

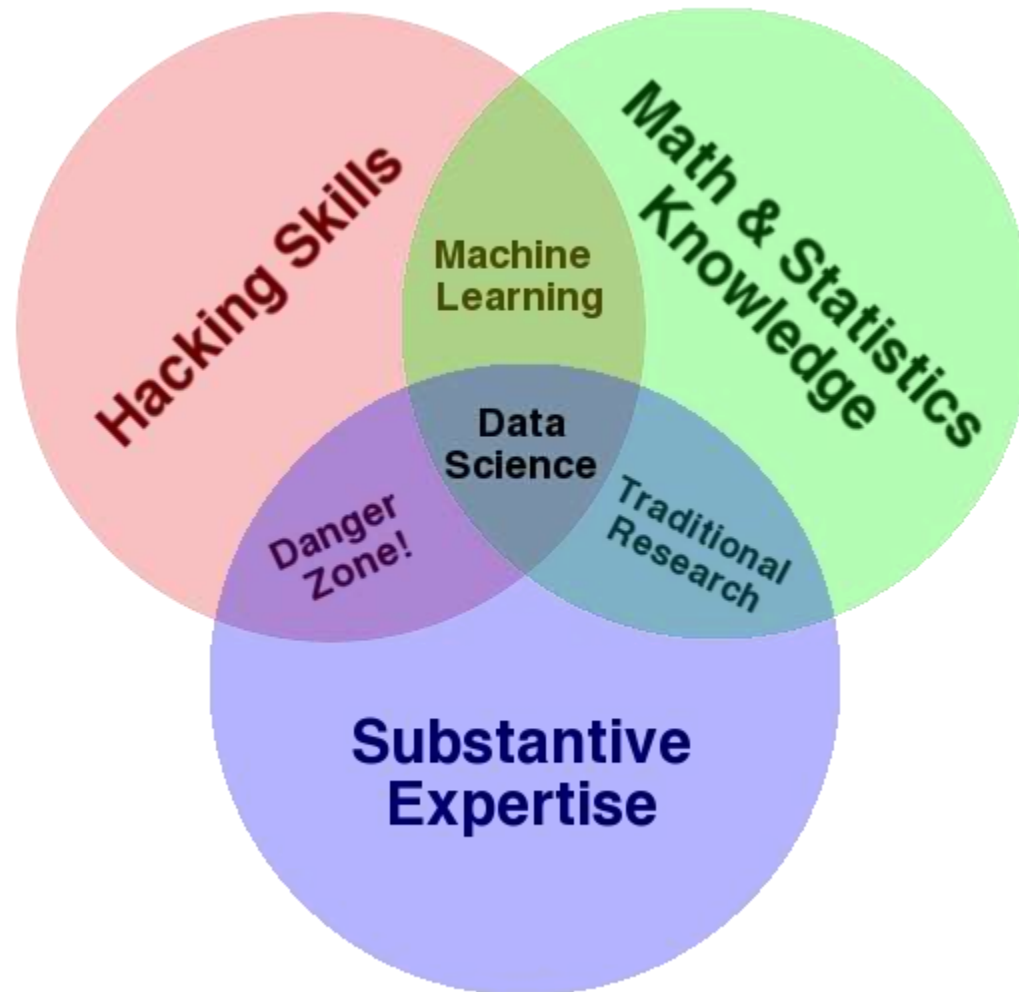
Revolution in Neural Networks: Introducing SPOCU

Milan Stehlík

Institute of Statistics, Universidad de Valparaíso, Valparaíso, Chile
Facultad de Ingenieria, Universidad Andres Bello, Chile



Ciencia de datos: una definición visual



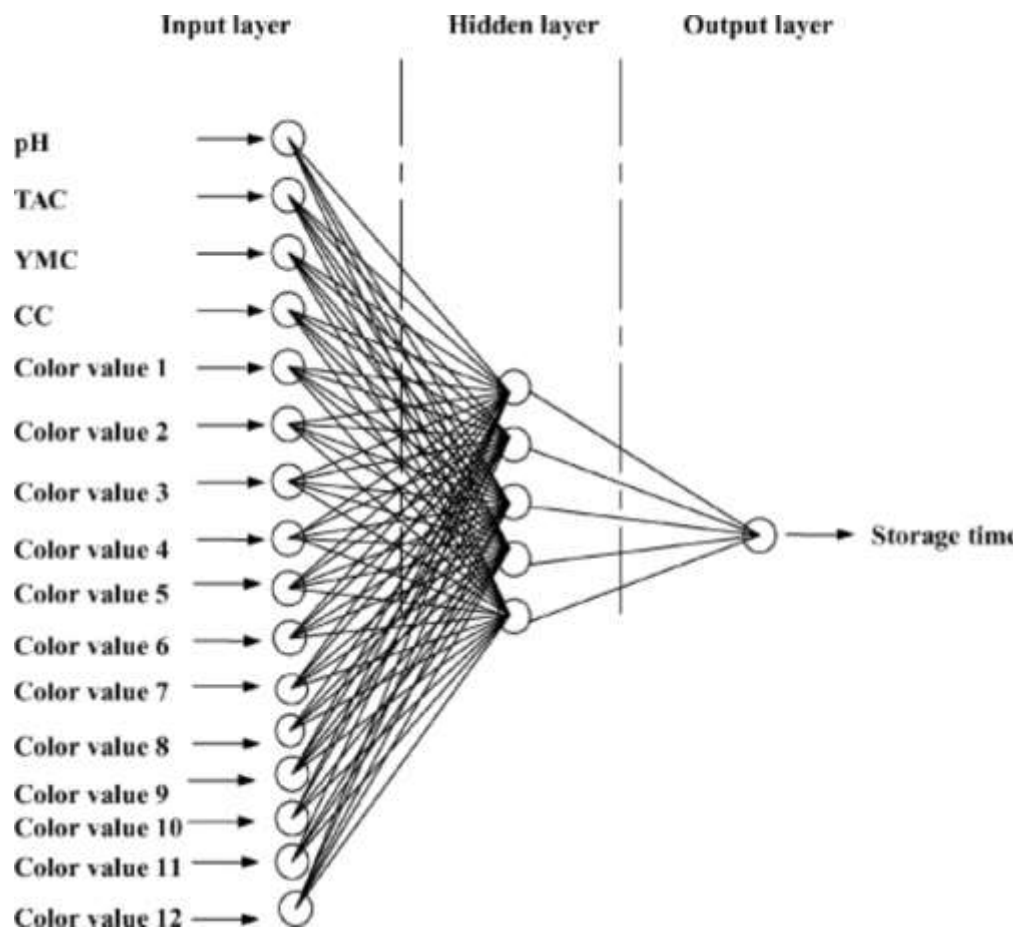
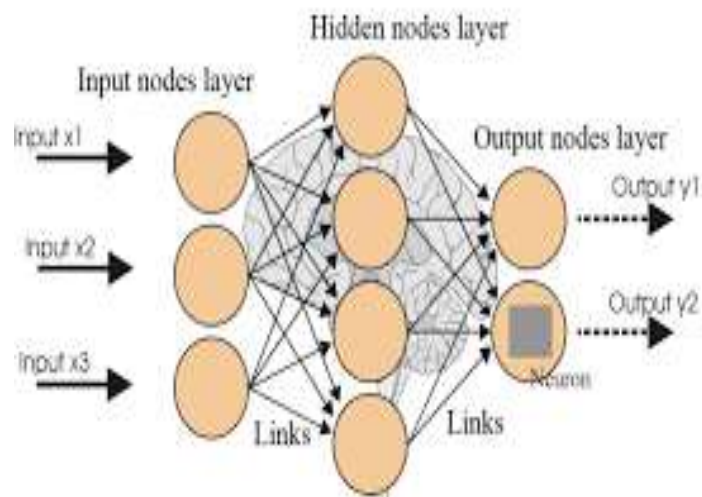
Neural networks and good working nets

