

m-ark-Support Vector Machine for Early Detection of Parkinson's Disease from Speech Signals

Luca Parisi,

Department of Computer Science,
Faculty of Engineering and Informatics,
University of Bradford
Bradford, BD7 1DP, United Kingdom
University of Auckland Rehabilitative Technologies
Association (UARTA)
University of Auckland
Auckland, 1010, New Zealand

Renfei Ma,

University of Auckland Rehabilitative Technologies
Association (UARTA)
University of Auckland
Auckland, 1010, New Zealand

Amir Zaernia,

Department of Biomedical and Electronics Engineering,
Faculty of Engineering and Informatics,
University of Bradford
Bradford, BD7 1DP, United Kingdom

Mansour Youseffi,

Department of Biomedical and Electronics Engineering,
Faculty of Engineering and Informatics,
University of Bradford
Bradford, BD7 1DP, United Kingdom

Received: January 20, 2021. Revised: March 26, 2021. Accepted: March 30, 2021. Published: April 5, 2021.

Abstract—Recent advances in the state-of-the-art open-source kernel functions for support vector machines (SVMs) have widened the choices of benchmark kernels for Machine Learning (ML)-based classification. However, it is still challenging to achieve margin maximisation in SVM, and further evidence is required to ensure such novel kernel functions can have translational applications with tangible impact. Noteworthy, m -arcsinh, freely available in scikit-learn, was preliminarily proven as a benchmark kernel function on 15 datasets in its seminal paper. Quantifying the benefit from leveraging this kernel in a specific application is essential to provide further evidence of its accuracy and reliability on real-life supervised ML-aided tasks. Thus, the predictive capability of SVM, including that with Lagrange multipliers for the first time coupled with m -arcsinh (m -ark-SVM with soft margin; m -arK-SVM with hard margin), is hereby assessed in aiding early detection of Parkinson's Disease (PD) from speech data. This is important to leverage the m -arcsinh kernel 'trick' to maximise the margin width and, therefore, the linear separability of input speech features via automated pattern recognition. In this study, we demonstrate the accuracy and reliability of m -ark-SVM to aid early diagnosis of PD, evaluated against other gold standard kernel functions. Two benchmark datasets from the University of California-Irvine (UCI) database, pre-processed solely via min-max normalisation, were used to discriminate between speech patterns of 72 healthy subjects and 211 patients with PD. Overtraining was avoided via cross validation and the models were developed and tested in Python 3.7. The supervised model (m -ark-SVM) could detect early Parkinson's Disease with 87.18% and 86.9% classification accuracy from the two datasets respectively (F1-

scores: 85 and 86.2% correspondingly). Furthermore, the model achieved high precision (89.2% and 86.8%) and specificity (87% and 86.8%). Thus, this study validates the application of m -arcsinh to aid real-life supervised ML-based classification, in particular early diagnosis of Parkinson's Disease from speech data.

Keywords: Support Vector Machine; Machine Learning; Parkinson's Disease; kernel function; m -arcsinh; classification.

I. INTRODUCTION

Parkinson's disease (PD) is a neurodegenerative disorder related to loss of dopamine-producing neurons in the basal ganglia. The primary symptoms of PD include tremor, muscular rigidity, bradykinesia, and postural instability with secondary symptoms, such as depression, executive dysfunctions, sleep disturbances, and autonomic impairments.

PD is the second most common neurodegenerative disease in the elderly, after Alzheimer's Disease, which affects approximately 2% of the people at the age of 65 and accounts for more than 6.3 million cases worldwide [1]. Up to 90% of patients with PD develop distinctive speech anomalies, referred to as hypokinetic dysarthria [2].

Common speech impairments resulting from PD include increased acoustic noise, reduced voice intensity and increased nasality, harsh and breathy vocal quality, mono-pitch, mono-loudness, and speech rate disturbances [3]. Previous research leveraged such vocal disorders to identify late-stage PD [4].

