiimoteY	0.19±0.02	0.18±0.03	0.20±0.01	0.34±0.04	0.25±0.02	0.36±0.03	0.39±0.02
AllGestureWiimoteZ	0.17±0.02	0.14±0.02	0.24±0.01	0.29±0.01	0.20±0.02	0.31±0.01	0.33±0.03
ArrowHead	0.46±0.04	0.42±0.06	0.63±0.04	0.68±0.04	0.45±0.03	0.67±0.05	0.68±0.05
Beef	0.33±0.11	0.33±0.16	0.43±0.20	0.45±0.20	0.40±0.11	0.40±0.12	0.47±0.14
BeetleFly	0.45±0.17	0.47±0.18	0.70±0.19	0.70±0.19	0.53±0.12	0.68±0.06	0.62±0.08
BirdChicken	0.70±0.10	0.65±0.09	0.68±0.13	0.65±0.18	0.70±0.06	0.72±0.09	0.72±0.09
BME	0.38±0.06	0.34±0.07	0.46±0.05	0.50±0.08	0.48±0.12	0.48±0.07	0.59±0.03
Car	0.36±0.08	0.42±0.04	0.43±0.05	0.53±0.04	0.40±0.04	0.52±0.09	0.50±0.05
CBF	0.37±0.04	0.36±0.03	0.35±0.03	0.41±0.02	0.46±0.04	0.42±0.03	0.45±0.03
Chinatown	0.75±0.03	0.80±0.04	0.82±0.03	0.95±0.01	0.92±0.01	0.93±0.02	0.96±0.02
ChlorineConcentration	0.41±0.01	0.46±0.01	0.48±0.02	0.55±0.01	0.50±0.02	0.52±0.02	0.58±0.01
CinCECGTorso	0.45±0.01	0.43±0.01	0.59±0.03	0.67±0.02	0.58±0.02	0.72±0.01	0.74±0.02
Coffee	0.37±0.18	0.50±0.14	0.64±0.09	0.64±0.13	0.32±0.13	0.68±0.15	0.64±0.16
Computers	0.64±0.05	0.62±0.08	0.61±0.05	0.67±0.05	0.66±0.06	0.65±0.04	0.66±0.05
CricketX	0.12±0.02	0.13±0.01	0.13±0.02	0.20±0.02	0.18±0.01	0.18±0.02	0.23±0.03
CricketY	0.09±0.02	0.09±0.02	0.15±0.02	0.18±0.03	0.16±0.03	0.20±0.03	0.21±0.02
CricketZ	0.10±0.02	0.17±0.01	0.14±0.01	0.25±0.01	0.20±0.01	0.21±0.01	0.27±0.02
Crop	0.21±0.00	0.34±0.01	0.40±0.01	0.51±0.01	0.42±0.01	0.45±0.01	0.54±0.01
DiatomSizeReduction	0.52±0.05	0.56±0.07	0.66±0.05	0.88±0.03	0.65±0.06	0.84±0.03	0.84±0.03
DistalPhalanxOutlineAgeGroup	0.74±0.03	0.80±0.01	0.72±0.01	0.78±0.01	0.78±0.02	0.77±0.02	0.77±0.01
DistalPhalanxOutlineCorrect	0.68±0.04	0.73±0.02	0.66±0.02	0.75±0.01	0.75±0.04	0.76±0.04	0.78±0.03
DistalPhalanxTW	0.67±0.03	0.69±0.05	0.65±0.03	0.73±0.04	0.72±0.04	0.73±0.03	0.72±0.04
DodgerLoopDay	0.20±0.09	0.20±0.06	0.25±0.08	0.30±0.07	0.24±0.06	0.24±0.06	0.33±0.05
DodgerLoopGame	0.49±0.05	0.53±0.06	0.50±0.07	0.51±0.10	0.53±0.02	0.49±0.04	0.51±0.06
DodgerLoopWeekend	0.74±0.10	0.73±0.10	0.71±0.07	0.75±0.03	0.86±0.09	0.76±0.08	0.83±0.07
Earthquakes	0.78±0.03	0.74±0.03	0.76±0.02	0.79±0.03	0.79±0.04	0.80±0.03	0.79±0.02
ECG200	0.56±0.07	0.61±0.07	0.57±0.03	0.62±0.08	0.65±0.02	0.66±0.06	0.67±0.06
ECG5000	0.50±0.01	0.59±0.01	0.59±0.01	0.71±0.01	0.61±0.01	0.69±0.01	0.74±0.01
ECGFiveDays	0.56±0.02	0.64±0.05	0.65±0.02	0.70±0.04	0.67±0.02	0.66±0.03	0.71±0.02
ElectricDevices	0.51±0.01	0.46±0.01	0.62±0.00	0.72±0.00	0.61±0.01	0.72±0.00	0.74±0.00
EOGHorizontalSignal	0.17±0.02	0.20±0.03	0.21±0.03	0.36±0.02	0.27±0.03	0.33±0.05	0.40±0.02
EOGVerticalSignal	0.11±0.02	0.14±0.02	0.11±0.02	0.21±0.02	0.15±0.03	0.19±0.04	0.23±0.04
EthanolLevel	0.24±0.02	0.23±0.02	0.29±0.02	0.30±0.01	0.24±0.02	0.29±0.01	0.29±0.03
FaceAll	0.22±0.01	0.23±0.01	0.24±0.01	0.35±0.01	0.32±0.01	0.36±0.02	0.39±0.02
FaceFour	0.33±0.08	0.52±0.09	0.31±0.05	0.57±0.05	0.37±0.09	0.41±0.05	0.48±0.08
FacesUCR	0.22±0.01	0.24±0.01	0.25±0.02	0.34±0.01	0.32±0.01	0.34±0.01	0.40±0.02
FiftyWords	0.10±0.03	0.12±0.02	0.17±0.00	0.27±0.02	0.15±0.02	0.24±0.02	0.26±0.03
Fish	0.25±0.04	0.32±0.03	0.30±0.06	0.46±0.06	0.34±0.04	0.44±0.03	0.47±0.03
FordA	0.49±0.02	0.50±0.02	0.56±0.01	0.58±0.01	0.50±0.01	0.56±0.02	0.58±0.01
FordB	0.49±0.01	0.50±0.02	0.52±0.02	0.56±0.02	0.49±0.01	0.52±0.01	0.55±0.01
FreezerRegularTrain	0.65±0.02	0.69±0.01	0.74±0.02	0.80±0.02	0.75±0.01	0.78±0.02	0.80±0.03
FreezerSmallTrain	0.66±0.01	0.70±0.01	0.74±0.02	0.79±0.02	0.76±0.02	0.78±0.01	0.81±0.01
Fungi	0.13±0.05	0.06±0.04	0.16±0.05	0.19±0.05	0.13±0.04	0.17±0.04	0.20±0.05
GestureMidAirD1	0.16±0.04	0.18±0.02	0.08±0.03	0.23±0.06	0.22±0.07	0.21±0.06	0.26±0.06
GestureMidAirD2	0.10±0.05	0.11±0.03	0.10±0.04	0.22±0.03	0.12±0.04	0.20±0.06	0.21±0.07
GestureMidAirD3	0.09±0.02	0.10±0.03	0.06±0.01	0.11±0.05	0.09±0.02	0.11±0.03	0.12±0.03

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 March 2026 doi:10.20944/preprints202603.1778.v1