- **John 21:6**
- *He told them, “Cast the net on the right side of the boat, and you will find some.” So they cast it there, and they were unable to haul it in because of the great number of fish.*

Activaciones y funciones de transferencia.

- Neurona sigmoidea, utilizada ya en 1972 en neuronas biológicas,
Hugh R. Wilson And Jack D. Cowan
- LIT-2016-1-SEE-023 project “Modeling complex dependencies: how to make strategic multicriterial decisions?/mODEC” **mejorando sigmoid**
- **mejorando RELU, SELU,...**

Lógica de red neuronal



SPOCU función de activación

J. Kiselak, Y. Lu, J. Svihira, P. Szepe, M. Stehlik (2020) **"SPOCU": scaled polynomial constant unit activation function**, Neural Computing and Applications DOI: 10.1007/s00521-020-05182-1

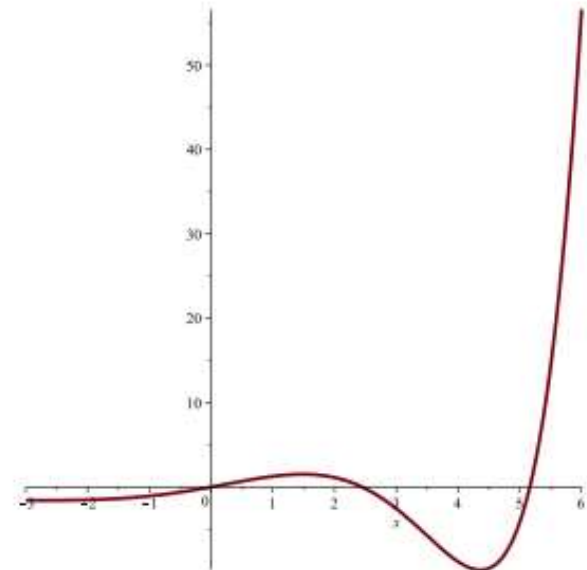
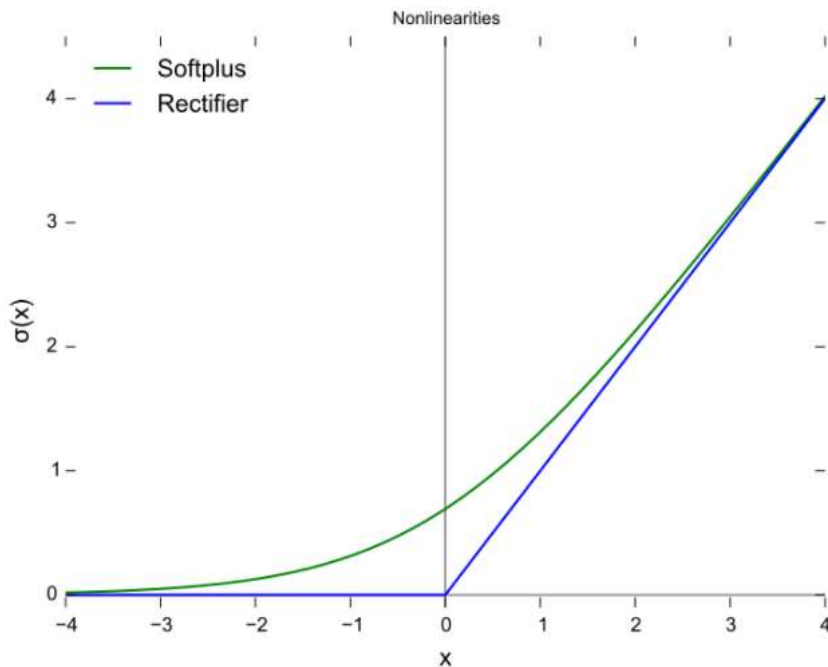


Figure 1: Activation function S

SPOCU función de activación

5. Scaled polynomial constant unit (SPOCU) activation function

Here, we define “scaled polynomial constant unit” (SPOCU) as the novel, well-motivated activation function. We also study its properties. The SPOCU activation function is given by

$$s(x) = \alpha h\left(\frac{x}{\gamma} + \beta\right) - \alpha h(\beta) \quad 11$$

where $\beta \in (0, 1)$, $\alpha, \gamma > 0$ and

$$h(x) = \begin{cases} r(c), & x \geq c, \\ r(x), & x \in [0, c), \\ 0, & x < 0, \end{cases} \quad 12$$

with $r(x) = x^3(x^5 - 2x^4 + 2)$ and $1 \leq c < \infty$. (We admit c goes to infinity with $r(c) \rightarrow \infty$.) Clearly s is continuous $s(0) = 0$ and $s'(x) = \frac{\alpha}{\gamma} h'\left(\frac{x}{\gamma} + \beta\right)$. Notice that for $c = 1$ one has $h'(1^+) = h'(1^-) = 0$ and $h'(0^+) = h'(0^-) = 0$ which implies that s' is continuous too. (This

Spocu Artículo 2019: 5336 Accesos, 72 Citas

Referencias

1. Achter JD, Webb CT (2006) Pair statistics clarify percolation properties of spatially explicit simulations. *Theor Popul Biol*, 69 (2): 155–164, ISSN 0040-5809. <https://doi.org/10.1016/j.tpb.2005.07.003>
2. Bucolo M, Buscarino A, Corradino C, Fortuna L, Frasca M (2019) Turing patterns in the simplest mcnm. *Nonlinear Theory Appl IEICE* 10(4):390–398. <https://doi.org/10.1587/nolta.10.390>
3. Chayes JT, Chayes L, Durrett R (1988) Connectivity properties of mandelbrot's percolation process. *Probab Theory Related Fields*, pp 307–324. <https://doi.org/10.1007/BF00319291> ISSN 1432-2064
4. Cybenko G (1989) Approximation by superpositions of a sigmoidal function. *Math Control Signals Syst* 2(4):303–314. <https://doi.org/10.1007/BF02551274> ISSN 0932-4194; 1435-568X/e
5. Dekking FM, Meester RWJ (1990) On the structure of mandelbrot's percolation process and other random cantor sets. *J Stat Phys* 58(5):1109–1126. <https://doi.org/10.1007/BF01026566> ISSN 1572-9613
6. Falconer K (2013) *Fractal geometry: mathematical foundations and applications*. Wiley. ISBN 9781118762868. URL <https://books.google.at/books?id=XJN7AgAAQBAJ>