However, currently, there is no consensus on which biomarkers can detect PD, and brain imaging scans do not allow for a definitive diagnosis either, resulting in a misdiagnosis rate of up to 25% by non-specialists [5]. Whilst there are no current cures for PD, early detection can significantly improve the quality of life of a patient. At the point of diagnosis, there is a 50 to 70% reduction in striatal dopamine. An early and timely detection of PD would potentially allow for treatment to slow the loss of dopamine and delay the symptoms due to PD being a progressive disease [6].

Typical pre-clinical studies on detection of PD rely on various signal processing techniques including signals via electromyography (EMG), movement analysis and cardiovascular oscillations [7]. Recent research has demonstrated the potential for leveraging Machine Learning (ML) to aid an early diagnosis of PD from speech patterns via feature-driven approaches [8].

The development of the open-source kernel function modified arcsinh or ‘m-arcsinh’ has provided a further benchmark kernel for ML-aided classification [9]. When coupled with the ‘Support Vector Machine’ (SVM) [10] and the shallow neural network ‘Multi-Layer Perceptron’ (MLP) [11], its reliability and computational efficiency have been proven on fifteen benchmark datasets [9]. Nevertheless, besides its reproducibility, it is essential that further studies can verify its replicability across different applications. In fact, at times, SVM struggle to find support vectors and optimal separating hyperplane (OSH) that maximize the margin width, thus lacking convergence [12, 13].

In such conditions, failure of SVM to converge could thus result in misclassifications, which may lead to misdiagnoses when applied in a clinical setting. Convergence can be aided by selecting an optimal kernel function to transform data into a higher dimensional space wherein they can be linearly separable [8, 14, 15]. As the traditional quadratic optimisation involved in SVM has been useful in presence of datasets whose features are not highly overlapping [16], the m-arcsinh kernel, able to cope with non-linearly separable data too, has been suggested to be coupled with SVM [9] with and without Lagrange

multipliers [17] to aid complex medical diagnostics, such as detecting early Parkinson’s Disease from speech data.

This study’s novelty lies in assessing the replicability of the m-arcsinh in such a clinical application, along with discerning the type of margin and the need for Lagrange multipliers in presence of such entangled medical datasets. In this research work, supervised ML techniques were leveraged to perform a classification task that involves entangled data with overlapping features, speech data with indicators of healthy and PD-related speech patterns [8, 14], to aid diagnosis of PD at early stages from their underlying vocal patterns.

II. AUTOMATED DETECTION OF PARKINSON’S DISEASE

A. Data and Software for Modelling

To demonstrate the accuracy and reliability of early detection of PD from speech patterns via the proposed algorithm, two datasets with speech features from 72 healthy subjects and 211 patients with PD were obtained from the University of California-Irvine (UCI) database [18, 19]. Both datasets underwent min-max normalisation to prepare them for the subsequent SVM-based classification. The SVM models were developed and evaluated via the scikit-learn library [20] in Python, including the m-arcsinh kernel function [9] compatible with scikit-learn and freely available on GitHub (https://github.com/luca-parisi/m-arcsinh_scikit-learn), and the CVXOPT library [21] for constrained convex optimisation (<https://cvxopt.org/>).

B. m-arcsinh-based SVM Algorithm

SVM is an OSH-based classifier that leverages the delineation of ‘anchor’ points or support vectors to achieve separability of the input data depending on the kernel function that is deployed to transform such data into a higher dimensional space to facilitate their classification [22], thus enabling automated learning based on input data [8, 14, 23]. These input data must be labelled, since the SVM is a supervised ML-based learning classifier, thus the class which each input belongs to is preliminarily known.

The choice of whether the margin should be soft or hard and whether leveraging Lagrange multipliers impacts the accuracy of the classification of such input data into the target classes [17, 24]. Thus, in this study, SVM was assessed with various kernel functions, including the recently developed m-arcsinh [9], types of margin and in presence or absence of Lagrange multipliers.