Table A22. *Cont.*

Dataset	Sig	GPF_S	GPF_U	GPF_{SU}	GPF_S + Sig	GPF_U + Sig	GPF_{SU} + Sig
GesturePebbleZ1	0.20±0.05	0.14±0.02	0.22±0.04	0.31±0.04	0.20±0.05	0.34±0.07	0.34±0.09
GesturePebbleZ2	0.28±0.03	0.14±0.04	0.21±0.04	0.32±0.03	0.24±0.03	0.36±0.04	0.34±0.03
GunPoint	0.53±0.04	0.49±0.07	0.76±0.02	0.74±0.05	0.53±0.07	0.70±0.08	0.72±0.08
GunPointAgeSpan	0.55±0.03	0.82±0.03	0.78±0.03	0.88±0.02	0.80±0.03	0.75±0.04	0.86±0.03
GunPointMaleVersusFemale	0.46±0.06	0.77±0.04	0.90±0.03	0.92±0.03	0.86±0.02	0.90±0.02	0.92±0.02
GunPointOldVersusYoung	0.83±0.05	0.93±0.01	1.00±0.00	1.00±0.00	0.97±0.02	1.00±0.00	1.00±0.00
Ham	0.43±0.08	0.54±0.08	0.58±0.06	0.58±0.03	0.50±0.05	0.51±0.06	0.56±0.05
HandOutlines	0.59±0.02	0.59±0.02	0.67±0.02	0.69±0.02	0.61±0.02	0.70±0.02	0.70±0.02
Haptics	0.24±0.07	0.27±0.02	0.30±0.02	0.30±0.05	0.25±0.05	0.26±0.03	0.29±0.04
Herring	0.56±0.13	0.54±0.06	0.44±0.10	0.44±0.09	0.61±0.07	0.52±0.09	0.55±0.05
HouseTwenty	0.60±0.06	0.68±0.05	0.68±0.09	0.73±0.08	0.63±0.05	0.64±0.05	0.69±0.07
InlineSkate	0.24±0.03	0.26±0.02	0.32±0.05	0.37±0.04	0.27±0.03	0.34±0.02	0.38±0.04
InsectEPGRegularTrain	0.55±0.06	0.67±0.04	0.99±0.01	0.99±0.01	0.70±0.05	0.97±0.02	0.98±0.01
InsectEPGSmallTrain	0.59±0.04	0.70±0.04	0.99±0.02	0.99±0.02	0.74±0.04	0.98±0.02	0.97±0.03
InsectWingbeatSound	0.11±0.01	0.14±0.01	0.12±0.01	0.23±0.02	0.19±0.01	0.18±0.01	0.24±0.02
ItalyPowerDemand	0.56±0.02	0.59±0.03	0.73±0.02	0.80±0.02	0.78±0.01	0.77±0.03	0.80±0.02
LargeKitchenAppliances	0.54±0.03	0.41±0.02	0.55±0.05	0.59±0.05	0.55±0.03	0.64±0.05	0.66±0.03
Lightning2	0.60±0.04	0.56±0.07	0.66±0.08	0.65±0.06	0.59±0.06	0.73±0.07	0.70±0.06
Lightning7	0.34±0.04	0.24±0.04	0.31±0.07	0.43±0.07	0.31±0.11	0.49±0.09	0.49±0.03
Mallat	0.18±0.02	0.21±0.01	0.42±0.01	0.62±0.02	0.29±0.01	0.50±0.01	0.60±0.01
Meat	0.50±0.10	0.48±0.12	0.58±0.07	0.67±0.08	0.60±0.07	0.63±0.14	0.68±0.03
MedicalImages	0.39±0.01	0.48±0.03	0.50±0.02	0.63±0.02	0.54±0.02	0.62±0.04	0.65±0.03
MelbournePedestrian	0.34±0.02	0.47±0.01	0.59±0.01	0.71±0.01	0.59±0.02	0.64±0.02	0.73±0.01
MiddlePhalanxOutlineCorrect	0.60±0.01	0.66±0.01	0.63±0.02	0.68±0.03	0.67±0.02	0.68±0.01	0.69±0.03
MiddlePhalanxOutlineAgeGroup	0.64±0.05	0.70±0.02	0.65±0.05	0.73±0.03	0.67±0.04	0.69±0.04	0.70±0.04
MiddlePhalanxTW	0.50±0.04	0.54±0.01	0.51±0.02	0.59±0.01	0.54±0.01	0.54±0.02	0.55±0.01
MixedShapesRegularTrain	0.63±0.03	0.56±0.02	0.44±0.03	0.75±0.02	0.70±0.03	0.77±0.02	0.78±0.02
MixedShapesSmallTrain	0.63±0.02	0.55±0.02	0.45±0.03	0.75±0.02	0.69±0.02	0.76±0.01	0.78±0.01
MoteStrain	0.72±0.03	0.65±0.02	0.60±0.02	0.73±0.03	0.85±0.01	0.79±0.02	0.86±0.02
NonInvasiveFetalECGThorax1	0.06±0.01	0.04±0.01	0.19±0.01	0.25±0.01	0.10±0.01	0.25±0.01	0.26±0.01
NonInvasiveFetalECGThorax2	0.09±0.01	0.07±0.01	0.23±0.01	0.36±0.01	0.15±0.01	0.33±0.01	0.35±0.01
OliveOil	0.37±0.08	0.47±0.08	0.45±0.15	0.57±0.12	0.52±0.19	0.38±0.08	0.48±0.13
OSULeaf	0.25±0.07	0.31±0.04	0.50±0.05	0.57±0.03	0.31±0.05	0.54±0.07	0.53±0.06
PhalangesOutlinesCorrect	0.57±0.01	0.66±0.02	0.62±0.01	0.71±0.02	0.66±0.02	0.67±0.01	0.71±0.01
Phoneme	0.06±0.01	0.08±0.01	0.13±0.01	0.18±0.01	0.09±0.01	0.15±0.02	0.18±0.01
PickupGestureWiimoteZ	0.24±0.06	0.13±0.05	0.37±0.12	0.53±0.06	0.24±0.08	0.43±0.09	0.40±0.09
PigAirwayPressure	0.06±0.02	0.07±0.02	0.12±0.02	0.13±0.03	0.08±0.02	0.11±0.02	0.12±0.03
PigArtPressure	0.02±0.01	0.03±0.01	0.12±0.05	0.15±0.04	0.05±0.02	0.10±0.03	0.12±0.03
PigCVP	0.02±0.01	0.03±0.01	0.10±0.03	0.10±0.03	0.01±0.01	0.06±0.02	0.06±0.02
PLAID	0.32±0.02	0.49±0.02	0.46±0.04	0.68±0.03	0.54±0.03	0.56±0.05	0.66±0.06
Plane	0.30±0.06	0.20±0.05	0.83±0.06	0.83±0.05	0.33±0.07	0.79±0.06	0.76±0.06
PowerCons	0.79±0.06	0.85±0.03	0.88±0.05	0.92±0.03	0.87±0.04	0.91±0.04	0.92±0.02
ProximalPhalanxOutlineCorrect	0.69±0.01	0.74±0.02	0.71±0.02	0.77±0.03	0.73±0.02	0.75±0.02	0.75±0.02
ProximalPhalanxOutlineAgeGroup	0.77±0.03	0.80±0.05	0.74±0.04	0.82±0.02	0.81±0.04	0.81±0.05	0.82±0.04
ProximalPhalanxTW	0.73±0.03	0.79±0.02	0.69±0.02	0.80±0.03	0.78±0.02	0.78±0.02	0.78±0.02
RefrigerationDevices	0.40±0.04	0.39±0.03	0.41±0.03	0.48±0.05	0.45±0.06	0.47±0.05	0.50±0.04
Rock	0.47±0.07	0.54±0.09	0.66±0.12	0.64±0.16	0.61±0.09	0.64±0.09	0.69±0.10
ScreenType	0.44±0.06	0.40±0.04	0.43±0.04	0.48±0.06	0.46±0.05	0.48±0.03	0.50±0.02
SemgHandGenderCh2	0.58±0.04	0.65±0.03	0.64±0.04	0.71±0.04	0.69±0.04	0.67±0.06	0.73±0.06