SPOCU_scaled_polynomial_const... 1 / 18 100%

Neural Computing and Applications (2021) 33:3385–3401
<https://doi.org/10.1007/s00521-020-05182-1>

ORIGINAL ARTICLE

"SPOCU": scaled polynomial constant unit activation function

Jozef Kiseľák^{1,2} · Ying Lu³ · Ján Švihra⁴ · Peter Szilp⁵ · Milan Stohák^{2,6,7}

Received: 15 February 2019 / Accepted: 11 July 2020 / Published online: 25 July 2020
© The Author(s) 2020

Abstract

We address the following problem: given a set of complex images or a large database, the numerical and computational complexity and quality of approximation for neural network may drastically differ from one activation function to another. A general novel methodology, scaled-polynomial constant unit activation function "SPOCU," is introduced and shown to work satisfactorily on a variety of problems. Moreover, we show that SPOCU can overcome already introduced activation functions with good properties, e.g., SELU and ReLU, on genetic problems. In order to explain the good properties of SPOCU, we provide several theoretical and practical motivations, including tissue growth model and memristive cellular nonlinear networks. We also provide estimation strategy for SPOCU parameters and its relation to generation of random type of Sierpinski carpet, related to the [ppp] model. One of the attractive properties of SPOCU is its genuine normalization of the output of layers. We illustrate SPOCU methodology on cancer discrimination, including mammary and prostate cancer and data from Wisconsin Diagnostic Breast Cancer dataset. Moreover, we compared SPOCU with SELU and ReLU on large dataset MNIST, which justifies usefulness of SPOCU by its very good performance.

Keywords Activation function · SPOCU · SELU · ReLU · Cancer discrimination · Percolation

Mathematics Subject Classification 68C35 · 62B45 · 92B20

1 Introduction

Neural computing and activations in neural networks are interdisciplinary topics, which range from neuroscience to theoretical statistical physics. Also, importantly enough, one of the most important theoretical and practical topics in

neural networks and neural computing is choice of the appropriate activation function. Activation function and its nonlinear ability represent core of the neural networks, both deep and shallow, with various architectures. In standard problems with a reasonable nonlinearity, the sigmoidal function is well justified; however, one faces the inverse problem of specification of the underlying

✉ Milan Stohák
milan.stohak@jku.sk

Jozef Kiseľák
jozef.kiselak@upjs.sk

Ying Lu
ylu1@stanford.edu

¹ School of Medicine, Stanford University, Stanford, USA

² Department of Urology, Jessenius Faculty of Medicine, Comenius University Bratislava, 80609 Martin, Slovak Republic

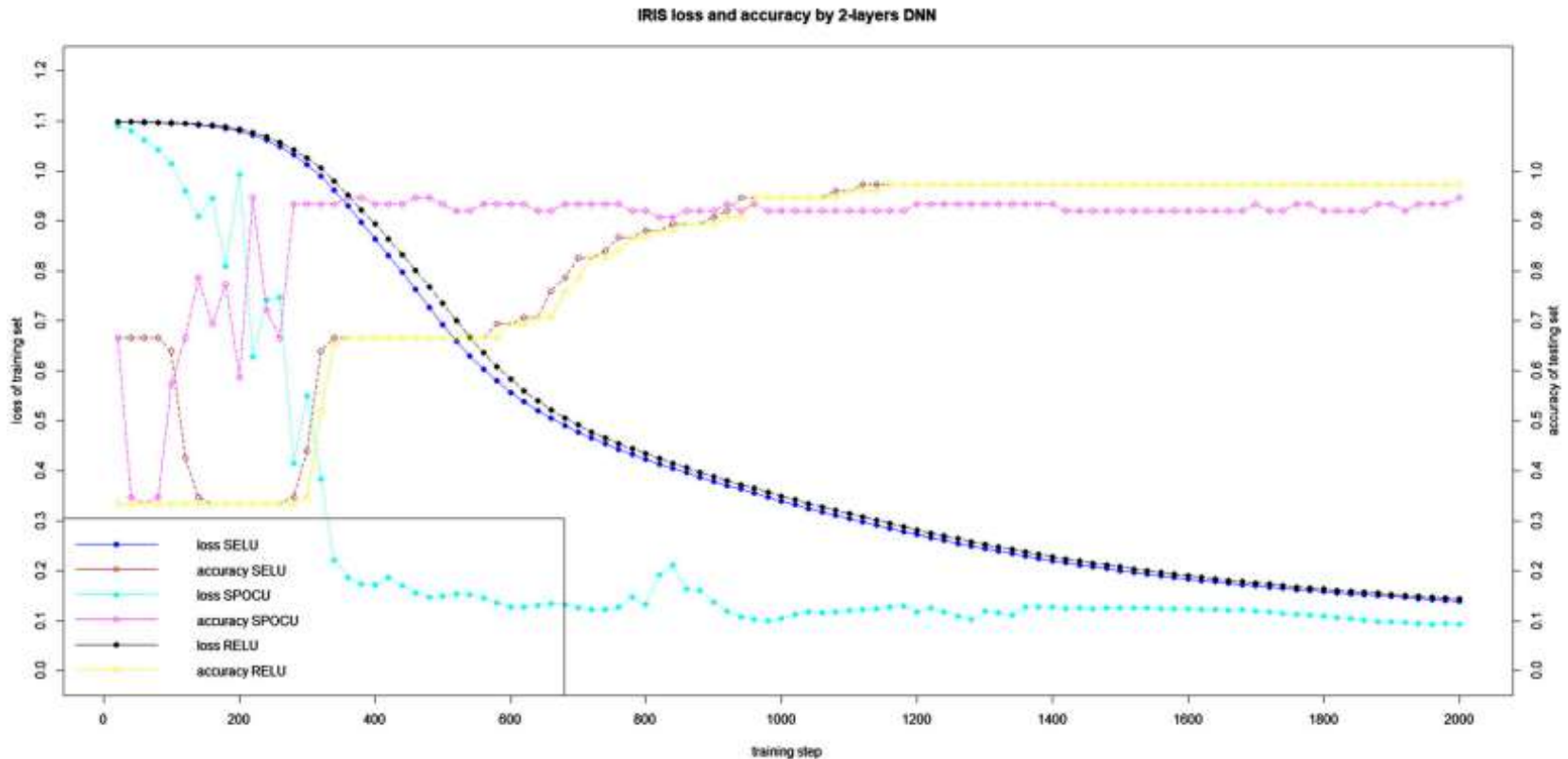
³ Department of Pathological Anatomy, University Hospital Martin, Kolárova 2, 036 39 Martin, Slovak Republic