C. Margins, Lagrange Multipliers, and other kernels

The regularisation parameter ‘C’ determines whether the margin criterion is soft or hard depending on its values, which has an impact on reaching convergence and determining the OSH [15, 24]. As SVM involves solving a constrained convex quadratic optimisation problem, various optimisation methods, such as the Wolfe Dual Principle under Karush Kuhn Tucker (KKT) conditions [17, 22], have been suggested for the objective function to solely depend on Lagrange multipliers, i.e., weights and intercept dropout. Such a convex problem is thus solved by leveraging Slater’s condition [21, 22] whereby a strong duality holds in that the dual and primal optima provide the solution.

To avoid overtraining or overfitting, thus preventing the model from ‘remembering’ the features seen during training and enabling learning that yields generalisation, stratified k-fold cross validation was applied with k=5 [25, 26].

Given its ability to maximise margin width and cope with non-linearly separable data, the m-arcsinh kernel function was coupled with SVM (**Fig. 1**) and evaluated against other gold standard kernels (linear, polynomial and Radial Basis Function (RBF)) [27], with and without Lagrange multipliers, as per the following types of margin:

- Soft margin: C=100 for both linear and m-arcsinh kernels (m-ark).
- Hard margin: C=10 for both linear and m-arcsinh kernels (m-arK).

C was set to 1 for both polynomial and RBF kernels. The gamma hyperparameter was 0.001 for the m-arcsinh-SVMs (both m-ark- and m-arK-SVM), whilst it was set to 1 when using a third-degree polynomial kernel. Gamma was set to ‘auto’, thus automatically determined as 1/number of input features for each of the two datasets

used, when using the RBF kernel function via the scikit-learn library. The gamma hyperparameter was set to the default ‘scale’ in sklearn [20], thus computed as 1/(number of features × variance of the input data array).

The equations for the m-arcsinh kernel function [9] and its derivative are available in (1) and (2) below.

$$\frac{\sqrt{|x|}}{12} \operatorname{arcsinh}(x), \tag{1}$$

where *x* is the input data

$$\sqrt{|x|} \times \frac{1}{12 \times \sqrt{x^2+1}} + \frac{x \times \operatorname{arcsinh}(x)}{24 \times |x|^{\frac{3}{2}}} \tag{2}$$

In **Table I**, the hyperparameters used in the SVM models tested in this study are listed based on the kernel function with which it was coupled, and the margin used.

TABLE I. LEARNING HYPERPARAMETERS FOR THE SVM MODELS BASED ON THE KERNEL FUNCTION AND MARGIN USED.

Kernel	Margin	C	gamma
Linear	Soft	100	‘scale’ in sklearn
Linear	Hard	10	‘scale’ in sklearn
m-arcsinh	Soft	100	0.001
m-arcsinh	Hard	10	0.001
Polynomial	Hard	1	1

RBF	Hard	1	'auto' in sklearn
-----	------	---	-------------------

D. Performance Measures to Validate Model Accuracy and Reliability

The classification performance of the SVM was evaluated quantitatively based on the trade-off between accuracy and reliability on the test sets [28-30] during the stratified k-fold cross validation applied to avoid overfitting or overtraining; the latter metric was assessed via multiple quantifiers, i.e., precision, recall and F1-score as their harmonic mean, thus representing their trade-off. Stratified k-fold cross validation was applied during training, thus providing average metrics across all testing sets for each of the five (k=5) folds [25, 26] the datasets were split into to validate the reliability of the predictions [31-34].

In binary classification problems, e.g., discriminating between healthy and PD-related speech patterns, classes can be defined as 'true' or 'false', resulting in four scenarios (**Table II**):

- true positive (TP), the occurrence of pathophysiological speech patterns correctly detected (e.g., due to PD).
- true negative (TN), physiological/healthy speech patterns correctly classified.
- false positive (FP), physiological/healthy speech patterns misclassified as if they were pathophysiological instead.