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 March 2026 doi:10.20944/preprints202603.1778.v1

Table A22. *Cont.*

Dataset	Sig	GPF_S	GPF_U	GPF_{SU}	GPF_S + Sig	GPF_U + Sig	GPF_{SU} + Sig
SemgHandMovementCh2	0.21±0.02	0.33±0.02	0.36±0.03	0.43±0.04	0.37±0.02	0.40±0.04	0.45±0.03
SemgHandSubjectCh2	0.28±0.04	0.34±0.02	0.35±0.04	0.46±0.03	0.41±0.03	0.42±0.02	0.47±0.01
ShakeGestureWiimoteZ	0.31±0.04	0.20±0.08	0.38±0.12	0.51±0.06	0.32±0.05	0.47±0.07	0.44±0.04
ShapeletSim	0.55±0.06	0.47±0.07	1.00±0.00	1.00±0.00	0.51±0.06	1.00±0.00	1.00±0.00
ShapesAll	0.14±0.03	0.13±0.02	0.18±0.03	0.36±0.04	0.27±0.05	0.33±0.05	0.39±0.04
SmallKitchenAppliances	0.58±0.04	0.50±0.03	0.64±0.03	0.69±0.05	0.63±0.03	0.69±0.06	0.71±0.03
SmoothSubspace	0.44±0.05	0.37±0.05	0.46±0.05	0.51±0.05	0.56±0.03	0.58±0.07	0.64±0.07
SonyAIBORobotSurface1	0.65±0.04	0.67±0.03	0.57±0.03	0.73±0.04	0.76±0.03	0.71±0.04	0.78±0.05
SonyAIBORobotSurface2	0.72±0.03	0.65±0.03	0.68±0.03	0.75±0.02	0.78±0.02	0.77±0.01	0.81±0.02
StarLightCurves	0.78±0.01	0.57±0.01	0.90±0.01	0.91±0.01	0.84±0.01	0.91±0.01	0.93±0.00
Strawberry	0.60±0.02	0.68±0.03	0.67±0.02	0.80±0.02	0.76±0.03	0.77±0.03	0.79±0.02
SwedishLeaf	0.23±0.02	0.21±0.03	0.35±0.02	0.53±0.02	0.40±0.02	0.51±0.02	0.58±0.02
Symbols	0.61±0.02	0.51±0.02	0.59±0.02	0.83±0.03	0.82±0.01	0.87±0.01	0.92±0.02
SyntheticControl	0.51±0.03	0.24±0.03	0.34±0.03	0.43±0.02	0.56±0.02	0.65±0.03	0.69±0.02
ToeSegmentation1	0.57±0.04	0.56±0.05	0.56±0.03	0.63±0.05	0.63±0.06	0.60±0.06	0.61±0.06
ToeSegmentation2	0.62±0.06	0.75±0.07	0.87±0.05	0.87±0.06	0.68±0.04	0.80±0.05	0.81±0.05
Trace	0.49±0.10	0.56±0.04	0.57±0.06	0.64±0.04	0.54±0.06	0.65±0.09	0.64±0.08
TwoLeadECG	0.58±0.02	0.63±0.02	0.66±0.03	0.67±0.02	0.69±0.03	0.68±0.03	0.71±0.02
TwoPatterns	0.27±0.01	0.31±0.01	0.26±0.01	0.33±0.01	0.33±0.02	0.28±0.01	0.34±0.02
UMD	0.43±0.06	0.32±0.07	0.38±0.03	0.47±0.08	0.48±0.08	0.47±0.06	0.52±0.08
UWaveGestureLibraryAll	0.20±0.01	0.27±0.01	0.18±0.00	0.42±0.02	0.45±0.01	0.32±0.01	0.53±0.02
UWaveGestureLibraryX	0.22±0.02	0.23±0.01	0.22±0.01	0.44±0.02	0.39±0.01	0.41±0.02	0.53±0.02
UWaveGestureLibraryY	0.29±0.01	0.29±0.01	0.18±0.01	0.42±0.01	0.45±0.01	0.40±0.01	0.55±0.01
UWaveGestureLibraryZ	0.23±0.01	0.30±0.01	0.24±0.01	0.49±0.02	0.40±0.00	0.46±0.02	0.54±0.02
Wafer	0.96±0.01	0.97±0.00	0.95±0.00	1.00±0.00	0.99±0.00	0.99±0.00	0.99±0.00
Wine	0.50±0.10	0.46±0.15	0.67±0.08	0.67±0.12	0.53±0.08	0.58±0.12	0.61±0.09
WordSynonyms	0.18±0.04	0.17±0.01	0.21±0.03	0.31±0.03	0.21±0.03	0.28±0.04	0.30±0.03
Worms	0.44±0.05	0.53±0.04	0.61±0.07	0.67±0.06	0.48±0.05	0.57±0.04	0.61±0.04
WormsTwoClass	0.59±0.06	0.66±0.06	0.68±0.05	0.73±0.04	0.60±0.04	0.69±0.04	0.67±0.07
Yoga	0.60±0.01	0.61±0.01	0.62±0.02	0.77±0.01	0.72±0.02	0.74±0.02	0.78±0.01

Appendix R. Results for Multivariate Classification

Table A23. Metrics Results (Accuracy) for Multivariate Classification with Dyadic *GPF* Curvature.

Dataset	Sig	GPF_S	GPF_U	GPF_{SU}	GPF_S + Sig	GPF_U + Sig	GPF_{SU} + Sig
ArticularyWordRecognition	0.98±0.02	0.98±0.01	0.95±0.02	0.98±0.01	0.99±0.01	0.99±0.01	0.99±0.01
AtrialFibrillation	0.27±0.17	0.17±0.15	0.20±0.16	0.10±0.08	0.30±0.07	0.27±0.20	0.27±0.17
BasicMotions	1.00±0.00	1.00±0.00	0.99±0.03	1.00±0.00	1.00±0.00	1.00±0.00	1.00±0.00
ERing	0.94±0.02	0.94±0.03	0.88±0.05	0.94±0.05	0.96±0.03	0.96±0.02	0.96±0.03
HandMovementDirection	0.22±0.07	0.39±0.12	0.26±0.05	0.32±0.04	0.33±0.07	0.24±0.04	0.31±0.10
Handwriting	0.39±0.03	0.38±0.05	0.26±0.03	0.47±0.03	0.42±0.04	0.44±0.03	0.47±0.04
Libras	0.84±0.04	0.73±0.06	0.61±0.04	0.80±0.06	0.85±0.07	0.86±0.05	0.87±0.03
LSST	0.51±0.01	0.53±0.01	0.59±0.01	0.58±0.01	0.54±0.02	0.58±0.01	0.58±0.01
NATOPS	0.89±0.01	0.88±0.05	0.82±0.05	0.86±0.04	0.88±0.03	0.89±0.03	0.89±0.01
PenDigits	0.96±0.00	0.98±0.00	0.90±0.00	0.98±0.00	0.98±0.00	0.98±0.00	0.99±0.00
RacketSports	0.83±0.04	0.84±0.03	0.79±0.07	0.87±0.04	0.85±0.03	0.88±0.04	0.92±0.02
SelfRegulationSCP1	0.81±0.04	0.82±0.03	0.73±0.04	0.83±0.04	0.87±0.02	0.84±0.03	0.87±0.02
SelfRegulationSCP2	0.51±0.05	0.51±0.03	0.50±0.03	0.55±0.03	0.53±0.05	0.51±0.04	0.51±0.04
StandWalkJump	0.21±0.12	0.33±0.13	0.50±0.26	0.40±0.20	0.26±0.09	0.29±0.13	0.40±0.15
UWaveGestureLibrary	0.87±0.04	0.87±0.01	0.67±0.03	0.90±0.02	0.90±0.03	0.90±0.04	0.90±0.03

Table A24. Metrics Results (Accuracy) for Multivariate Classification with Dyadic *GPF* Chord + Curvature.