Citas

1. Ouhamme, S., Hadi, Y. & Ullah, A. An efficient forecasting approach for resource utilization in cloud data center using CNN-LSTM model. *Neural Comput & Applic* **33**, 10043–10055 (2021). <https://doi.org/10.1007/s00521-021-05770-9>
2. Eltrass, A.S., Tayel, M.B. & Ammar, A.I. Automated ECG multi-class classification system based on combining deep learning features with HRV and ECG measures. *Neural Comput & Applic* **34**, 8755–8775 (2022). <https://doi.org/10.1007/s00521-022-06889-z>
3. Johnstone, C., Sulungu, E.D. Application of neural network in prediction of temperature: a review. *Neural Comput & Applic* **33**, 11487–11498 (2021). <https://doi.org/10.1007/s00521-020-05582-3>
4. Khessiba, S., Blaiech, A.G., Ben Khalifa, K. et al. Innovative deep learning models for EEG-based vigilance detection. *Neural Comput & Applic* **33**, 6921–6937 (2021). <https://doi.org/10.1007/s00521-020-05467-5>
5. Dlamini, G., Fahim, M. DGM: a data generative model to improve minority class presence in anomaly detection domain. *Neural Comput & Applic* **33**, 13635–13646 (2021). <https://doi.org/10.1007/s00521-021-05993-w>
6. Owolabi, K.M., Baleanu, D. Emergent patterns in diffusive Turing-like systems with fractional-order operator. *Neural Comput & Applic* **33**, 12703–12720 (2021). <https://doi.org/10.1007/s00521-021-05917-8>

SPOCU función de activación

SPOCU ha mostrado resultados mucho mejores que Relu y SELU en la clasificación de imágenes de cáncer.



SPOCU función de activación

Act. func.	Range	Smooth.	Nonlin.	Monot.	Monot. deriv.	Id. near 0	SNN
SPOCU	$[-\alpha r(\beta), \alpha(1 - r(\beta))]$	C^1	Yes	Yes	No	Yes	Yes
ReLU	$[0, \infty)$	C^0	Piecewise lin.	Yes	Yes	No	No
SELU	$(-\lambda\alpha, \infty)$	C^0	Yes	Yes	No	No	Yes

Experimentos:

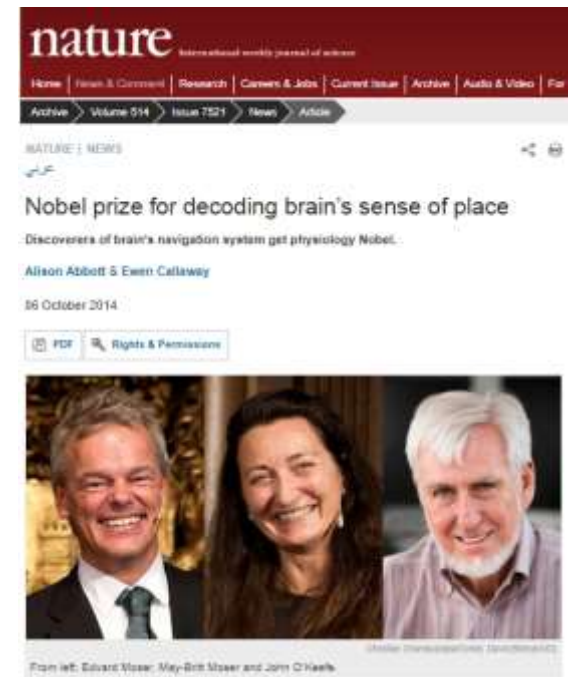
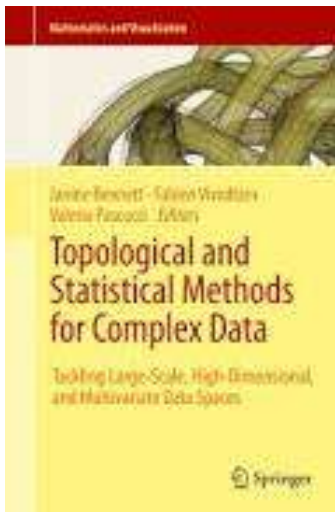
- en conjunto con los estudiantes: León Benjamín, Bastian Barazza, Castro Fierro Felipe; Ronald Ecklmaier, ...
- Mejora del modelo de bajo rendimiento, SNN (con Ronald Ecklmair, Bastian Barazza)
- LSTMs (Memoria larga a corto plazo): con Leon Benjamin
- -códigos: Keras Tensor Flow, fuente disponible para la función de activación de SPOCU, excepto que se ha publicado una implementación de SPOCU para Pytorch para la comunidad en un repositorio de Github.<https://github.com/pacocp/spocu-pytorch>
- In literature there are e.g. configurations to use spocu in both layers, and specifically when we used for the LSTM layer
:Recurrent_Activation=spocu, recurrent_regularizer=L1, Kernel_regularizer=L1, Bias_regularizer=L2, there was an initial moment where the loss takes values, slowly decreasing until before the first epoch ends in NAN, (gradient death).

prerequisites

base 1

Combinar análisis de datos topológicos y estadísticos: ¡un camino hacia el éxito!

- La extracción de información de conjuntos de datos de gran dimensión, incompletos y ruidosos suele ser un desafío.
- Muy útil para el procesamiento de big data.
- Neuro-computación



base 2

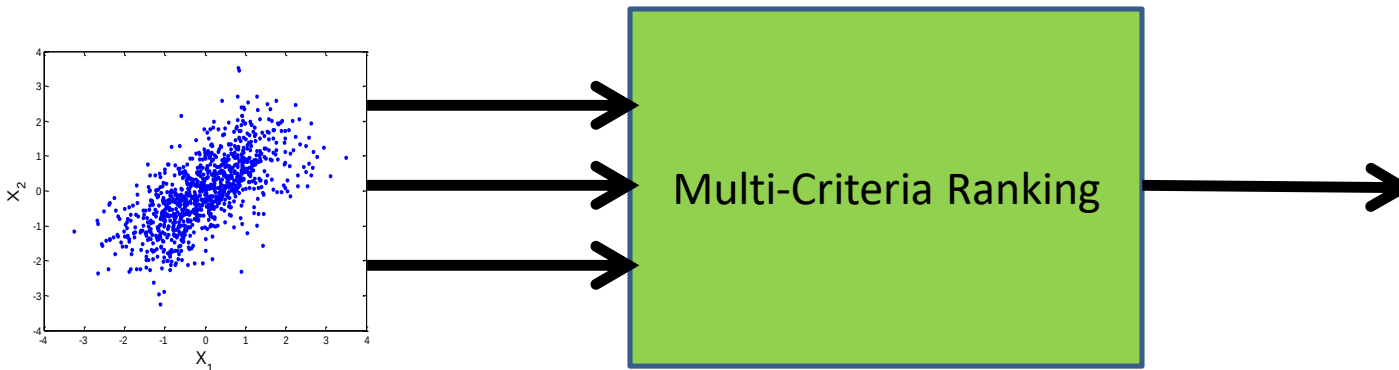
Aprendizaje por agregación

p.Ej. Agregación de probabilidades

La teoría de la información y las técnicas de aprendizaje automático agregan muchas entradas de probabilidad de expertos y proporcionan una salida única