- false negative (FN), the occurrence of a pathophysiological speech patterns misclassified as if they were physiological/healthy instead.

TABLE II. CONFUSION MATRICES RELATING THE ACTUAL AND PREDICTED SPEECH PATTERNS AS TO WHETHER THEY ARE INDICATORS OF PATHOPHYSIOLOGICAL SPEECH (E.G., DUE TO PARKINSON'S DISEASE).

		Actual	
		True (pathophysiological speech patterns)	False (physiological speech patterns)
Predicted	Positive (pathophysiological speech patterns)	TP	FP
	Negative (physiological speech patterns)	FN	TN

Besides the average classification accuracy across all folds during cross validation, the classification performance of all variations of the SVM with and without Lagrange multipliers tested were assessed via two further performance metrics as indicators of reliability, i.e., precision and recall (**Tables III and IV**).

As a more comprehensive quantifier of the model reliability, the F1-score was computed too (**Tables III and IV**), which corroborates or disproves the test classification-related outcomes.

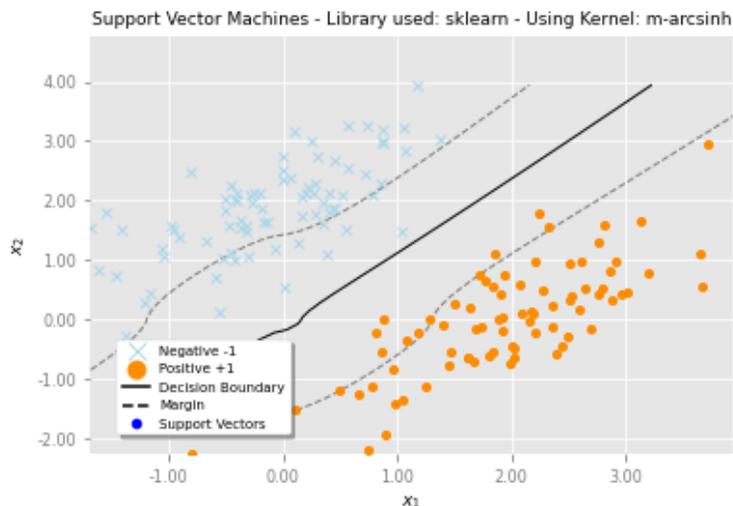


Figure 1. The m-arcsinh kernel function, coupled with a soft margin-SVM, illustrating its non-linear decision boundary and the corresponding margins. As the field 'support_vectors_' is currently empty for custom kernels in scikit-learn (<https://scikit-learn.org/stable/modules/svm.html>), support vectors are not shown, but, for clarity, they would lie along the decision boundary.

III. M-ARCSINH WITH SOFT MARGIN-SVM FOR SUPERVISED MACHINE LEARNING-BASED AUTOMATED DETECTION OF PARKINSON’S DISEASE

A quantitative analysis of classification accuracy and reliability for automated early detection of PD from speech patterns [8] was performed considering two main modelling scenarios (**Tables III** and **IV**):

- When leveraging SVM with various kernels and baseline hyperparameters but without Lagrange multipliers ('SVM').
- When using the above-mentioned SVM but with the addition of Lagrange multipliers ('Lm-SVM').

Both accuracy and reliability metrics have been reported in **Tables III** and **IV**, as well as the overall or average classification performance in percentages.

TABLE III. COMPARISON OF CLASSIFICATION PERFORMANCE AMONGST ALL SVM MODELS EVALUATED ON THE PARKINSON’S DISEASE SPEECH DATASET BY LITTLE ET AL. [18] FROM THE UNIVERSITY OF OXFORD.

