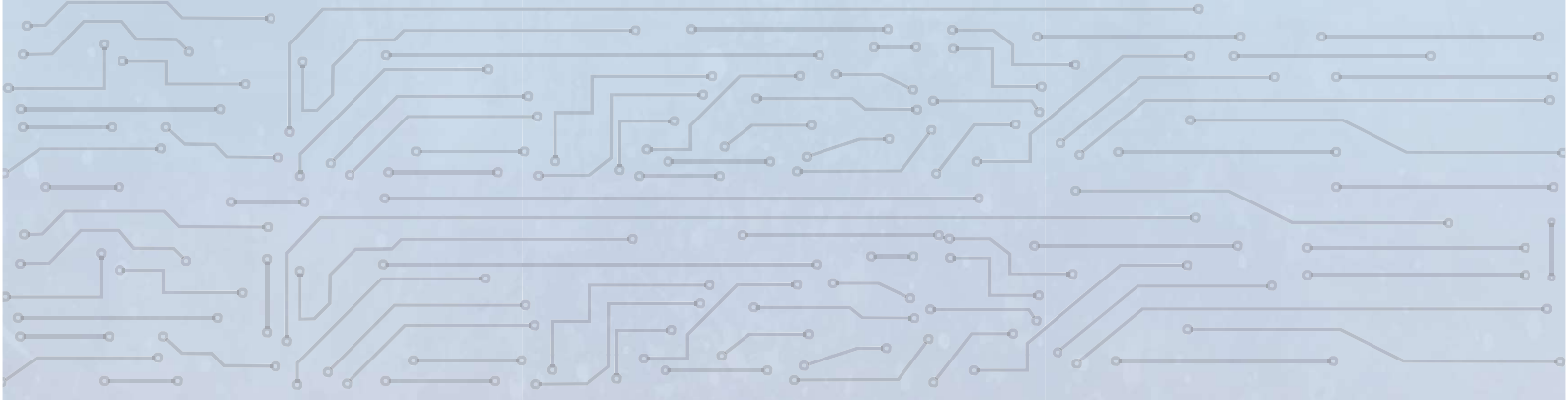




ACTAS DE LA XVII CONFERENCIA DE LA ASOCIACIÓN ESPAÑOLA PARA LA INTELIGENCIA ARTIFICIAL

*Oscar Luaces; Francisco Herrera; José A. Gámez; Luis Martínez;
Edurne Barrenechea; José Riquelme; Alicia Troncoso;
Bruno Baruque; Mikel Galar; Héctor Quintián;
Emilio Corchado (Eds.)*

CAEPIA'16
Salamanca

MAEB
Salamanca

LODISCO
Salamanca

TAMIDA
Salamanca

FINO
Salamanca

Ediciones Universidad

Salamanca

AQUILAFUENTE

224

©

Ediciones Universidad de Salamanca y
de cada autor

Motivo de cubierta:
Diseñadora María Alonso Miguel

1.º edición: septiembre, 2016
ISBN: 978-84-9012-632-5 (PDF)

Ediciones Universidad de Salamanca
www.eusal.es
eusal@usal.es

Realizado en España – Made in Spain

*Todos los derechos reservados.
Ni la totalidad ni parte de este libro
pueden reproducirse ni transmitirse sin permiso escrito de
Ediciones Universidad de Salamanca*

Obra sometida a proceso de
evaluación mediante sistema de revisión por pares a ciegas
a tenor de las normas del congreso

Ediciones Universidad de Salamanca es miembro de la UNE
Unión de Editoriales Universitarias Españolas
www.une.es

CEP

VIII Simposio Teoría y Aplicaciones de Minería de Datos TAMIDA 2016

Co-Presidentes de TAMIDA 2016

- Alicia Troncoso Lora, Universidad Pablo de Olavide de Sevilla
- José C. Riquelme, Universidad de Sevilla