Dataset	Sig	GPF_S	GPF_U	GPF_{SU}	GPF_S + Sig	GPF_U + Sig	GPF_{SU} + Sig
ArticularyWordRecognition	0.98±0.02	0.98±0.01	0.95±0.02	0.98±0.00	0.99±0.01	0.99±0.01	0.99±0.00
AtrialFibrillation	0.27±0.17	0.13±0.07	0.10±0.13	0.07±0.08	0.23±0.08	0.33±0.18	0.20±0.16
BasicMotions	1.00±0.00	0.99±0.03	0.97±0.03	1.00±0.00	1.00±0.00	1.00±0.00	1.00±0.00
ERing	0.94±0.02	0.95±0.02	0.93±0.04	0.96±0.04	0.96±0.02	0.96±0.02	0.98±0.03
HandMovementDirection	0.21±0.07	0.36±0.05	0.27±0.04	0.38±0.09	0.30±0.06	0.27±0.06	0.30±0.02
Handwriting	0.39±0.03	0.40±0.02	0.28±0.02	0.46±0.06	0.43±0.04	0.45±0.03	0.46±0.03
Libras	0.84±0.04	0.83±0.06	0.79±0.05	0.86±0.05	0.87±0.06	0.91±0.05	0.91±0.05
LSST	0.51±0.01	0.52±0.01	0.59±0.01	0.58±0.01	0.54±0.01	0.57±0.01	0.58±0.01
NATOPS	0.89±0.01	0.87±0.05	0.81±0.04	0.86±0.03	0.91±0.02	0.89±0.03	0.90±0.01
PenDigits	0.96±0.00	0.98±0.00	0.97±0.00	0.99±0.00	0.99±0.00	0.99±0.00	0.99±0.00
RacketSports	0.83±0.04	0.85±0.02	0.81±0.07	0.88±0.02	0.87±0.03	0.88±0.04	0.92±0.03
SelfRegulationSCP1	0.81±0.04	0.85±0.04	0.73±0.04	0.83±0.03	0.86±0.01	0.84±0.04	0.86±0.01
SelfRegulationSCP2	0.51±0.05	0.52±0.03	0.50±0.02	0.51±0.07	0.51±0.04	0.51±0.05	0.51±0.03
StandWalkJump	0.21±0.12	0.43±0.23	0.53±0.26	0.46±0.25	0.29±0.13	0.35±0.23	0.40±0.15
UWaveGestureLibrary	0.87±0.04	0.87±0.02	0.74±0.03	0.91±0.04	0.89±0.02	0.91±0.03	0.92±0.03

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 March 2026 doi:10.20944/preprints202603.1778.v1

Appendix S. Results for Univariate Regression

Table A25. Metrics Results (MSE) for Univariate Regression with *GPF* Curvature.

Dataset	Sig	GPF_S	GPF_U
AcousticContaminationMadrid	61.53±7.76	58.15±13.14	57.67±7.70
AluminiumConcentration	309995.60±49535.15	240506.59±15635.05	256368.98±29717.59
BoronConcentration	8.00±9.45	8.38±9.24	8.15±8.94
CalciumConcentration	18645852.29±3671773.88	15108561.15±7462659.01	15717603.83±4999897.18
CopperConcentration	6.57±3.05	5.78±2.36	4.89±1.76
Covid19Andalusia	0.00±0.00	0.00±0.00	0.00±0.00
Covid3Month	0.00±0.00	0.00±0.00	0.00±0.00
DhakaHourlyAirQuality	4003.48±207.67	1150.35±95.42	1319.36±132.17
FloodModeling1	0.00±0.00	0.00±0.00	0.00±0.00
FloodModeling2	0.00±0.00	0.00±0.00	0.00±0.00
FloodModeling3	0.00±0.00	0.00±0.00	0.00±0.00
GasSensorArrayAcetone	0.12±0.02	0.10±0.01	0.09±0.02
GasSensorArrayEthanol	0.11±0.01	0.13±0.02	0.11±0.03
IronConcentration	9728.31±2075.66	8900.87±1684.51	8861.11±1896.00
LPGasMonitoringHomeActivity	2.40±0.18	2.16±0.22	2.32±0.28
MagnesiumConcentration	307714.20±83339.25	254205.25±41681.94	224462.66±34634.11
ManganeseConcentration	15136.06±1445.95	18141.32±2975.15	16000.21±3226.45
MethaneMonitoringHomeActivity	0.46±0.05	0.38±0.05	0.52±0.06
NaturalGasPricesSentiment	0.01±0.00	0.02±0.00	0.01±0.00
PhosphorusConcentration	1069.23±215.29	1169.82±151.10	1074.27±147.66
PotassiumConcentration	145944.46±36935.37	119946.88±23080.75	140356.99±32998.71
SodiumConcentration	460872.86±477180.29	373611.47±514993.81	628124.94±639590.67
SulphurConcentration	59741.07±41266.19	58320.28±55546.61	85389.66±62163.36
WaveDataTension	336.81±34.75	338.79±16.55	482.74±17.15
WindTurbinePower	744821.80±20001.26	643705.31±64257.65	479848.59±21275.23
ZincConcentration	3.38±1.56	3.31±1.13	3.56±1.94

Table A25. *Cont.*

GPF_{SU}	GPF_S + Sig	GPF_U + Sig	GPF_{SU} + Sig
46.71±10.29	46.00±4.83	48.91±3.26	41.38±5.68
188478.67±19630.62	211941.33±30980.30	209528.51±22801.16	186610.21±21679.53
8.54±8.73	7.26±9.13	8.15±8.79	8.32±8.67
14194447.42±3441879.70	13973874.09±5242397.03	15078684.24±3667271.16	13893351.61±3012127.68
4.23±2.32	4.60±2.38	4.19±2.15	4.11±2.23
0.00±0.00	0.00±0.00	0.00±0.00	0.00±0.00
0.00±0.00	0.00±0.00	0.00±0.00	0.00±0.00
833.84±65.70	831.80±78.28	977.77±66.68	613.65±45.95
0.00±0.00	0.00±0.00	0.00±0.00	0.00±0.00
0.00±0.00	0.00±0.00	0.00±0.00	0.00±0.00
0.00±0.00	0.00±0.00	0.00±0.00	0.00±0.00
0.07±0.01	0.08±0.01	0.08±0.02	0.07±0.01
0.07±0.02	0.08±0.01	0.09±0.01	0.07±0.01
7637.46±1513.36	7758.44±1489.99	7404.43±1391.89	7086.33±1375.56
1.83±0.20	1.73±0.21	1.94±0.23	1.66±0.18
171003.54±42813.21	213602.05±52726.09	185363.28±33153.48	160260.33±36491.56
12747.80±3265.09	12267.06±2504.44	11722.62±1886.37	11066.32±1673.43
0.30±0.04	0.30±0.05	0.40±0.07	0.29±0.05
0.01±0.00	0.01±0.00	0.01±0.00	0.01±0.00
874.73±196.29	884.61±197.30	892.17±212.70	831.89±210.04
99474.00±32574.01	104340.21±28313.56	108759.30±29781.00	97677.38±33795.37
412465.14±493428.42	418386.64±489014.29	449638.52±476201.18	411510.32±485618.20
69178.45±58335.95	49642.65±47368.83	65799.98±52450.85	62925.17±53581.86
286.32±18.67	282.18±19.04	278.58±21.19	264.66±19.08
365964.12±24882.55	454677.10±31785.39	359038.14±14910.82	315060.93±21163.35
2.82±1.04	2.87±1.06	2.71±1.47	2.63±1.00

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 March 2026 doi:10.20944/preprints202603.1778.v1

Table A26. Metrics Results for Univariate Regression with *GPF* Chord + Curvature.