Classifier	Kernel	AC C (%)	PR (%)	RE (%)	F1-score (%)	Average performance (%)
SVM	Poly	89.8	89.8	89.8	89.2	89.6
SVM	m-ark	87.2	89.2	87.0	85.0	87.1
SVM	linear	86.6	87.2	86.8	85.6	86.6
SVM	Linear	86.1	87.2	86.4	85.2	86.2
SVM	RBF	75.9	62.0	75.8	65.8	69.9
SVM	m-arK	75.4	56.6	75.2	64.6	68.0
Lm-SVM	m-ark	24.6	6.2	24.8	9.6	16.3
Lm-SVM	m-arK	24.6	6.2	24.8	9.6	16.3
Lm-SVM	Linear	24.6	6.2	24.8	9.6	16.3

Lm-SVM	linear	24.6	6.2	24.8	9.6	16.3
Lm-SVM	Poly	24.6	6.2	24.8	9.6	16.3

SVM=Support Vector Machine; Lm-SVM=SVM with Lagrange multipliers; ACC=average test classification accuracy; PR=precision; RE=recall; Poly=polynomial kernel; m-ark=m-arcsinh kernel with soft margin; m-arK=m-arcsinh kernel with hard margin; linear=linear kernel with soft margin; Linear=linear kernel with hard margin; RBF=Radial Basis Function kernel.

TABLE IV. COMPARISON OF CLASSIFICATION PERFORMANCE AMONGST ALL SVM MODELS EVALUATED ON THE PARKINSON’S DISEASE SPEECH DATASET BY SAKAR ET AL. [19] FROM INSTANBUL UNIVERSITY.

Classifier	Kernel	ACC (%)	PR (%)	RE (%)	F1-score (%)	Average performance (%)
SVM	Poly	86.9	87.0	86.8	86.8	86.9
SVM	m-ark	86.9	86.8	86.8	86.2	86.7
SVM	m-arK	86.4	86.8	86.4	85.2	86.2
SVM	linear	82.8	82.8	82.6	82.8	82.8
SVM	Linear	82.6	82.6	82.4	82.6	82.5
SVM	RBF	74.6	55.6	74.6	63.6	67.1
Lm-SVM	m-ark	25.4	6.4	25.4	10.2	16.9
Lm-SVM	m-arK	25.4	6.4	25.4	10.2	16.9
Lm-SVM	Linear	25.4	6.4	25.4	10.2	16.9
Lm-SVM	linear	25.4	6.4	25.4	10.2	16.9
Lm-SVM	Poly	25.4	6.4	25.4	10.2	16.9

SVM=Support Vector Machine; Lm-SVM=SVM with Lagrange multipliers; ACC=average test classification accuracy; PR=precision; RE=recall; Poly=polynomial kernel; m-ark=m-arcsinh kernel with soft margin; m-arK=m-arcsinh kernel with hard margin; linear=linear kernel with soft margin; Linear=linear kernel with hard margin; RBF=Radial Basis Function kernel.

III. DISCUSSION AND CONCLUSION

The supervised ML-based algorithm alone (SVM) could detect early PD with 89.6% and 86.9% overall classification performance (**Tables III and IV**) when using the polynomial kernel on the PD-related speech datasets from Little *et al.* [18] and Sakar *et al.* [19], respectively. The m-arcsinh [9], when coupled with SVM with soft margin (m-ark), led to the second-best classification performance with 87.1% and 86.7%, correspondingly (**Tables III and IV**). The linear kernel also performed better with the SVM having a soft margin when classifying both datasets (86.6% and 82.8% overall performance, **Tables III and IV**). The use of Lagrange multipliers in SVM [17, 22] led to the worst consistent classification performance across both datasets, resulting in inaccurate OSH-based decision boundaries (16.3% and 16.9% overall classification performance, **Tables III and IV**).

This study validates the use of m-ark when coupled with SVM without Lagrange multipliers for supervised ML-based classification to aid early detection of Parkinson’s Disease from speech features.