Comité Científico

- Alicia Troncoso Lora, Universidad Pablo de Olavide
- José C. Riquelme, Universidad de Sevilla
- Emilio Corchado, Universidad de Salamanca
- Francesc J. Ferri, Universidad de Valencia
- Francisco Herrera, Universidad de Granada
- Héctor Pomares, Universidad de Granada
- Ignacio Rojas, Universidad de Granada
- José A. Lozano, Universidad del País Vasco
- José del Campo, Universidad de Málaga
- Juan J. Rodríguez, Universidad de Burgos
- José Hernández Orallo, Universidad Politécnica de Valencia
- Juan J. del Coz, Universidad de Oviedo
- Karina Gibert, Universidad Politécnica de Cataluña
- Marta Zorrilla, Universidad de Cantabria
- Miquel Sànchez i Marrè, Universidad Politécnica de Cataluña
- M^a José del Jesús, Universidad de Jaén
- M^a José Ramírez, Universidad Politécnica de Valencia
- Nicolás García Pedrajas, Universidad de Córdoba
- Olatz Arbelaiz, Universidad del País Vasco
- Óscar Cerdón, Universidad de Granada
- Pablo Varona, Universidad Autónoma de Madrid
- Pedro Larrañaga, Universidad Politécnica de Madrid
- Sebastián Ventura, Universidad de Córdoba
- Juan Nepomuceno, Universidad de Sevilla
- Francisco Martínez Álvarez, Universidad Pablo de Olavide
- Gualberto Asencio, Universidad Pablo de Olavide
- María Martínez Ballesteros, Universidad de Sevilla
- Marta Arias, Universidad Politécnica de Cataluña
- Jesús Aguilar, Universidad Pablo de Olavide
- Cristóbal Carmona, Universidad de Burgos

Un Estudio sobre el Preprocesamiento para Redes Neuronales Profundas y Aplicación sobre Reconocimiento de Dígitos Manuscritos	
DANIEL PERALTA, ANDRÉS HERRERA-POYATOS, FRANCISCO HERRERA	867
Instance selection for regression by discretization	
ÁLVAR ARNAIZ-GONZÁLEZ, JOSÉ F. DIEZ-PASTOR, CÉSAR GARCA-OSORIO, JUAN J.RODRÍGUEZ	877
Instance Selection for Regression: adapting DROP	
ÁLVAR ARNAIZ-GONZÁLEZ, JOSÉ F. DIEZ-PASTOR, CÉSAR GARCA-OSORIO, JUAN J.RODRÍGUEZ	879
Tutorial on practical tips of the most in uential data preprocessing algorithms in data mining	
SALVADOR GARCÍA, JULIÁN LUENGO, FRANCISCO HERRERA	881

Aprendizaje no supervisado:

Mining the Dataflow in Interactive Systems: Automatic Generation of User Behavior Patterns	
RAFAEL DUQUE, JOSÉ L. MONTAÑA, CRISTINA TÎRNĂUCĂ	885
Análisis visual de técnicas no supervisadas de <i>deep learning</i> con el paquete dlvisR	
DAVID CHARTE, FRANCISCO CHARTE, FRANCISCO HERRERA	895

Parte V.- III Jornadas de Fusión de la Información y ensembles (FINO 2016)

Combinando diferentes tipos de información en la clasicación de imágenes	
JUAN I. FORCEN, MIGUEL PAGOLA, EDURNE BARRENECHEA, MIKEL GALAR	907
Intercambio de Clases de acuerdo a la Distancia al Enemigo más Cercano para Problemas con Clases altamente No Balanceadas	
SERGIO GONZÁLEZ, MARCELINO LÁZARO, ANIBAL R. FIGUEIRAS-VIDAL, SALVADOR GARCÍA, FRANCISCO HERRERA	917
CHI-BD: Sistema de Clasificación Basado en Reglas Difusas para Problemas de Clasificación Big Data	
MIKEL ELKANO, MIKEL GALAR, JOSÉ SANZ, HUMBERTO BUSTINCE	927
Ordering-Based Pruning for Improving the Performance of Ensembles of Classiers in the Framework of Imbalanced Datasets	
A. FERNÁNDEZ, M. GALAR, E. BARRENECHEA, H. BUSTINCE, F. HERRERA	937