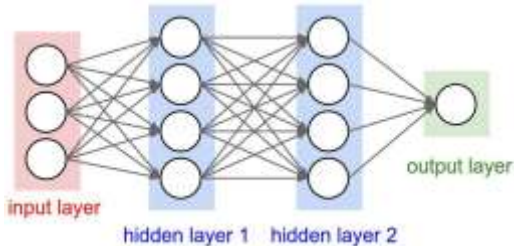
Ejemplos: sistema de decisión multicriterio - fusión de varias preferencias en conflicto en una sola decisión

- componentes estadísticos, estocásticos, deterministas y caóticos, p. PSO, ANN, híbrido NN



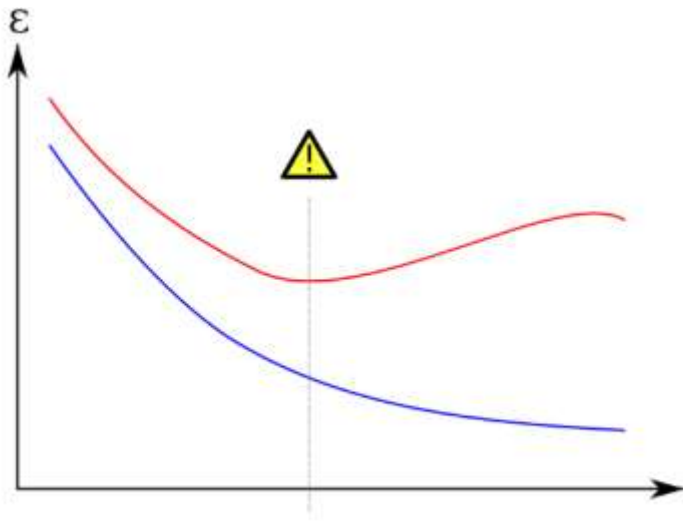
base 3

Sobre las necesidades estadísticas urgentes del aprendizaje automático



A veces, hay una gran diferencia entre el error ocurrido en el conjunto de datos de entrenamiento y el error encontrado en un nuevo conjunto de datos no visto.

SUV eléctrico Tesla se estrelló contra una barrera en U.S. Autopista 101 en Mountain View, California. 23/03/2018 murió el conductor



Training error in blue, Validation error in red (Overfitting) as a function of the number of cycles. Credits:



base 4

Robustez y optimización de hiperparámetros

- Los hiperparámetros son los parámetros cuyo valor se define antes del comienzo del proceso de aprendizaje. Cambiar el valor de dichos parámetros por una pequeña cantidad puede invocar un gran cambio en el rendimiento de su modo
- [LIT proyectos:](#)
- [Modeling Complex Dependencies: How to Make Strategic Multi-Criterial Decisions](#)
- [Autonomous Driving in Austria \(AUD/A\) – Predictions by, and a challenge for Criminal Law](#)
- [Long Short-Term Memory For Self-Driving Cars.](#)

base 4

Robustez y optimización de hiperparámetros

DEXPSO

A Double Exponential Particle Swarm Optimization

- Desarrollamos un novedoso algoritmo basado en enjambres utilizando la distribución exponencial doble.

Demostramos que DExPSO puede converger al óptimo global y mejorar con respecto a PSO.

Comparamos el rendimiento de DExPSO con varios algoritmos basados en enjambres.

Proporcionamos aplicaciones de DExPSO a problemas de diseño en bioestadística.

Reconstruimos los conocidos resultados de "Convergencia - Búsqueda Global".



M. Stehlik, Weng Kee, Ping-Yang, Kiselak, (2024) A Double Exponential Particle Swarm Optimization with non-uniform variates as stochastic tuning and guaranteed convergence to a global optimum with sample applications to finding optimal exact designs in biostatistics, Applied Soft Computing