Dataset	Sig	GPF_S	GPF_U
AcousticContaminationMadrid	61.53±7.76	43.97±6.64	48.03±2.68
AluminiumConcentration	309995.60±49535.15	189895.85±27669.82	206694.03±26888.40
BoronConcentration	8.00±9.45	8.12±8.80	7.41±8.85
CalciumConcentration	18645852.29±3671773.88	13884799.82±6123973.77	14258573.93±3371931.76
CopperConcentration	6.57±3.05	4.22±2.16	3.93±2.12
Covid19Andalusia	0.00±0.00	0.00±0.00	0.00±0.00
Covid3Month	0.00±0.00	0.00±0.00	0.00±0.00
DhakaHourlyAirQuality	4003.48±207.67	897.67±52.42	531.32±56.02
FloodModeling1	0.00±0.00	0.00±0.00	0.00±0.00
FloodModeling2	0.00±0.00	0.00±0.00	0.00±0.00
FloodModeling3	0.00±0.00	0.00±0.00	0.00±0.00
GasSensorArrayAcetone	0.12±0.02	0.07±0.00	0.08±0.02
GasSensorArrayEthanol	0.11±0.01	0.08±0.02	0.09±0.01
IronConcentration	9728.31±2075.66	7914.60±1403.02	6894.86±1485.91
LPGasMonitoringHomeActivity	2.40±0.18	1.78±0.17	1.91±0.20
MagnesiumConcentration	307714.20±83339.25	204874.80±45767.81	158886.43±31915.06
ManganeseConcentration	15136.06±1445.95	12719.85±3255.21	12309.46±3057.64
MethaneMonitoringHomeActivity	0.46±0.05	0.30±0.05	0.40±0.06
NaturalGasPricesSentiment	0.01±0.00	0.01±0.00	0.01±0.00
PhosphorusConcentration	1069.23±215.29	881.54±187.09	773.34±249.94
PotassiumConcentration	145944.46±36935.37	107086.06±25165.45	103516.85±30869.39
SodiumConcentration	460872.86±477180.29	491744.35±526949.25	658521.71±684219.15
SulphurConcentration	59741.07±41266.19	55710.56±55153.60	64804.50±51373.63
WaveDataTension	336.81±34.75	279.17±16.73	375.79±20.72
WindTurbinePower	744821.80±20001.26	505059.15±18069.67	342824.81±17796.66
ZincConcentration	3.38±1.56	2.92±0.98	2.85±1.81

Table A26. *Cont.*

GPF_{SU}	GPF_S + Sig	GPF_U + Sig	GPF_{SU} + Sig
36.69±5.46	42.34±4.27	48.99±3.66	37.23±4.31
172887.19±20287.55	181143.67±28653.83	191454.75±17586.37	163614.09±27781.64
8.30±8.53	7.69±8.88	7.75±8.76	8.33±8.43
12790269.15±3492922.08	13804504.47±5162619.56	14250320.95±3033006.52	13207212.84±2701305.59
3.74±2.26	4.19±2.10	3.88±2.32	3.81±2.34
0.00±0.00	0.00±0.00	0.00±0.00	0.00±0.00
0.00±0.00	0.00±0.00	0.00±0.00	0.00±0.00
321.31±30.47	705.17±70.80	462.40±66.69	282.78±31.15
0.00±0.00	0.00±0.00	0.00±0.00	0.00±0.00
0.00±0.00	0.00±0.00	0.00±0.00	0.00±0.00
0.00±0.00	0.00±0.00	0.00±0.00	0.00±0.00
0.07±0.01	0.07±0.01	0.08±0.02	0.07±0.01
0.07±0.01	0.07±0.01	0.08±0.01	0.07±0.01
6675.34±1359.05	7338.21±1317.72	6515.82±1500.23	6491.08±1469.70
1.64±0.14	1.66±0.18	1.81±0.21	1.65±0.18
121585.38±35750.35	185156.06±45242.64	135509.80±33955.13	122210.18±36933.47
11213.92±2748.90	10945.49±2061.44	10489.46±1859.22	10124.64±1755.86
0.27±0.04	0.28±0.04	0.38±0.07	0.27±0.04
0.01±0.00	0.01±0.00	0.01±0.00	0.01±0.00
637.66±220.61	850.42±207.56	675.97±227.17	640.42±225.52
79401.36±19039.48	101251.39±27271.66	89657.80±27918.17	81439.21±20964.85
481635.21±523095.09	485883.90±516550.24	448782.36±476282.94	471817.14±511647.58
64528.64±54352.36	52109.00±49225.76	64156.21±51227.36	61617.08±52027.64
262.22±16.41	267.81±20.16	264.69±19.17	253.44±16.66
267873.38±26752.13	425998.25±37554.78	300269.56±15811.40	254201.11±33932.73
2.71±1.10	2.85±1.07	2.64±1.60	2.66±1.12

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 March 2026 doi:10.20944/preprints202603.1778.v1

Appendix T. Results for Multivariate Regression

Table A27. Metrics Results (MSE) for Multivariate Regression with *GPF* Curvature.

Dataset	Sig	GPF_S	GPF_U
BarCrawl6min	0.00±0.00	0.00±0.00	0.00±0.00
BeijingIntAirportPM25Quality	4137.23±759.66	3617.07±539.14	3376.95±332.18
BinanceCoinSentiment	0.14±0.02	0.14±0.03	0.15±0.01
BitcoinSentiment	0.06±0.01	0.04±0.01	0.04±0.01
CardanoSentiment	0.09±0.02	0.10±0.03	0.10±0.03
ChilledWaterPredictor	8726299.78±4380433.00	7901263.98±3798174.62	6672556.43±2854366.40
DailyOilGasPrices	4.12±1.41	2.76±1.03	2.25±1.08
ElectricityPredictor	301987.00±32066.41	296333.83±41218.81	292771.64±40383.20
EthereumSentiment	0.07±0.02	0.06±0.01	0.06±0.01
HotwaterPredictor	2135863.82±593013.23	1698336.67±411117.58	1882371.68±406071.83
HouseholdPowerConsumption1	174093.58±28516.12	53586.38±10595.83	36556.00±9206.76
HouseholdPowerConsumption2	1854.81±142.80	1602.07±173.07	1098.84±139.44
IEEEPPG	247.17±27.68	147.91±27.14	119.06±5.64
MetroInterstateTrafficVolume	340140.08±23292.80	349592.90±22304.57	348590.22±22488.21
OccupancyDetectionLight	16674.43±2575.90	8810.62±3678.44	13963.22±3538.52
PrecipitationAndalusia	0.22±0.06	0.29±0.07	0.30±0.06
SierraNevadaMountainsSnow	5193.89±1494.43	5455.08±1621.92	2056.62±221.00
SolarRadiationAndalusia	0.54±0.07	0.64±0.07	0.67±0.08
SteamPredictor	1400871699.82±566068401.34	1189162666.70±426571113.46	1113319004.69±501814009.80