Future work is underway to understand if hyperparameter optimisation can level out differences in overall classification performance of the m-ark- and m-arK-based SVMs. Additional training data will be required to minimise the overlapping regions of the transformed data to discriminate between physiological and PD-related speech patterns more reliably.

ACKNOWLEDGMENT

The authors would like to thank the University of Bradford (Faculty of Engineering & Informatics) and the University of Auckland Rehabilitative Technologies Association (UARTA) for giving them the opportunity to develop this collaborative research work. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

REFERENCES

- [1] Bhat, S., Acharya, U. R., Hagiwara, Y., Dadmehr, N., & Adeli, H. (2018). Parkinson's disease: Cause factors, measurable indicators, and early diagnosis. *Computers in biology and medicine*, 102, 234-241.
- [2] Duffy, J. R. (2013). *Motor Speech disorders-E-Book: Substrates, differential diagnosis, and management*. Elsevier Health Sciences.
- [3] Brabenec, L., Mekyska, J., Galaz, Z., & Rektorova, I. (2017). Speech disorders in Parkinson's disease: early diagnostics and effects of medication and brain stimulation. *Journal of neural transmission*, 124(3), 303-334.
- [4] Hlavnička, J., Čmejla, R., Tykalová, T., Šonka, K., Růžička, E., & Ruzs, J. (2017). Automated analysis of connected speech reveals early biomarkers of Parkinson's disease in patients with rapid eye movement sleep behaviour disorder. *Scientific reports*, 7(1), 1-13.
- [5] Adams, W. R. (2017). High-accuracy detection of early Parkinson's Disease using multiple characteristics of finger movement while typing. *PLoS one*, 12(11), e0188226.
- [6] Rees, R. N., Acharya, A. P., Schrag, A., & Noyce, A. J. (2018). An early diagnosis is not the same as a timely diagnosis of Parkinson's disease. *F1000Research*, 7.
- [7] Lahmiri, S., & Shmuel, A. (2019). Detection of Parkinson's disease based on voice patterns ranking and optimized support vector machine. *Biomedical Signal Processing and Control*, 49, 427-433.
- [8] Parisi, L., RaviChandran, N., & Manaog, M. L. (2018). Feature-driven machine learning to improve early diagnosis of Parkinson's disease. *Expert Systems with Applications*, 110, 182-190.
- [9] Parisi, L. (2020). m-arcsinh: An Efficient and Reliable Function for SVM and MLP in scikit-learn. *arXiv preprint arXiv:2009.07530*.
- [10] Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine learning*, 20(3), 273-297.
- [11] Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature*, 323(6088), 533-536.
- [12] Lin, H. T., & Lin, C. J. (2003). A study on sigmoid kernels for SVM and the training of non-PSD kernels by SMO-type methods. *submitted to Neural Computation*, 3(1-32), 16.
- [13] Parisi, L., Biggs, P. R., Whatling, G. M., & Holt, C. A. (2015, July). A novel comparison of artificial intelligence methods for diagnosing knee osteoarthritis. In *XXV congress of the international society of biomechanics* (pp. 1227-1229).
- [14] Parisi, L., & Manaog, M. L. (2016). Preliminary validation of the Lagrangian support vector machine learning classifier as clinical decision-making support tool to aid prediction of prognosis in patients with hepatitis. In *The 16th international conference on biomedical engineering, National University of Singapore (NUS)*.
- [15] Vert, R., Vert, J. P., & Schölkopf, B. (2006). Consistency and Convergence Rates of One-Class SVMs and Related Algorithms. *Journal of Machine Learning Research*, 7(5).
- [16] Vapnik, V. N. (1999). An overview of statistical learning theory. *IEEE transactions on neural networks*, 10(5), 988-999.
- [17] Suykens, J. A., & Vandewalle, J. (1999). Least squares support vector machine classifiers. *Neural processing letters*, 9(3), 293-300.
- [18] Little, M., McSharry, P., Roberts, S., Costello, D., & Moroz, I. (2007). Exploiting nonlinear recurrence and fractal scaling properties for voice disorder detection. *Nature Precedings*, 1-1.
- [19] Sakar, C. O., Serbes, G., Gunduz, A., Tunc, H. C., Nizam, H., Sakar, B. E., ... & Apaydin, H. (2019). A comparative analysis of speech signal processing algorithms for Parkinson's disease classification and the use of the tunable Q-factor wavelet transform. *Applied Soft Computing*, 74, 255-263.
- [20] Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., & Vanderplas, J. (2011). "Scikit-learn: Machine learning in Python." *Journal of machine learning research*, 12 (Oct), 2825–2830.
- [21] Vandenberghe, L. (2010). The CVXOPT linear and quadratic cone program solvers. *Online: <http://cvxopt.org/documentation/coneprog.pdf>*.
- [22] Shao, Y. H., & Deng, N. Y. (2012). A coordinate descent margin based-twin support vector machine for classification. *Neural networks*, 25, 114-121.
- [23] Parisi, L., & Manaog, M. L. (2017). A minimum viable machine learning-based speech processing solution for facilitating early diagnosis of Parkinson's disease. In *MATLAB conference*.
- [24] Wen, Z., Shi, J., Li, Q., He, B., & Chen, J. (2018). ThunderSVM: A fast SVM library on GPUs and CPUs. *The Journal of Machine Learning Research*, 19(1), 797-801
- [25] Kohavi, R. (1995, August). A study of cross-validation and bootstrap for accuracy estimation and model selection. In *IJCAI* (Vol. 14, No. 2, pp. 1137-1145).
- [26] Parisi, L., & Manaog, M. L. (2017). The importance of selecting appropriate k-fold cross-validation and training