base 4

Robustez y optimización de hiperparámetros

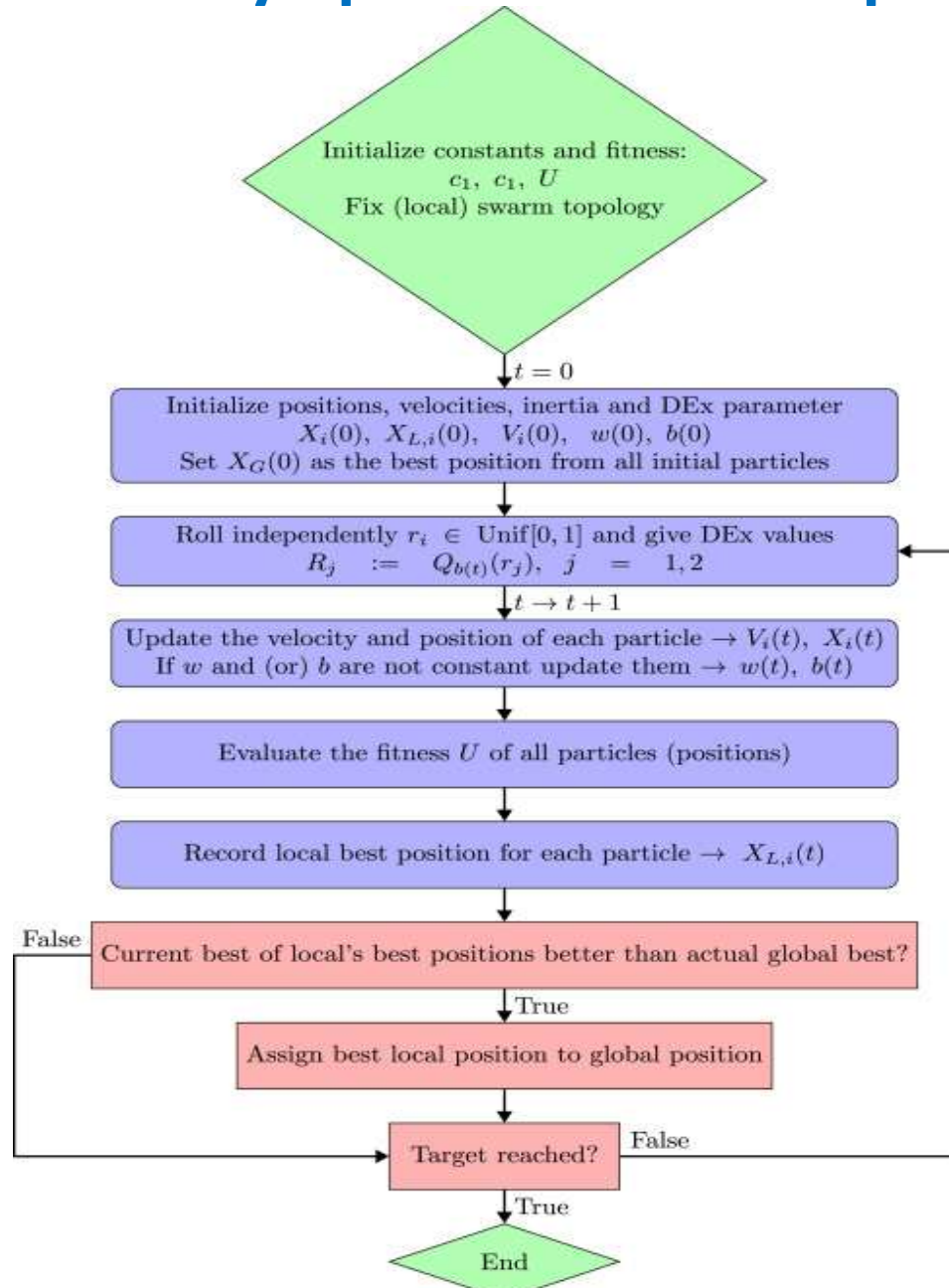
Desventajas de PSO

- - 1) Tendencia a convergencia rápida en puntos medios óptimos.
 - 2) Convergencia lenta en la búsqueda refinada.

Ventajas de DEXPSO

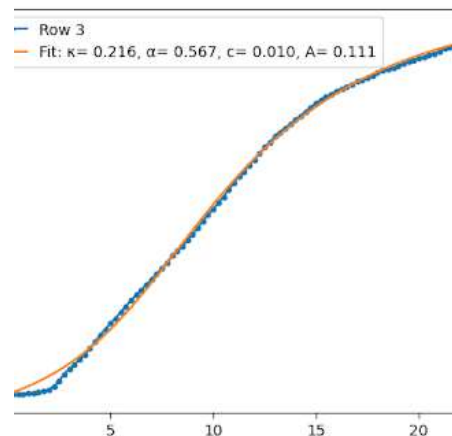
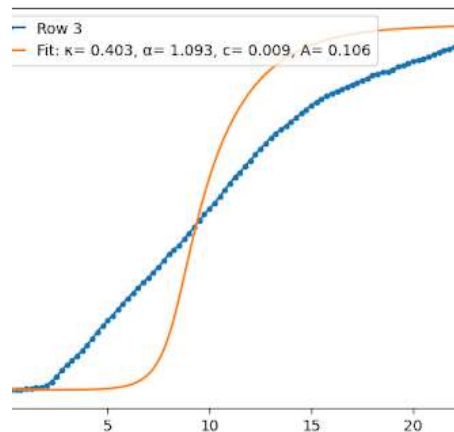
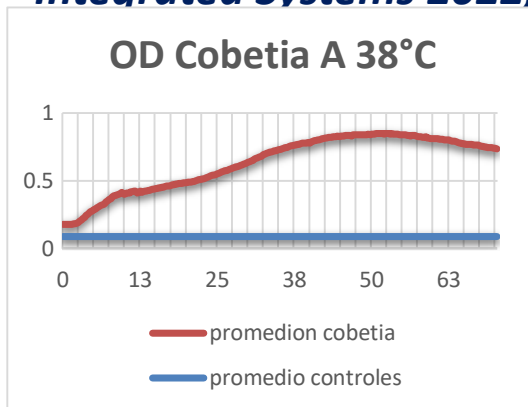
- a) Mejora sustancial de las desventajas 1 y 2 del PSO
 - b) Insensible a la escala de las variables de diseño
- 3) Muy pocos parámetros de algoritmo
 - 4) Algoritmo de búsqueda global muy eficiente

base 4: Robustez y optimización de hiperparámetros



¿Por qué Dexpso y SPOCU? Soluciones adaptativas para los desafíos de la sensibilidad de los parámetros en el modelado del crecimiento de Cobetia Marina. (A.Dinamarca, M.Stehlik, C. Ibacache, J. Kiselak, K. Gonzales, I. Plaza, B. Barazza, B. Leon)

- 2023-2026, co-Director, proyecto Exploracion ANID GRT-AC: A Global Gene Regulation Based Tool For Modelling The Next Generation Of Biological-Silicon Integrated Systems 2022, Nr. 13220184



Base 5

Requisitos para hardware de alto rendimiento

- Formación y conjunto de datos para una solución de Deep Learning

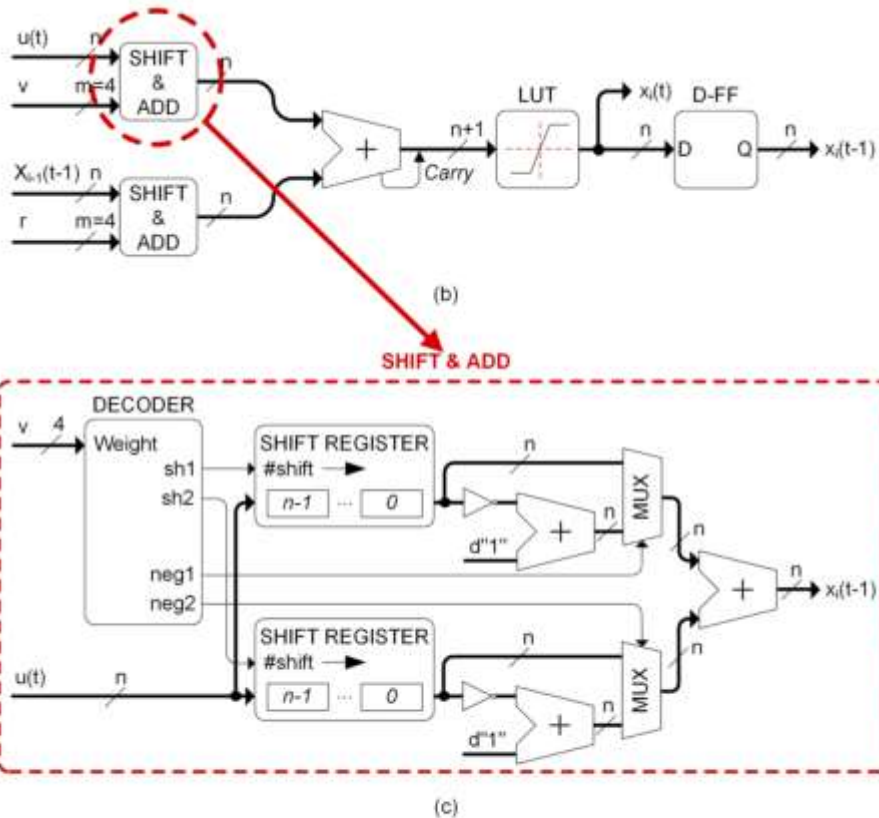
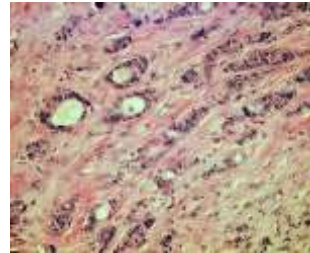
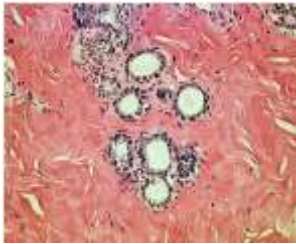