Table A27. *Cont.*

GPF_{SU}	GPF_S + Sig	GPF_U + Sig	GPF_{SU} + Sig
0.00±0.00	0.00±0.00	0.00±0.00	0.00±0.00
3273.90±432.18	3298.42±528.46	3222.79±457.13	3153.23±427.94
0.14±0.02	0.13±0.02	0.14±0.02	0.14±0.02
0.04±0.01	0.04±0.01	0.04±0.01	0.04±0.01
0.09±0.03	0.09±0.02	0.09±0.02	0.09±0.03
7083638.44±2941651.61	8408185.20±4251227.45	7638476.35±3300385.87	7742217.85±3496750.71
2.02±0.94	3.25±1.17	2.88±1.26	2.66±1.17
288927.24±41791.03	297763.95±33607.53	295018.90±40269.94	296475.92±41186.33
0.06±0.01	0.06±0.01	0.05±0.01	0.05±0.01
1677707.46±456419.50	1816073.00±531828.44	2003807.37±562254.12	1773567.68±521277.24
37606.37±9174.06	59899.22±12934.36	43371.93±10404.73	43939.62±10445.89
1077.43±133.53	1664.40±141.52	1151.28±147.76	1154.45±143.76
105.42±16.25	205.24±30.46	158.31±15.21	147.64±15.67
346661.05±19836.69	337478.25±16205.86	334002.42±23051.01	337620.17±20412.37
7563.00±2316.72	6973.80±4102.29	11945.34±1316.76	6129.86±3017.11
0.28±0.06	0.21±0.06	0.21±0.06	0.21±0.06
2059.22±195.56	5083.23±1667.61	2010.17±359.55	2069.17±315.52
0.62±0.06	0.53±0.07	0.53±0.08	0.52±0.07
1078004825.38±415606504.90	1404054079.71±538211660.54	1339283187.38±546121713.29	1359680850.98±569389589.20

Table A28. Metrics Results (MSE) for Multivariate Regression with *GPF* Chord + Curvature.

Dataset	Sig	GPF_S	GPF_U
BarCrawl6min	0.00±0.00	0.00±0.00	0.00±0.00
BeijingIntAirportPM25Quality	4137.23±759.66	3470.39±512.87	3333.05±319.45
BinanceCoinSentiment	0.14±0.02	0.13±0.02	0.13±0.02
BitcoinSentiment	0.06±0.01	0.04±0.01	0.05±0.01
CardanoSentiment	0.09±0.02	0.09±0.03	0.10±0.03
ChilledWaterPredictor	8726299.78±4380433.00	7964258.76±3595897.44	6859144.40±2912898.04
DailyOilGasPrices	4.12±1.41	2.95±1.23	2.13±0.95
ElectricityPredictor	301987.00±32066.41	297632.97±39418.34	290666.84±41458.12
EthereumSentiment	0.07±0.02	0.06±0.01	0.06±0.01
HotwaterPredictor	2135863.82±593013.23	1703706.28±398277.83	1820447.96±388466.61
HouseholdPowerConsumption1	174093.58±28516.12	52682.22±10137.04	33336.46±7811.10
HouseholdPowerConsumption2	1854.81±142.80	1610.67±191.74	1036.85±130.29
IEEEPPG	247.17±27.68	154.26±27.38	125.50±7.47
MetroInterstateTrafficVolume	340140.08±23292.80	334650.51±24923.92	333286.29±19869.22
OccupancyDetectionLight	16674.43±2575.90	9792.82±4200.65	13921.11±3585.51
PrecipitationAndalusia	0.22±0.06	0.21±0.05	0.25±0.04
SierraNevadaMountainsSnow	5193.89±1494.43	5157.70±1492.28	2100.60±221.92
SolarRadiationAndalusia	0.54±0.07	0.56±0.07	0.63±0.08
SteamPredictor	1400871699.82±566068401.34	1176636721.12±419559963.88	968691409.99±452249190.89

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 March 2026 doi:10.20944/preprints202603.1778.v1

Table A28. *Cont.*

GPF_{SU}	GPF_S + Sig	GPF_U + Sig	GPF_{SU} + Sig
0.00±0.00	0.00±0.00	0.00±0.00	0.00±0.00
3247.92±394.15	3302.91±511.56	3197.44±448.61	3156.88±421.65
0.12±0.02	0.13±0.02	0.13±0.02	0.13±0.02
0.04±0.01	0.04±0.01	0.04±0.01	0.04±0.01
0.09±0.03	0.10±0.02	0.09±0.02	0.09±0.03
7365422.08±3092870.85	8489141.01±4322332.99	7683509.70±3249433.37	7675844.37±3312634.94
2.12±0.96	3.37±1.14	2.74±1.16	2.67±1.19
289757.98±45353.91	299386.14±34144.93	293298.25±39135.66	295516.79±43615.59
0.06±0.01	0.05±0.01	0.05±0.01	0.06±0.01
1560768.19±461405.25	1797093.67±546039.82	1982165.43±566396.34	1702877.05±539886.29
33944.51±7002.41	58310.44±12110.54	39077.37±9080.06	39484.04±9571.27
1037.69±122.11	1661.87±143.79	1131.47±158.22	1121.46±154.55
112.40±18.01	205.73±29.23	159.44±15.15	149.15±17.73
332427.81±18552.33	338684.03±20321.36	333351.60±19963.89	334575.83±18190.00
8112.74±2758.04	7062.25±4096.78	11871.93±1263.34	6041.98±2646.59
0.20±0.04	0.21±0.06	0.21±0.06	0.21±0.06
2047.05±170.58	5080.67±1595.60	2037.00±351.52	2060.31±370.45
0.53±0.04	0.53±0.06	0.52±0.07	0.51±0.06
1015881030.78±498422057.10	1403061555.46±539163892.55	1296234073.01±536802371.45	1284800955.29±538588396.78

Table A29. Mean ± std accuracy (5-fold CV) on 15 UEA datasets across all baseline methods.

Dataset	Catch22	ROCKET	MiniROCKET	MultiROCKET	HYDRA	TSFresh	Shapelet	RDST	Interval	BOSS
AWR	0.97 ± 0.01	0.98 ± 0.01	0.98 ± 0.01	0.99 ± 0.01	0.98 ± 0.01	0.45 ± 0.44	0.98 ± 0.01	0.99 ± 0.01	0.91 ± 0.03	0.97 ± 0.01
AtrFib	0.23 ± 0.17	0.23 ± 0.08	0.13 ± 0.12	0.17 ± 0.11	0.10 ± 0.08	0.20 ± 0.16	0.10 ± 0.08	0.13 ± 0.12	0.23 ± 0.17	0.20 ± 0.12
BasicMot	1.00 ± 0.00	0.99 ± 0.03	1.00 ± 0.00	1.00 ± 0.00	1.00 ± 0.00	0.95 ± 0.10	0.99 ± 0.03	1.00 ± 0.00	1.00 ± 0.00	0.95 ± 0.05
ERing	0.98 ± 0.02	0.99 ± 0.02	0.99 ± 0.02	0.99 ± 0.01	0.99 ± 0.01	0.99 ± 0.01	0.98 ± 0.02	0.98 ± 0.01	0.94 ± 0.04	0.99 ± 0.01
HandMov	0.22 ± 0.05	0.30 ± 0.03	0.29 ± 0.06	0.29 ± 0.05	0.28 ± 0.03	0.33 ± 0.04	0.40 ± 0.04	0.35 ± 0.02	0.37 ± 0.05	0.27 ± 0.03
Handwrt	0.47 ± 0.06	0.67 ± 0.05	0.62 ± 0.04	0.64 ± 0.04	0.61 ± 0.04	0.62 ± 0.04	0.68 ± 0.04	0.75 ± 0.02	0.51 ± 0.02	0.69 ± 0.02
LSST	0.68 ± 0.02	0.60 ± 0.02	0.63 ± 0.02	0.64 ± 0.02	0.63 ± 0.01	0.51 ± 0.15	0.61 ± 0.01	0.63 ± 0.01	0.54 ± 0.01	0.46 ± 0.01
Libras	0.92 ± 0.05	0.91 ± 0.03	0.93 ± 0.04	0.93 ± 0.02	0.94 ± 0.01	0.48 ± 0.40	0.88 ± 0.05	0.93 ± 0.02	0.81 ± 0.06	0.88 ± 0.05
NATOPS	0.89 ± 0.04	0.86 ± 0.04	0.87 ± 0.04	0.83 ± 0.05	0.84 ± 0.04	0.63 ± 0.28	0.82 ± 0.03	0.86 ± 0.05	0.75 ± 0.05	0.81 ± 0.06
PenDig	0.98 ± 0.00	0.99 ± 0.00	0.99 ± 0.00	0.99 ± 0.00	0.99 ± 0.00	0.99 ± 0.00	0.99 ± 0.00	0.99 ± 0.00	0.86 ± 0.01	0.97 ± 0.01
Racket	0.88 ± 0.02	0.90 ± 0.02	0.88 ± 0.06	0.86 ± 0.04	0.85 ± 0.04	0.70 ± 0.24	0.87 ± 0.03	0.89 ± 0.03	0.82 ± 0.05	0.86 ± 0.04
SCP1	0.85 ± 0.03	0.85 ± 0.02	0.88 ± 0.03	0.89 ± 0.03	0.87 ± 0.05	0.87 ± 0.03	0.84 ± 0.02	0.88 ± 0.02	0.83 ± 0.02	0.71 ± 0.04
SCP2	0.51 ± 0.05	0.47 ± 0.04	0.49 ± 0.03	0.51 ± 0.03	0.48 ± 0.04	0.50 ± 0.07	0.49 ± 0.04	0.51 ± 0.06	0.51 ± 0.04	0.51 ± 0.04
SWJ	0.40 ± 0.11	0.37 ± 0.03	0.36 ± 0.14	0.54 ± 0.22	0.37 ± 0.16	0.23 ± 0.09	0.29 ± 0.13	0.43 ± 0.24	0.21 ± 0.12	0.19 ± 0.13
UWave	0.92 ± 0.03	0.94 ± 0.02	0.92 ± 0.03	0.94 ± 0.02	0.93 ± 0.02	0.45 ± 0.24	0.93 ± 0.02	0.95 ± 0.02	0.90 ± 0.03	0.80 ± 0.02