algorithms in improving postoperative discharge decision-making via artificial intelligence. In *2017 AUT mathematical sciences symposium* (Vol. 1, No. 1, p. 16).

- [27] Parisi, L., RaviChandran, N., & Manaog, M. L. (2020). A novel hybrid algorithm for aiding prediction of prognosis in patients with hepatitis. *Neural Computing and Applications*, 32(8), 3839-3852.
- [28] Parisi, L., & RaviChandran, N. (2020). Evolutionary Denoising-Based Machine Learning for Detecting Knee Disorders. *Neural Processing Letters*, 52(3), 2565-2581.
- [29] Parisi, L., & RaviChandran, N. (2020). Evolutionary feature transformation to improve prognostic prediction of hepatitis. *Knowledge-Based Systems*, 200, 106012.
- [30] Parisi, L., RaviChandran, N., & Manaog, M. L. (2018). Decision support system to improve postoperative discharge: A novel multi-class classification approach. *Knowledge-Based Systems*, 152, 1-10.
- [31] Parisi, L. (2014). Neural networks for distinguishing the performance of two hip joint implants on the basis of hip implant side and ground reaction force. *Int. J. Medical, Heal. Pharm. Biomed. Eng*, 8(10), 659-663.
- [32] Parisi, L. (2014). Exploiting kinetic and kinematic data to plot cyclograms for managing the rehabilitation process of BKAs by applying neural networks. *Int J Biomed Biol Eng*, 8(10), 664-668.
- [33] Parisi, L., Neagu, D., Ma, R., & Campean, F. (2020). QReLU and m-QReLU: Two novel quantum activation functions to aid medical diagnostics. *arXiv preprint arXiv:2010.08031*.
- [34] Parisi, L., Ma, R., RaviChandran, N., & Lanzillotta, M. (2020). hyper-sinh: An Accurate and Reliable Function from Shallow to Deep Learning in TensorFlow and Keras. *arXiv preprint arXiv:2011.07661*.

Creative Commons Attribution License 4.0 (Attribution 4.0 International, CC BY 4.0)

This article is published under the terms of the Creative Commons Attribution License 4.0
https://creativecommons.org/licenses/by/4.0/deed.en_US