Fig. 2 (a-c) Neuron design: (a) general circuit design of the neuron; (b) reduced implementation scheme when the weight resolution is limited to few bits ($m = 4$) and the multipliers are replaced by simple shift-and-add blocks; (c) description of the shift-and-add block

Base 6

ANÁLISIS DE DATOS TOPOLÓGICOS

Ejemplo: LA DISCRIMINACIÓN DEL CÁNCER



Pruebas y discriminación cáncer de mama-mastopatía

La mastopatía diabética son lesiones no cancerosas en la mama que se diagnostican con mayor frecuencia en mujeres premenopáusicas con diabetes tipo 1. (Keyoung, 2009)

Stehlik et al. Fractal and Stochastic Geometry Inference for Breast Cancer: Case Study with Modelo de fractales aleatorios y proceso de interacción de Quermass, Statistics in Medicine, 34, 2636-2661

solo microscopios utilizados durante mucho tiempo. (Mattfeldt, UniClinic Ulm, Alemania)
Vidakovic et al. índices de discriminación del 50-55% ("ligeramente mejor que lanzar una moneda")

Stehlik et al (2016) y las tasas de clasificación correcta de SVM son más del 80%

Podemos usar topología de datos: Quermass Procesos de interacción

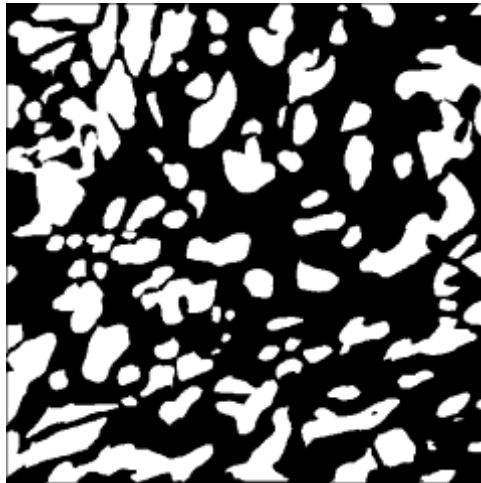
Problemas de discriminación: bajos índices de clasificaciones correctas



- Sólo microscopios utilizados desde hace mucho tiempo. (Mattfeldt, UniClinic Ulm, Alemania)
- Vidakovic et al. **tasas de discriminación** del 50 al 55% ("un poco mejor que lanzar una moneda al aire")
- Stehlik et al 2016: las tasas de clasificación correcta de SVM son superiores al 80 %
- La tecnología de metilación del ADN ya está disponible...
- Podemos utilizar topología: Procesos de interacción Quermass

FracLac

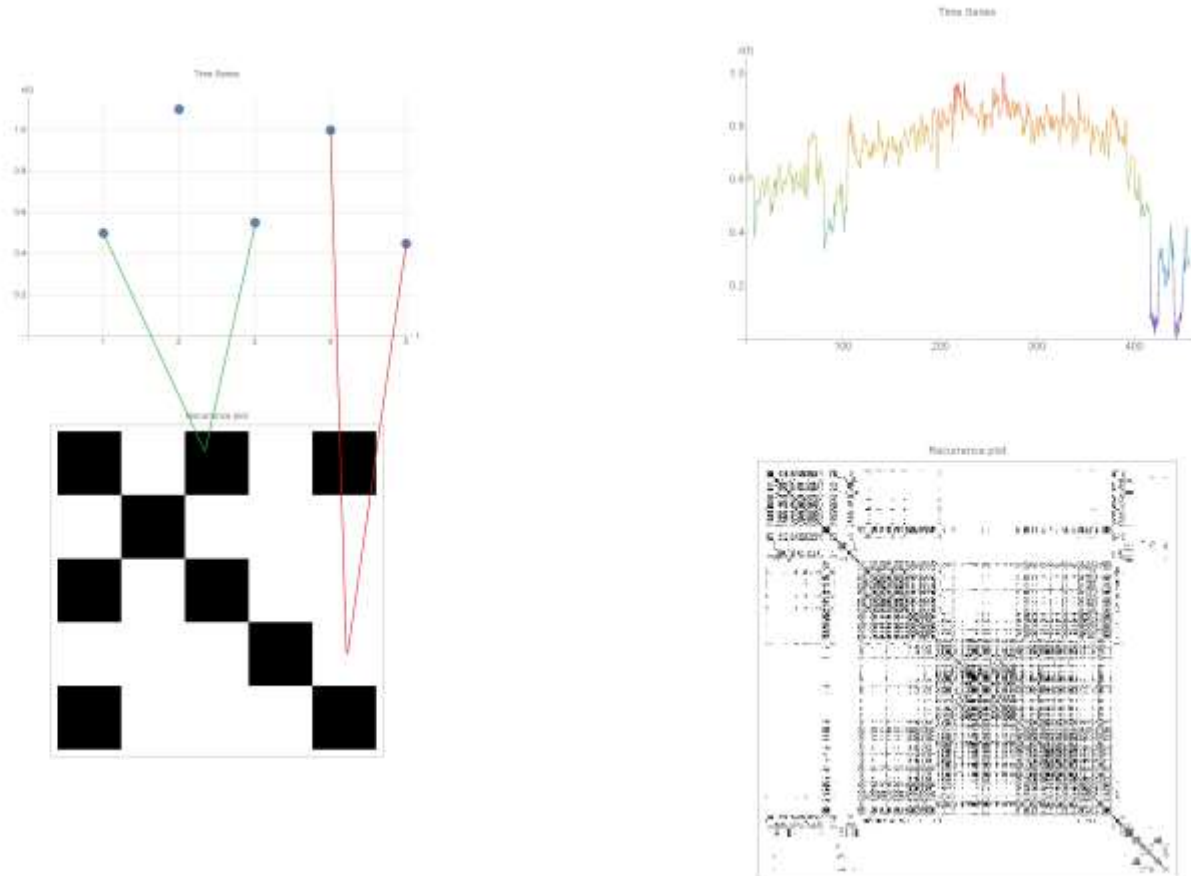
- Análisis de tejido mamario



FracLac Box Count Results

File	Edit	Font			
FracLac 2.5 Release 1e: file/slice/roi		Foreground pixels	Total pixels	Mean D	
mamca.tifS1 (0,0:512x512) 1,1--511,511		93880	262144	1,715	
masto.tifS1 (0,0:512x512) 309,1--511,374		17125	262144	1,4963	