Table A30. Mean ± std accuracy (5-fold CV) on 15 UEA datasets across *GPF* variants.

Dataset	Sig	GPF_S	GPF_U	GPF_{SU}	GPF_S+Sig	GPF_U+Sig	GPF_{SU}+Sig
AWR	0.98 ± 0.01	0.99 ± 0.01	0.97 ± 0.01	1.00 ± 0.00	0.98 ± 0.01	0.98 ± 0.01	0.99 ± 0.01
AtrFib	0.40 ± 0.17	0.30 ± 0.16	0.27 ± 0.17	0.23 ± 0.17	0.37 ± 0.16	0.30 ± 0.19	0.33 ± 0.15
BasicMot	1.00 ± 0.00	1.00 ± 0.00	1.00 ± 0.00	1.00 ± 0.00	1.00 ± 0.00	1.00 ± 0.00	1.00 ± 0.00
ERing	0.97 ± 0.01	0.98 ± 0.02	0.94 ± 0.03	0.98 ± 0.01	0.97 ± 0.01	0.98 ± 0.01	0.97 ± 0.01
HandMov	0.30 ± 0.04	0.29 ± 0.06	0.27 ± 0.04	0.28 ± 0.04	0.30 ± 0.07	0.27 ± 0.02	0.32 ± 0.03
Handwrt	0.59 ± 0.04	0.60 ± 0.03	0.39 ± 0.02	0.61 ± 0.02	0.59 ± 0.03	0.62 ± 0.04	0.61 ± 0.04
LSST	0.62 ± 0.02	0.62 ± 0.01	0.63 ± 0.01	0.66 ± 0.02	0.62 ± 0.01	0.65 ± 0.02	0.65 ± 0.02
Libras	0.90 ± 0.05	0.86 ± 0.03	0.91 ± 0.03	0.90 ± 0.02	0.90 ± 0.05	0.92 ± 0.04	0.92 ± 0.03
NATOPS	0.91 ± 0.04	0.91 ± 0.05	0.89 ± 0.04	0.91 ± 0.04	0.93 ± 0.04	0.93 ± 0.04	0.92 ± 0.04
PenDig	0.98 ± 0.00	0.98 ± 0.00	0.98 ± 0.00	0.99 ± 0.00	0.98 ± 0.00	0.99 ± 0.00	0.99 ± 0.00
Racket	0.85 ± 0.05	0.85 ± 0.05	0.82 ± 0.03	0.86 ± 0.06	0.87 ± 0.05	0.87 ± 0.05	0.86 ± 0.05
SCP1	0.87 ± 0.03	0.86 ± 0.02	0.78 ± 0.02	0.89 ± 0.03	0.87 ± 0.03	0.89 ± 0.02	0.90 ± 0.02
SCP2	0.53 ± 0.03	0.51 ± 0.03	0.51 ± 0.06	0.53 ± 0.04	0.52 ± 0.02	0.52 ± 0.05	0.53 ± 0.03
SWJ	0.11 ± 0.09	0.21 ± 0.12	0.27 ± 0.27	0.25 ± 0.17	0.22 ± 0.06	0.22 ± 0.06	0.29 ± 0.12
UWave	0.92 ± 0.03	0.92 ± 0.04	0.69 ± 0.05	0.95 ± 0.03	0.92 ± 0.03	0.94 ± 0.03	0.94 ± 0.03

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 March 2026 doi:10.20944/preprints202603.1778.v1

Table A31. Feature counts for all methods across 15 UEA datasets. For *GPF* variants, feature counts are before RF selection. C22 = Catch22, RKT = ROCKET, mRKT = MiniROCKET, mRKT+ = MultiROCKET, HYD = HYDRA, TSF = TSFresh, Shp = Shapelet, Int = Interval.

Dataset	C22	RKT	mRKT	mRKT+	HYD	TSF	Shp	RDST	Int	BOSS	Sig	GPF$_S$	GPF$_U$	GPF$_{SU}$	GPF$_S$+Sig	GPF$_U$+Sig	GPF$_{SU}$+Sig
AWR	216	2,000	924	7,392	5,120	6,993	981	30,000	84	202,923	5,256	2,556	2,556	5,112	7,812	7,812	10,368
AtrFib	48	2,000	924	7,392	7,168	1,554	139	30,000	175	24,189	272	120	120	240	392	392	512
BasicMot	144	2,000	924	7,392	4,096	4,662	580	20,121	70	13,315	2,352	1,128	1,128	2,256	3,480	3,480	4,608
ERing	96	2,000	924	7,392	4,096	3,108	712	25,464	56	637	1,056	496	496	992	1,552	1,552	2,048
HandMov	240	2,000	924	7,392	6,144	7,770	990	30,000	140	291,790	6,480	3,160	3,160	6,320	9,640	9,640	12,800
Handwrt	72	2,000	924	7,392	5,120	2,331	593	30,000	84	75,553	600	276	276	552	876	876	1,152
LSST	144	2,000	924	7,392	3,072	4,662	941	30,000	42	960	2,352	1,128	1,128	2,256	3,480	3,480	4,608
Libras	48	2,000	924	7,392	3,072	1,554	691	16,563	42	320	272	120	120	240	392	392	512
NATOPS	576	2,000	924	7,392	3,072	18,648	986	21,774	49	3,804	37,056	18,336	18,336	36,672	55,392	55,392	73,728
PenDig	48	2,000	924	7,392	1,024	1,554	988	30,000	14	154	272	120	120	240	392	392	512
Racket	144	2,000	924	7,392	2,048	4,662	963	4,434	35	960	2,352	1,128	1,128	2,256	3,480	3,480	4,608
SCP1	144	2,000	924	7,392	7,168	4,662	994	30,000	203	202,445	2,352	1,128	1,128	2,256	3,480	3,480	4,608
SCP2	168	2,000	924	7,392	8,192	5,439	998	30,000	231	216,243	3,192	1,540	1,540	3,080	4,732	4,732	6,272
SWJ	96	2,000	924	7,392	9,216	3,108	168	30,000	350	43,337	1,056	496	496	992	1,552	1,552	2,048
UWave	72	2,000	924	7,392	6,144	2,331	887	30,000	119	75,222	600	276	276	552	876	876	1,152