Parte VI.- Doctoral Consortium

Data mining techniques for time series processing Application to industrial sound classification	
JAVIER ROMERO, AMALIA LUQUE, ALEJANDRO CARRASCO	941
Classification from imbalanced datasets. A framework for improving the application of sampling strategies	
MOHAMED S. KRAIEM	947
Reducción de datos aplicada a Big Data	
SERGIO RAMÍREZ GALLEGÓ	953
Detección de objetos en secuencias de video mediante modelos estadísticos y neuronales	
MIGUEL A. MOLINA-CABELLO, EZEQUIEL LÓPEZ-RUBIO, RAFAEL MARCOS LUQUE-BAENA	959
Semantic Computer Vision tools for mission-critical visual communications over bandwidth limited networks	
LAURA LÓPEZ FUENTES	965

Parte VII.- Concurso de Apps

PROA: una Aplicación Android para el Diseño de Rutas Personalizadas	
MARINA TORRES, DAVID A. PELTA, JOSÉ L. VERDEGAY	973
Competición CAEPIA-APP: DOWNTOWN-APP	
INMACULADA AYALA, LORENZO MANDOW, MERCEDES AMOR, LIDIA FUENTES	981

Análisis visual de técnicas no supervisadas de *deep learning* con el paquete *dlvisR*

David Charte, Francisco Charte, and Francisco Herrera

Universidad de Granada, Granada, España

`fdavidcl@correo.ugr.es`, `francisco@fcharte.com`, `herrera@ugr.es`

Resumen Las técnicas de *deep learning* aplicadas al aprendizaje no supervisado han demostrado su utilidad y potencial, pero carecen del nivel de interpretabilidad que pueden proporcionar otros algoritmos. Además, el ajuste de los parámetros de funcionamiento de este tipo de métodos suele realizarse de forma automática, y no se obtienen explicaciones de cómo influyen en el comportamiento de los modelos y los resultados que estos ofrecen. En este trabajo se presenta una herramienta desarrollada para la plataforma R, el paquete *dlvisR*. Este proporciona un conjunto de utilidades para la visualización de las variables obtenidas internamente por este tipo de modelos respecto de parámetros ajustables por el usuario. Además, un estudio sobre algunos conjuntos de datos reafirma la hipótesis de que la modificación de dichos parámetros tiene consecuencias observables visualmente, cuyo análisis podría aportar conocimiento de interés.

Keywords: deep learning · aprendizaje no supervisado · R · software

1. Introducción

El aprendizaje no supervisado abarca multitud de problemas ampliamente estudiados que tienen diversas aplicaciones presentes en distintos campos, como el tratamiento de imágenes y reconocimiento de objetos [1], análisis semántico [2] y sintáctico del lenguaje [3] o el preprocesamiento de datos y pre-entrenamiento para una posterior fase de aprendizaje [4].

Una familia de técnicas que se pueden utilizar para realizar estas tareas son las de *deep learning* (DL), capaces de construir modelos complejos para los datos a partir de numerosas pequeñas representaciones sencillas. Se pueden englobar por tanto dentro del conjunto de herramientas de *representation learning*, pero se diferencian del resto de estas en que son capaces de aprender varias capas de representaciones, cada una en base a la anterior [5,6]. En los últimos años, gracias al aumento de la capacidad de procesamiento disponible, de la calidad del software abierto dedicado a DL y del volumen de datos publicados para su tratamiento, se ha producido un resurgimiento de este tipo de técnicas, especialmente desde la *Deep Belief Network* [7].