Stehlik et al 2018, 2021, RECURRENCE ANALYSIS OF METHANE EMISSIONS FROM WETLAND ECOSYSTEM, *AIP, Cambridge Un. Press*,



Superioridad especial de SPOCU

- SPOCU puede superar funciones de activación como SELU y ReLU en problemas genéricos **con una mejora del 30%**.
- La selección óptima de la función de activación y la gestión de hiperparámetros **son cruciales**. La función SPOCU ofrece flexibilidad y un rendimiento superior en tareas de aprendizaje automático (Babimore et al., 2021, Vives et al., 2021 o Mesellem et al., 2021, etc.).
- Los resultados experimentales muestran mejoras en el diagnóstico de cáncer y la dinámica de adsorción de contaminantes. **Las funciones de transferencia estándar presentan limitaciones en configuraciones complejas.**

Superioridad especial de SPOCU

- Ilustramos la metodología SPOCU sobre discriminación por cáncer, incluido el cáncer de mama y de próstata y datos del conjunto de datos de Wisconsin Diagnostic Breast Cancer, **Menos falsos positivos y falsos negativos.**
- Además, comparamos SPOCU con SELU y ReLU en un gran conjunto de datos MNIST, lo **que justifica la utilidad de SPOCU por su muy buen rendimiento (complejidad y tiempo computacional)**

Research in Statistics

NEW OPEN ACCESS JOURNAL LAUNCHING 2022

Research in Statistics is a broad open access journal publishing original research in all areas of statistics and probability. The journal focuses on broadening existing research fields, and in facilitating international collaboration, and is devoted to the international advancement of the theory and application of advanced statistical techniques. This includes both theoretical and official statistics, statistics for financial data, and administrative statistics. The journal welcomes original and significant contributions in the area of statistics whether on empirical or mathematical foundations, or their applications to any domain of statistical and information technology, and more generally to any field of investigation where statistical methodologies are relevant. Applied papers demonstrating the usefulness of statistical methodology in practical problems are particularly welcome. Papers submitted for consideration should provide novel contributions in statistical theory or applications as they relate to the following interdisciplinary topics:

- Statistical Theory
- Statistical Modeling
- Medical Statistics and Epidemiology
- Biostatistics
- Agricultural Statistics
- Environmental Statistics
- Physical Sciences Statistics
- Engineering Statistics
- Extremes and its applications
- Risk Theory
- Behavioural and Psychological Statistics
- Economics & Business Statistics
- Statistics Education
- Statistical Surveys
- Statistical Computing, Data Visualization, and Simulation
- Social Sciences
- Demography
- History of Statistics
- Theory of Fuzzy Measures, Fuzzy Probability and Information and its Statistical Applications
- Randomness, Stochastic Processes, and Statistical Applications

Editor-in-Chief: Milan Stehlik

Institute of Applied Statistics, Johannes Kepler University Linz, Austria
Instituto de Estadística, Universidad de Valparaíso, Chile
milan.stehlik@jku.at



Taylor & Francis Group
an informa business

lanzamiento 6.10.2022

Política de intercambio de datos

Data Sharing Policy

Esta revista aplica la Política de intercambio de datos básicos de Taylor & Francis. Se alienta a los autores a compartir o abrir los datos que respaldan los resultados o análisis presentados en su artículo cuando esto no viole la protección de los sujetos humanos u otras preocupaciones válidas de privacidad o seguridad.

Creemos que los datos deben ser Encontrables, Accesibles, Interoperables y Reproducibles (FAIR).

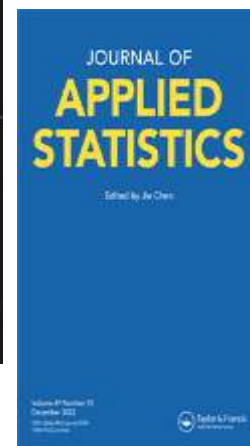
Hacer que sus datos estén disponibles puede aumentar la visibilidad de su investigación, aumentando su impacto y asegurando que obtenga crédito por todo su trabajo. También mejora la solidez del proceso de investigación, apoyando la validación, la transparencia de la investigación, la reproducibilidad y la



Estadísticas en Taylor & Francis



Al ofrecer libros y revistas sobre **recopilación, análisis, visualización e interpretación de datos cuantitativos y cualitativos**, reconocemos el papel clave que desempeña la estadística en una variedad de disciplinas científicas. Nuestro objetivo es promover su importancia a través de una proliferación de recursos y asociaciones con sociedades como la American Statistical Association en beneficio de investigadores, estudiantes y profesionales.



Muchas Gracias!

References

Bamimore, A., Sobowale, N.B., Osunleke, A.S. et al. Offset-free neural network-based nonlinear model predictive controller design using parameter adaptation. *Neural Comput and Applic* 33, 10235-10257 (2021).

Vives-Boix, V., Ruiz-Fernandez, D. Fundamentals of artificial metaplasticity in radial basis function networks for breast cancer classification. *Neural Comput and Applic* 33, 12869-12880 (2021).

Mesellem, Y., Hadj, A.A.E., Laidi, M. et al. Computational intelligence techniques for modeling of dynamic adsorption of organic pollutants on activated carbon. *Neural Comput and Applic* 33, 12493-12512 (2021).

Kiselak, J.; Lu, Y.; Svihra, P.; Szepe, M. ; Stehlik, M. "SPOCU": scaled polynomial constant unit activation function, *Neural Comput and Applic* 33, (2020) 3385-3401.

Stehlík, M. , Ping-Yang, Ch. Wong, WK, Kiselaak, J. (2024) A Double Exponential Particle Swarm Optimization with non-uniform variates as stochastic tuning and guaranteed convergence to a global optimum with sample applications to finding optimal exact designs in biostatistics, *Applied Soft Computing*, <https://doi.org/10.1016/j.asoc.2024.111913>

Dinamarca,A, Stehlik, M. et al. (2025) Comprehensive and deep learning classification for analyses of biological complexity of growth and biofilms of *Cobetia marina* under different temperature growths. *PLOS ONE*, in revision.