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 March 2026 doi:10.20944/preprints202603.1778.v1

References

1. Morrill, J.; Fermanian, A.; Kidger, P.; Lyons, T. A Generalised Signature Method for Multivariate Time Series Feature Extraction. 2021.
2. Christ, M.; Braun, N.; Neuffer, J.; Kempa-Liehr, A.W. Time series feature extraction on basis of scalable hypothesis tests (tsfresh–a python package). *Neurocomputing* **2018**, *307*, 72–77.
3. Lubba, C.H.; Sethi, S.S.; Knaute, P.; Schultz, S.R.; Fulcher, B.D.; Jones, N.S. catch22: CAnonical Time-series CHaracteristics, 2019, [arXiv:cs.IR/1901.10200].
4. Dempster, A.; Schmidt, D.F.; Webb, G.I. Quant: A minimalist interval method for time series classification. *Data Mining and Knowledge Discovery* **2024**, *38*, 2377–2402.
5. Middlehurst, M.; Large, J.; Bagnall, A. The canonical interval forest (CIF) classifier for time series classification. In Proceedings of the 2020 IEEE international conference on big data (big data). IEEE, 2020, pp. 188–195.
6. Schäfer, P.; Leser, U. WEASEL 2.0: a random dilated dictionary transform for fast, accurate and memory constrained time series classification. *Machine Learning* **2023**, *112*, 4763–4788.
7. Schäfer, P. The BOSS is concerned with time series classification in the presence of noise. *Data Mining and Knowledge Discovery* **2015**, *29*, 1505–1530.
8. Hills, J.; Lines, J.; Baranauskas, E.; Mapp, J.; Bagnall, A. Classification of time series by shapelet transformation. *Data mining and knowledge discovery* **2014**, *28*, 851–881.
9. Guillaume, A.; Vrain, C.; Elloumi, W. Random dilated shapelet transform: A new approach for time series shapelets. In Proceedings of the International Conference on Pattern Recognition and Artificial Intelligence. Springer, 2022, pp. 653–664.
10. Karlsson, I.; Papapetrou, P.; Boström, H. Generalized random shapelet forests. *Data mining and knowledge discovery* **2016**, *30*, 1053–1085.
11. Dempster, A.; Schmidt, D.F.; Webb, G.I. Hydra: Competing convolutional kernels for fast and accurate time series classification. *Data Mining and Knowledge Discovery* **2023**, *37*, 1779–1805.
12. Dempster, A.; Petitjean, F.; Webb, G.I. ROCKET: exceptionally fast and accurate time series classification using random convolutional kernels. *Data Mining and Knowledge Discovery* **2020**, *34*, 1454–1495. https://doi.org/10.1007/s10618-020-00701-z.
13. Dempster, A.; Schmidt, D.F.; Webb, G.I. MiniRocket: A Very Fast (Almost) Deterministic Transform for Time Series Classification. In Proceedings of the Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery. ACM, 2021, KDD '21, p. 248–257. https://doi.org/10.1145/3447548.3467231.
14. Tan, C.; Dempster, A.; Bergmeir, C.; Webb, G. MultiRocket: multiple pooling operators and transformations for fast and effective time series classification. *Data Mining and Knowledge Discovery* **2022**, *36*, 1623–1646. https://doi.org/10.1007/s10618-022-00844-1.
15. Ismail Fawaz, H.; Lucas, B.; Forestier, G.; Pelletier, C.; Schmidt, D.F.; Weber, J.; Webb, G.I.; Idoumghar, L.; Muller, P.A.; Petitjean, F. Inceptiontime: Finding alexnet for time series classification. *Data Mining and Knowledge Discovery* **2020**, *34*, 1936–1962.
16. Schäfer, P.; Leser, U. Multivariate time series classification with WEASEL+ MUSE. *arXiv preprint arXiv:1711.11343* **2017**.
17. Zhang, X.; Gao, Y.; Lin, J.; Lu, C.T. Tapnet: Multivariate time series classification with attentional prototypical network. In Proceedings of the Proceedings of the AAAI conference on artificial intelligence, 2020, Vol. 34, pp. 6845–6852.
18. Karim, F.; Majumdar, S.; Darabi, H.; Harford, S. Multivariate LSTM-FCNs for time series classification. *Neural networks* **2019**, *116*, 237–245.
19. Chen, K.T. Integration of paths–A faithful representation of paths by noncommutative formal power series. *Transactions of the American Mathematical Society* **1958**, *89*, 395–407.
20. Lyons, T.J.; Caruana, M.; Lévy, T. *Differential equations driven by rough paths*; Springer, 2007.
21. Friz, P.K.; Victoir, N.B. *Multidimensional stochastic processes as rough paths: theory and applications*; Vol. 120, Cambridge University Press, 2010.
22. Moore, P.; Lyons, T.; Gallacher, J.; Initiative, A.D.N. Using path signatures to predict a diagnosis of Alzheimer's disease. *PloS one* **2019**, *14*, e0222212.
23. Yang, W.; Jin, L.; Liu, M. Deepwriterid: An end-to-end online text-independent writer identification system. *IEEE Intelligent Systems* **2016**, *31*, 45–53.
24. Ibrahim, M.R.; Lyons, T. ImageSig: A signature transform for ultra-lightweight image recognition. In Proceedings of the Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022, pp. 3649–3659.

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 March 2026 doi:10.20944/preprints202603.1778.v1

25. Ni, H.; Szpruch, L.; Sabate-Vidales, M.; Xiao, B.; Wiese, M.; Liao, S. Sig-Wasserstein GANs for time series generation. In Proceedings of the Proceedings of the Second ACM International Conference on AI in Finance, 2021, pp. 1–8.
26. Levin, D.; Lyons, T.; Ni, H. Learning from the past, predicting the statistics for the future, learning an evolving system. *arXiv preprint arXiv:1309.0260* **2013**.
27. Le Jan, Y.; Qian, Z. Stratonovich's signatures of Brownian motion determine Brownian sample paths. *Probability Theory and Related Fields* **2013**, *157*, 209–223.
28. Hambly, B.; Lyons, T. Uniqueness for the signature of a path of bounded variation and the reduced path group. *Annals of Mathematics* **2010**, pp. 109–167.
29. Cass, T.; Salvi, C. Lecture notes on rough paths and applications to machine learning. *arXiv preprint arXiv:2404.06583* **2024**.
30. Brahim Belhaouari, S.; Kahalan, Y.C.; Aksikas, I.; Hamdi, A.; Belhaouari, I.; Haoudi, E.N.; Bensmail, H. Generalizing the Cross Product to N Dimensions: A Novel Approach for Multidimensional Analysis and Applications. *Mathematics (2227-7390)* **2025**, *13*.
31. Chevyrev, I.; Kormilitzin, A. A primer on the signature method in machine learning. *arXiv preprint arXiv:1603.03788* **2016**.
32. Demšar, J. Statistical comparisons of classifiers over multiple data sets. *The Journal of Machine learning research* **2006**, *7*, 1–30.
33. Bagnall, A.; Lines, J.; Bostrom, A.; Large, J.; Keogh, E. The great time series classification bake off: a review and experimental evaluation of recent algorithmic advances. *Data mining and knowledge discovery* **2017**, *31*, 606–660.
34. Chen, K.T. Integration of paths–A faithful representation of paths by noncommutative formal power series. *Transactions of the American Mathematical Society* **1958**, *89*, 395–407.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.